

STEVEN G. MURRAY

NEXT-GENERATION TOOLS FOR NEXT-GENERATION SURVEYS



# NEXT-GENERATION TOOLS FOR NEXT-GENERATION SURVEYS

STEVEN G. MURRAY

Presented for the degree of Doctor of Philosophy,  
School of Physics  
University of Western Australia

October 2015

SUPERVISORS:

Chris Power

Aaron Robotham



THE UNIVERSITY OF  
WESTERN AUSTRALIA

*Achieve International Excellence*



---

## DECLARATION

---

This thesis is my own work, and no part of it has been submitted for a degree at this, or at any other, university.

*Perth, October 2015*

---

Steven G. Murray



To Emily, Joshua and Caleb.  
May you never be forced to read this.





---

## ABSTRACT

---

The next generation of large-scale galaxy surveys, across the electromagnetic spectrum, loom on the horizon as explosively game-changing datasets, in terms of our understanding of cosmology and structure formation. We are on the brink of a torrent of data that is set to both confirm and constrain current theories to an unprecedented level, and potentially overturn many of our conceptions.

One of the great challenges of this forthcoming deluge is to extract maximal scientific content from the vast array of raw data. This challenge requires not only well-understood and robust physical models, but a commensurate network of software implementations with which to efficiently apply them.

The halo model, a semi-analytic treatment of cosmological spatial statistics down to nonlinear scales, provides an excellent mathematical framework for exploring the nature of dark matter. This thesis presents a next-generation toolkit based on the halo model formalism, intended to fulfil the requirements of next-generation surveys.

Our toolkit comprises three tools: (i) `HMF`, a comprehensive and flexible calculator for halo mass functions (HMFs) within extended Press-Schechter theory, (ii) the MRP distribution for extremely efficient analytic characterisation of HMFs, and (iii) `HALOMOD`, an extension of `HMF` which provides support for the full range of halo model components.

In addition to the development and technical presentation of these tools, we apply each to the task of physical modelling. With `HMF`, we determine the precision of our knowledge of the HMF, due to uncertainty in our knowledge of the cosmological parameters, over the past decade of *cosmic microwave background* (CMB) experiments. We place rule-of-thumb uncertainties on the predicted HMF for the Planck cosmology, and find that current limits on the precision are driven by modeling uncertainties rather than those from cosmological parameters.

With the MRP, we create and test a method for robustly fitting the HMF to observed masses with arbitrary measurement uncertainties on a per-object basis. We find that our method reduces estimation uncertainty on parameters by over 50%, and correctly accounts for Eddington bias even in extremely poorly measured data. Additionally, we use the analytical properties of the MRP to obtain asymptotically correct forms for the stellar-mass halo-mass relation, in the subhalo abundance matching scheme.

Finally, with `HALOMOD`, we explore the viability of the halo model as a test of warm dark matter (WDM) via galaxy clustering. Examining three distinct scale regimes, we find that the clustering of galaxies at the smallest resolvable scales may provide a valuable independent probe in the coming era.



---

## PUBLICATIONS

---

This thesis is constructed as a “series of papers” in compliance with the rules for PhD thesis submission from the graduate research school at the University of Western Australia.

Two chapters have resulted in publications, and the remaining chapters are presented in a form suitable for submission.

Publications arising from this thesis:

- **HMFCalc: An online tool for calculating dark matter halo mass functions.**

S. G. Murray, C. Power and A. S. G. Robotham

A & C 3-4, 23-24, 2013

arXiv: 1306.6721

⇒ S. G. M. contribution:  $\sim 80\%$

- **How well do we know the halo mass function?**

S. G. Murray, C. Power and A. S. G. Robotham

MNRAS 434, L61-L65, 2013

arXiv: 1306.5140

⇒ S. G. M. contribution:  $\sim 80\%$

---

Chris Power

---

Aaron Robotham



---

## ACKNOWLEDGMENTS

---

I presume it is often the case that one finishes a work of such a significant magnitude as this knowing that a host of people, organisations and circumstances were utterly instrumental in bringing about its completion, but also knowing that the true extent of their support is beyond knowing. Certainly, that is the case for me. I guess I may take solace in the hope that this is neither the conclusion nor pinnacle of my life's work, and that in future works I might more fully repay the generosity of all those who have stood with and behind me.

I must thank my two primary supervisors, Chris Power and Aaron Robotham, who have put up with a seemingly endless list of questions, requests, blank stares and mis-treatments of little *h* throughout the years. I could not have had a pair of more supportive and complementary supervisors: Chris, your big-picture outlook has carried the project, and Aaron, your attention to detail has gotten it done. Thank you.

Thanks also goes to all the students and staff at ICRAR who have made this time such a positive experience. Special thanks to the DIA team who helped enormously with getting my first web-application off the ground.

As far as collaborations go, I have had the privilege of sharing a great side-project with Santi Avila, whose company and discussion have been invigorating, and whose translations in the Spanish cafeteria potentially life-saving. Also, I cannot miss those who laboured through a plethora of back-and-forwards emails in attempts to tie down problems: Chris Blake, David Palamara, Florian Beutler and others. Thank you!

Of course, I would not even be thinking about completing a Ph.D. thesis without the avid and enduring support of my parents. You guys are beyond amazing. Thank you for everything you sacrificed throughout the years, and the way you taught me to view life, and love understanding.

This Ph.D. has formed about 74% of my married life, and that time has also seen three ridiculously awesome children spring into existence. I am forever grateful to my beautiful wife, Jeanette, whose unfailing love and support has given me the best of times, and plucked me out of the worst. Words here could not possibly convey the sheer brilliance with which you go about life. I love you, and here's to the life beyond!

Finally, there is Another who cannot be ignored. Without Him, all of this loses its substance.

Soli Deo Gloria.



---

## CONTENTS

---

1	INTRODUCTION	1
1.1	General Introduction	1
1.1.1	Historical perspective	1
1.1.2	The state of affairs	4
1.1.3	On the brink	7
1.2	Structure Formation	8
1.2.1	The Cosmological Expansion	8
1.2.2	The seeds of structure	10
1.2.3	Linear Growth	11
1.2.4	Press-Schechter Formalism	12
1.2.5	Extended PS	15
1.2.6	WDM	16
1.3	Observations	17
1.3.1	Cluster surveys	17
1.3.2	Galaxy redshift surveys	19
1.4	Next-generation tools	21
2	hmfcalc : AN ONLINE TOOL FOR CALCULATING DARK MATTER HALO MASS FUNCTIONS	23
2.1	Introduction	23
2.2	The Halo Mass Function (HMF)	24
2.2.1	Fitting Functions	25
2.2.2	Redshift Dependence	27
2.2.3	Warm Dark Matter Models	28
2.3	Implementation	29
2.3.1	HMFcalc's engine, hmf	30
2.3.2	HMFcalc's web interface and the Django Framework	34
2.3.3	Usage	35
2.3.4	Comparison To Other Codes	38
2.4	Example Applications	39
2.4.1	Box Size for One Halo of Mass M	39
2.4.2	The Impact of Finite Box Size	43
2.5	Future Development	45
2.5.1	Extending Range of Dark Matter and Dark Energy Models	45
2.5.2	Calibrating Synthetic Galaxy Surveys	45
2.5.3	Dynamic and Adaptable User Interface	46
2.6	Summary	46
3	HOW WELL DO WE KNOW THE HALO MASS FUNCTION?	51
3.1	Introduction	51

3.2	Methodology . . . . .	52
3.3	Results . . . . .	54
3.4	Summary . . . . .	60
4	AN EMPIRICAL MASS FUNCTION DISTRIBUTION	63
4.1	Introduction . . . . .	63
4.2	The MRP Model . . . . .	66
4.2.1	Cumulative Distribution . . . . .	66
4.2.2	Normalisation . . . . .	67
4.2.3	Re-parameterisations . . . . .	67
4.2.4	Comparison to EPS fits . . . . .	68
4.3	Fitting MRP . . . . .	69
4.3.1	Basic framework . . . . .	70
4.3.2	Common Issues . . . . .	70
4.3.3	Fitting to simulations . . . . .	72
4.3.4	MRP in the presence of baryons . . . . .	73
4.3.5	Fitting halos with measurement uncertainty . . . . .	80
4.4	Dependence on Physical Parameters . . . . .	84
4.5	Stellar-Mass Halo-Mass Relation . . . . .	89
4.6	Conclusion . . . . .	94
5	halomod : A NEW AND FLEXIBLE HALO MODEL CALCULATOR	97
5.1	Introduction . . . . .	97
5.2	The Halo Model . . . . .	100
5.2.1	Clustering Framework . . . . .	101
5.2.2	HOD models . . . . .	103
5.2.3	Extension of framework to galaxies . . . . .	105
5.2.4	Derived Quantities . . . . .	106
5.2.5	Matter Power Spectrum . . . . .	106
5.2.6	The Mass Function . . . . .	108
5.2.7	Bias . . . . .	109
5.2.8	Halo Density Profiles . . . . .	112
5.2.9	Concentration-Mass Relation . . . . .	114
5.2.10	Halo exclusion . . . . .	116
5.3	The HALOMOD library . . . . .	118
5.3.1	Overview . . . . .	119
5.3.2	Frameworks . . . . .	123
5.3.3	Components . . . . .	126
5.3.4	Extra functionality . . . . .	141
5.4	Example Application . . . . .	143
5.4.1	Setup and methodology . . . . .	143
5.4.2	Results . . . . .	143
5.5	Future Development . . . . .	147
5.6	Summary . . . . .	151



6	CONNECTING GALAXIES TO WARM DARK MATTER HALOS: CLUSTERING SIGNATURES	153
6.1	Introduction	153
6.2	Methods	155
6.2.1	The halo model formalism	155
6.2.2	WDM halo model	158
6.2.3	Test Statistics	165
6.3	WDM Signatures in Varying Regimes	168
6.3.1	Large-scale 2-halo term	168
6.3.2	Small-scale 1-halo term	169
6.3.3	Transition scales	175
6.4	Results	177
6.4.1	Model I: NFW	177
6.4.2	Model II: Einasto	178
6.4.3	Analysis of WAVES	180
6.5	Conclusions	181
6.5.1	Summary of results	181
6.5.2	Future work	182
7	SUMMARY AND CONCLUSIONS	185
7.1	Summary	185
7.2	Conclusions	188
7.2.1	Next-generation tools	188
7.2.2	Next-generation surveys	189
	BIBLIOGRAPHY	193
	<b>Appendix</b>	<b>205</b>
A	MRP SUPPLEMENTS	205
A.1	Recurrence Relation	205
A.2	Re-parameterisations	205
A.3	Truncation Mass Dependence and Extensions	208
A.4	Jacobians and Hessians of MRP	209
A.4.1	Method I: Binned data	209
A.4.2	Method II: Per-object data	210
A.4.3	Derivatives of $g$ and $q$	210
A.5	An Idealised Analytic Model	212
A.5.1	Development of Model	212
A.5.2	Analysis of Simplest Model	215
A.5.3	Addition of Noise and Systematic Bias	215
A.6	stan code	220
B	halomod SUPPLEMENTS	223
B.1	Hankel Transform	223

B.2	Projected Correlation Function Limits . . . . .	225
-----	---	-----

---

## LIST OF FIGURES

---

Figure 1.1	CfA ‘Stick-man’ . . . . .	4
Figure 1.2	Schematic of the Press-Schechter peaks argument . . . . .	13
Figure 1.3	Illustration of the integral to determine the fractional volume of collapsed matter in PS theory. . . . .	14
Figure 1.4	Depth of several completed and future surveys . . . . .	20
Figure 2.1	Large selection of halo mass functions at redshift zero . . . . .	27
Figure 2.2	Projected dark matter density in slices of a CDM and WDM simulation . . . . .	28
Figure 2.3	HMF for CDM and three WDM models . . . . .	29
Figure 2.4	Window function integration limits schematic . . . . .	32
Figure 2.5	HMFCALC web-application flow-chart . . . . .	36
Figure 2.6	Box size required to form at least one halo of mass $> M$ . . . . .	41
Figure 2.7	Number of halos expected in simulation of size $L^3$ . . . . .	42
Figure 2.8	Effect of finite box-size . . . . .	44
Figure 3.1	Several HMF fitting functions in the Planck cosmology . . . . .	54
Figure 3.2	Errors on the amplitude and slope of the HMF for a range of cosmologies . . . . .	55
Figure 3.3	HMFs with uncertainty regions . . . . .	56
Figure 3.4	Variation of HMF fits with simulation parameters . . . . .	58
Figure 3.5	Contribution of single parameters to the variance in the HMF . . . . .	59
Figure 3.6	WDM HMFs with error region overplotted . . . . .	60
Figure 4.1	Comparison of MRP to Tinker fitting function . . . . .	68
Figure 4.2	Residuals of MRP fit to $\nu^2$ GC simulations . . . . .	74
Figure 4.3	Joint posteriors of MRP parameters estimated simultaneously on four $\nu^2$ GC simulations . . . . .	75
Figure 4.4	Joint posterior of MRP parameters estimated over OWLS simulations . . . . .	76
Figure 4.5	Relative variation between baryon models’ HMFs in cosmo-OWLS . . . . .	77
Figure 4.6	MRP parameters determined from the various cosmo-OWLS simulations . . . . .	79
Figure 4.7	Diagram of estimated masses from algorithm with arbitrary uncertainty . . . . .	82
Figure 4.8	Recovered HMFs from best-fit MRP, for $S = 0.8$ mass uncertainties . . . . .	83
Figure 4.9	Joint posteriors for $S = 0.2$ case . . . . .	83
Figure 4.10	Summary of all methods to estimate MRP parameters from uncertain masses . . . . .	85
Figure 4.11	Logarithmic mass mode $\mathcal{H}_T$ as a function of redshift. . . . .	86
Figure 4.12	Independent physical dependence of MRP parameters . . . . .	88

Figure 4.13	Estimated uncertainty in parameterisations for the physical dependence of the MRP parameters . . . . .	89
Figure 4.14	Schematic of stellar mass function . . . . .	91
Figure 4.15	Numerically calculated SMHM with 3 parameterisations from the literature . . . . .	92
Figure 5.1	Halo occupation for several parameterisations . . . . .	105
Figure 5.2	Comparison of transfer function models . . . . .	108
Figure 5.3	Several HMF fits . . . . .	109
Figure 5.4	Effect of the HMF fit on the two-point correlation function . . . . .	110
Figure 5.5	Several bias functions from the literature . . . . .	112
Figure 5.6	Effect of halo exclusion, including several models . . . . .	118
Figure 5.7	Comparison of transfer function models, normalised at large scales	127
Figure 5.8	Mass variance using three different filters . . . . .	130
Figure 5.9	Concentration-mass-redshift relation for various models in the literature . . . . .	139
Figure 5.10	Marginalised Joint-posteriors for the “basic” run . . . . .	144
Figure 5.11	Residuals for all fits in example application . . . . .	145
Figure 5.12	Marginalised joint-posteriors for a run with cosmology free to vary with uniform priors . . . . .	148
Figure 5.13	Marginalised joint-posteriors for a run with cosmology free to vary with covariant normal priors . . . . .	149
Figure 5.14	Marginalised joint-posteriors for the HOD parameters with the inclusion of both covariant cosmology and nuisance parameters from the HMF, bias and concentration . . . . .	150
Figure 6.1	Comparison of $z = 0$ CDM and WDM power spectra . . . . .	160
Figure 6.2	Mass variance in CDM and WDM using tophat and sharp- $k$ window functions . . . . .	161
Figure 6.3	HMFs for CDM and WDM using tophat and sharp- $k$ window functions . . . . .	162
Figure 6.4	Concentration-mass relations for CDM and WDM . . . . .	165
Figure 6.5	BAO peak structure in CDM and WDM . . . . .	169
Figure 6.6	Ratio of BAO peak-to-trough distances in CDM vs. WDM . . . . .	170
Figure 6.7	Effective concentration for CDM and WDM . . . . .	171
Figure 6.8	Schematic of 1-halo integral . . . . .	172
Figure 6.9	Schematic of 1-halo correlations in five regions . . . . .	173
Figure 6.10	Dependence of the correlation ratio on HOD parameters . . . . .	174
Figure 6.11	Effective bias in WDM vs. CDM . . . . .	175
Figure 6.12	Transition scale correlation ratios between WDM and CDM . . . . .	176
Figure 6.13	Effect of scale-dependent bias and halo exclusion on transition scale correlation ratios . . . . .	177
Figure 6.14	Maximum deviation from WDM of best-fit CDM models in 1-halo and transition scales for NFW haloes . . . . .	178
Figure 6.15	Deviation at $r_{p,\min}$ from WDM of best-fit CDM models in 1-halo and transition scales for NFW haloes . . . . .	179

Figure 6.16	Maximum deviation from WDM of best-fit CDM models in 1-halo and transition scales for Einasto haloes . . . . .	179
Figure 6.17	Deviation at $r_{p,\min}$ from WDM of best-fit CDM models in 1-halo and transition scales for Einasto haloes . . . . .	180
Figure 7.1	Analytics for HMFALC over the past two years. . . . .	190
Figure 7.2	Comparison of current and future extra-galactic spectroscopic facilities . . . . .	191
Figure 7.3	Comparison of current and future extra-galactic redshift surveys	192
Figure A.1	MRP re-parameterisations . . . . .	207
Figure A.2	Dependence of MRP parameters on truncation mass . . . . .	209
Figure A.3	Analytic expectation of uncertainty in MRP parameters . . . . .	216
Figure A.4	Analytic expectation of parameter correlations . . . . .	217
Figure A.5	Expected variance from simple sampled model . . . . .	218
Figure A.6	Expected parameter bias if parameters are ideal . . . . .	219
Figure B.1	Analysis of appropriate values of $h$ in Hankel transform . . . . .	225

---

## LIST OF TABLES

---

Table 2.1	Description of Run Parameters in HMFALC . . . . .	37
Table 2.2	Description of Cosmological Parameters in HMFALC . . . . .	38
Table 2.3	Compilation of HMF fitting functions . . . . .	47
Table 4.1	Basic properties and derived quantities of MRP . . . . .	67
Table 4.2	Details of New Numerical Galaxy Catalogue . . . . .	72
Table 4.3	Table showing formulae for the physical dependence of the MRP parameters . . . . .	87
Table 4.4	Summary of SMHM relations in the literature . . . . .	91
Table 4.5	Best-fit parameters for the SMHM . . . . .	94
Table 5.1	Summary of parameters included in HaloModel . . . . .	124
Table 5.2	All included properties of HaloModel . . . . .	125
Table 5.3	Extra parameters and properties introduced in ProjectedCF . . .	126
Table 5.4	API summary of the Transfer component. . . . .	128
Table 5.5	Summary of included Transfer models. . . . .	128
Table 5.6	API summary of the Filter component. . . . .	130
Table 5.7	Summary of included Filter models . . . . .	131
Table 5.8	API summary of the FittingFunction component. . . . .	131
Table 5.9	API summary of the Bias component. . . . .	132
Table 5.10	Summary of included Bias models . . . . .	133
Table 5.11	API summary of the Profile component . . . . .	136
Table 5.12	Summary of included Profile models . . . . .	137
Table 5.13	API summary of the CMRelation component. . . . .	139

Table 5.14	Summary of concentration-mass-redshift relations implemented in HALOMOD. . . . .	139
Table 5.15	API summary of the HOD component. . . . .	140
Table 5.16	Summary of included HOD parameterisations . . . . .	140
Table 5.17	Summary of base model for MCMC runs in example application	145
Table 5.18	Summary of the variable parameters used in MCMC runs . . . .	146
Table A.1	MRP re-parameterisations . . . . .	206

---

## LISTINGS

---

Listing A.1	Stan code for fitting MRP to masses with arbitrary measurement uncertainties . . . . .	220
-------------	--	-----

---

## INTRODUCTION

---

We are on a precipice. It has been a century since the astonishing discovery of the general theory of relativity, and the unprecedented unveiling of the Universe that followed has offered a breathtaking glimpse into its very structure. But on the horizon, looming ever closer, is the promise of a grand new revolution. This revolution will pour over us at the rate of terabytes per second, and transform our understanding of the way that fluctuations at the smallest imaginable scales in the infant Universe could give rise to the magnificent galactic structures of today. To harness this potential, we need a corresponding paradigm-shift in our toolset – software and methods applicable to the biggest questions ever asked.

### 1.1 GENERAL INTRODUCTION

#### 1.1.1 *Historical perspective*

In the last century, physical cosmology has seen something akin to the Copernican revolution. Up until the beginning of the previous century, astronomers predominantly believed in a static universe containing just our own Galaxy, obeying Newtonian gravitation. In quick succession, these notions were challenged and overthrown.

In 1915, Einstein published his *general theory of relativity* (Einstein, 1915; Einstein, 1916), changing the way we think about the fundamental force of gravity, which has a significant impact in cosmology. While this ground-breaking work was being done, astronomers still thought it inconceivable that the Universe extended beyond our own Milky Way. This paradigm began to be challenged, climaxing in what is known as the ‘Great Debate’ in 1920 between Harlow Shapley (in favour of just the Milky Way) and Heber Curtis (in favour of extra-galactic nebulae) (Curtis, 1921; Shapley, 1921), before finally breaking when Slipher and Hubble spectacularly showed the existence of extragalactic objects by their redshifts (Slipher, 1915; Hubble, 1929). Not only this, but Hubble’s interpretation of the redshifts as Doppler shifts, or recessional velocities, led to what is known as Hubble’s Law – that the Universe itself is expanding (Lemaître, 1927).

These stunning discoveries were certainly unexpected. They completely changed the face of cosmology, birthing the era of *physical cosmology*, which marries the more abstract notions of cosmic structure with theories from other physical fields. And in this new paradigm, the seeds of perhaps yet more surprising discoveries were beginning to emerge.

In 1933, Zwicky noticed that the galaxies in the Coma cluster moved too quickly to be bound together as they were, and he inferred the presence of extra ‘dark matter’ to provide the necessary gravity (Zwicky, 1933). At the time, the field of quantum mechanics was just beginning to make exciting discoveries, and it was thought that this would eventually explain the missing matter. The problem did not disappear however, and some forty years later, Rubin and Ford, while studying the rotation curves of the Andromeda Nebula, came to the same conclusion – the presence of ‘dark matter’ (Rubin and Ford, 1970). In fact, even in these early times, it was (correctly) believed that the dark matter outweighed the visible matter by a factor of about four.

While the gradual confirmation of this mysterious component was an interesting discovery of itself, it was not long before it became the crucial tool of a larger paradigm: the theory of *structure formation*. In particular, a view emerged in which the formation of galaxies was conceptually a two-step process: first, the broad-brush arrangement of dark matter was laid down purely by gravitational interactions, which resulted in a network of gravitationally bound clouds of dark matter, termed ‘halos’. And second, the gas in these halos cooled and condensed into visible galaxies via both gravity and dissipation (White and Rees, 1978). This two-step demarcation is still an important conceptual tool today.

Importantly, the segregation of the gravitational from the dissipational aspects of the process meant that the main features of the entire field of visible galaxies could be explained purely in terms of the simple physics of gravity, against the background of an expanding cosmos. Equivalently, it meant that the distribution of galaxies and their properties could be principally understood by the distribution of their host halos.

At the same time, statistical and physical models of this distribution were being developed. Seminal work by Press and Schechter (1974) provided an analytical model for the number density of halos as a function of their mass. This provided a quantitative prescription for a whole spectrum of objects, formed by hierarchical merging. Conversely, Peebles and collaborators promoted the correlation functions as measures of the *spatial* distribution of mass (Peebles, 1980). These functions quantify how far from completely random a set of points are under certain configurations (eg. the two-point correlation function measures the excess likelihood, over random, of finding *pairs* at a given separation). Together, these would form a basic description of the halo (and therefore galaxy) field.

However, two crucial parts of the theory were still highly unknown. First, the predictions of present-day structure were highly contingent on the state of the ‘initial’ density field. The work of Press and Schechter (1974) and Peebles (1980) had assumed that the initial conditions were such that there was an extremely smooth field, onto which very small Gaussian-distributed fluctuations were imposed. Furthermore, these fluctuations were proposed to have a spatial power spectrum in the form of a simple power-law.



However, there was no physical motivation for these simple assumptions. This began to change in 1981 with the introduction of *inflationary theory* Guth (1981). In this model, primarily proposed to account for the so-called *horizon* and *flatness* problems, quantum interactions in the very early Universe were vigorously stretched within a period of exponential Universal growth. Fortuitously, the standard prediction was that the fluctuations resembled a Gaussian-distributed power-law at the end of the period, providing a solid physical motivation for the theories of structure formation.

Secondly, the very nature of the dark matter was under debate. The typical understanding in early work was that the dark matter was something quite mundane, such as low-mass stars that had escaped detection. Quite aside from its particular nature were questions of its *behaviour* – does it interact with the visible matter by gravity only? Does it interact with itself? Does it have high or low thermal velocity? A lack of consensus on these points constrained the theories of structure formation to be primarily qualitative.

However, with the advent of particle physics, a new idea emerged and soon garnered widespread consensus: dark matter may be fundamentally different than ordinary ‘baryonic’ matter (Bergström, 2000). In this way, a type of matter truly representative of the two-step structure formation paradigm was possible – one which interacted with baryonic matter almost entirely through gravity only and was otherwise collisionless. Such a candidate was simple to model, and soon enough computer simulations could trace the evolution of the idealised dark matter field through the history of the cosmos.

These computer simulations proved invaluable for making robust predictions of the large-scale structure of the Universe, and testing them against the observed distribution (Davis et al., 1985). Beginning with the physically motivated initial conditions of inflation, and with a firm hypothesis for the behaviour of the dark matter, the simulations could integrate the equations of motion for each particle in a box defined by the cosmology derived from general relativity. For the first time it was seen directly that the dark matter truly coalesced into a spectrum of halos, roughly following the predictions of Press and Schechter, and arranged in a network of filaments and voids which could be described by the correlation functions.

An example of the spectacular ways in which these kinds of predictions could be observationally verified is shown in figure 1.1, which shows the famous ‘stick-man’ from the CfA (Center for Astrophysics) redshift survey during the 1980’s (Davis et al., 1985; Huchra et al., 1988). This figure is iconic as the first observational indication that galaxies in the Universe are not arranged randomly, but instead are organized into groups, filaments, clusters and superclusters, as predicted by the theory.

The convergence of the several components of the structure formation theory – inflation for the initial conditions, a non-baryonic dark matter component and analytic descriptions of the evolution of the dark matter halos – to conform to the newly uncovered cosmic structure was an unrivalled success. Furthermore, it proved to be useful in understanding the nature of the dark matter.

Until the results of the CfA survey, two opposing candidates for the dark matter were typically considered. The first was “cold” dark matter, referring to a heavy particle with negligibly small random motion in the early Universe. Such a particle was almost completely described by the expanding cosmos and gravity, and provided for hierarchical

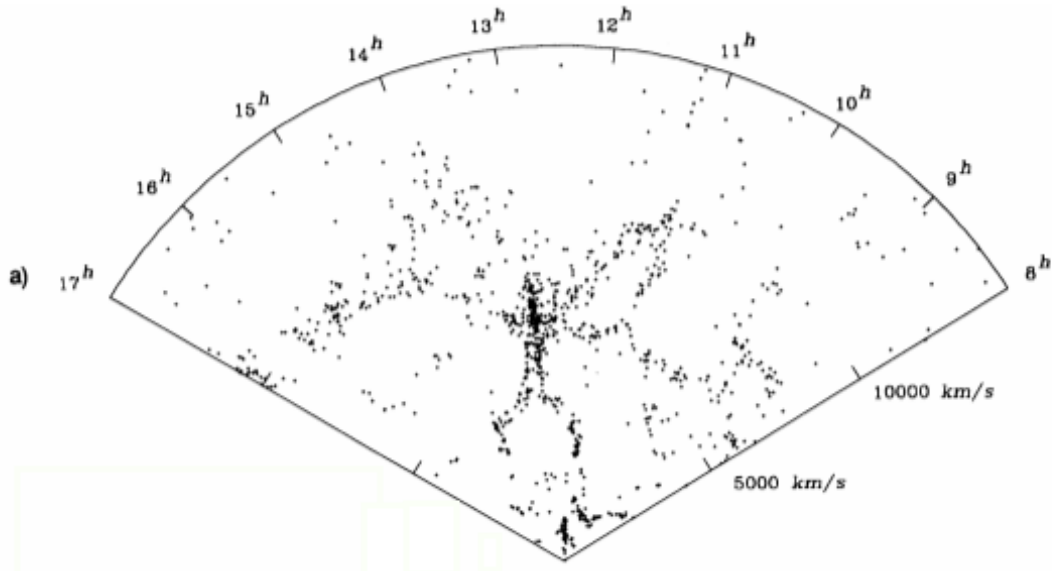


Figure 1.1: The famous ‘Stick-man’ figure from the CfA survey, from Lapparent, Geller, and Huchra (1986).

growth of structure. The second was “hot” dark matter, referring to a light particle with random motions close to the speed of light in the early Universe. This motion would cause the particles to be defined primarily by their random motions in the early Universe, eliminating the effects of gravity on scales comparable to the distance the particle would travel before slowing down. This meant that small scales would see no structure formation early on, and these would later form by the break-up of larger objects.

The CfA survey results were instrumental in showing that the “hot” dark matter was a poor representation of the cosmic distribution of galaxies, and the subsequent general acceptance of cold dark matter (hereafter CDM). There was still a surprising discovery to made, however.

#### 1.1.2 The state of affairs

Around the beginning of the millennium, two independent groups, analysing the distances and redshifts to samples of type Ia supernovae, came to the robust conclusion that the Universe was not only expanding, but accelerating. This required a component of energy in the Universe that was non-gravitational, and found its simplest expression in an allowed constant in the solution to Einstein’s field equations, typically called the ‘cosmological constant’,  $\Lambda$ . The Universe now contained *two* mysterious dark components, neither of which had an agreed-upon physical source, and which together comprised about 96% of the total mass-energy density.

The flourishing of primarily two kinds of observations in the last two decades have thoroughly entrenched this cosmological model of dark energy and dark matter – called  $\Lambda$ CDM after its components – as the standard paradigm. Firstly, ever-increasing numbers of galaxy positions in the relatively nearby Universe have been mapped by redshift surveys akin to the CfA (eg. 2dFGRS, Colless et al. 2001; SDSS, SDSS Collaboration 2009; GAMA, Driver et al. 2011). Measurements of the correlation functions in these surveys, over a wide range of scales have served to confirm and constrain the two-step

galaxy formation process driven primarily by gravity, and have even successfully located large-scale features in the distribution relating to baryonic effects in the early Universe (Eisenstein et al., 2005).

Perhaps even more stunning has been the direct observation of the initial conditions, imprinted on the cosmic microwave background radiation (CMB; Smoot et al., 1992) as minute temperature fluctuations. Ever-finer studies of this radiation (Spergel et al., 2003; Planck Collaboration, 2014a), a relic of the epoch of recombination, have revealed a spectrum of fluctuations – the seeds of galaxies and clusters – consistent to a high degree with the general predictions from inflation: a Gaussian distribution with a near power-law power spectrum. The 2D maps from these experiments also provide exquisite measurements of the relative densities of each component – dark matter, baryonic matter and dark energy.

Book-ended by the observations of large-scale structure at early times via the CMB and late-times via the galaxy distribution, the  $\Lambda$ CDM model of cosmological expansion and structural growth performs impressively well. Indeed, it is remarkable that a theory containing just six parameters – the baryon and dark matter densities, the primordial power spectrum index and amplitude, the characteristic angular size of fluctuations in the CMB, and the optical depth of cosmological re-ionization – is able to accurately describe the CMB, galaxy positions, supernovae measurements, and several other large-scale features of the Universe. So accurate is this simple model that proposed extensions and generalisations are statistically unnecessary within the precision of current data.

However, despite the unprecedented success of  $\Lambda$ CDM on large-scales, there is a growing body of evidence that it is lacking in its description of small-scales (roughly of the size of galaxies; cf. Primack and Cline, 2009). For instance, it predicts a higher density core of dark-matter haloes than is observed (Dubinski and Carlberg, 1991; Moore et al., 1999), a larger number of small galaxies surrounding galaxies such as our own Milky Way than is observed (Moore et al., 1999; Klypin et al., 1999; McGaugh, Barker, and Blok, 2003), more dwarf galaxies located in cosmological voids than observed (Peebles, 2001; Smith and Markovič, 2011), higher rotational velocities of satellite galaxies than is observed (Boylan-Kolchin, Bullock, and Kaplinghat, 2011) and fails to predict observed planar structures in the satellite distribution around galaxies like the Milky Way (Ibata et al., 2013).

Several ideas have been proposed in order to remedy these apparent discrepancies. The most obvious candidate is that the approximation of the two-step process breaks down in these regimes, and the effects of the complex interactions between baryons modifies the distribution expected from CDM-only theory. Certainly, the additional pressure associated with baryons would serve to reduce the precipitation of structure on small scales, alleviating tensions with observations. Considerable effort has been spent to understand the effects of baryons on structure formation (Cui et al., 2012; Cui, Borgani, and Murante, 2014; Martizzi et al., 2014; Mohammed et al., 2014; Bocquet et al., 2015), and it potentially provides a satisfying natural solution to the small-scale problems.

However, complementary solutions arise from the consideration of alternative forms of dark matter. Viable alternatives must have properties which align with CDM on large scales, but naturally produce less structure on small scales. One option considered has

been a form of dark matter which is self-interacting (i.e. relaxing the ‘collisionless’ assumption). This would serve a similar purpose to the inclusion of baryons, creating an extra pressure that limits the condensation of small-scale structures (Moore et al., 2000; Dave et al., 2001; Rocha et al., 2013). More pertinent to this thesis is the conjecture of a return to the arguments of dark matter temperature.

While hot dark matter has been conclusively ruled out by observations of large-scale structure in the local Universe, a milder version of the same argument is quite interesting. So-called “warm” dark matter (WDM) would have significant, but not extreme random velocities in the early Universe. It would be able to escape density fluctuations on very small scales, but would decelerate too quickly to escape larger ones. It would thus be characterised by a suppression of small-scale structure, but would have precisely the same properties as CDM on large scales. It is thus an extremely interesting alternative to CDM in order to explain the small-scale discrepancies in conjunction with baryonic processes.

To test these questions requires pushing our understanding of the formation of structure – so well tested at large scales – to the small-scale limit. Unfortunately, this endeavour is met by a host of difficulties; firstly, standard analytical methods begin to fail since the nonlinearities induced by gravitational effects exceed the reach of perturbation theory techniques; secondly, robustly characterising scales so small within cosmological volumes in numerical simulations is extremely challenging due to the required resolution; and thirdly, gravitational processes are no longer completely dominant in these regimes, and we must consider the effects of the formation and feedback of galaxies and other baryonic interactions.

We can make progress in the first two challenges by way of extensions of the old arguments of Press and Schechter (1974, hereafter PS). These arguments bypass the difficulties associated with tracing the nonlinear evolution of individual particles by assuming all of the mass is within spherically symmetric halos. By treating these halos as the basic units of structure formation, and providing descriptions of their properties based on analytic and phenomenological arguments, the entire density field can be reconstructed. For example, if one knows the *halo mass function* (i.e. the distribution of number density as a function of mass, predicted by PS), the spatial clustering of halos with respect to the background density field, and a formula for the density profile within a halo of a given mass, the fully nonlinear matter density field may be reconstructed, far beyond the reach of perturbation theory. This methodology is commonly referred to as the *halo model*, and provides invaluable insight into nonlinear structure formation when calibrated with more detailed numerical simulations.

The third difficulty relates to the problem of linking the visible matter to the underlying dark matter field. This problem is clearly crucial, since theories deal primarily with the simple dark matter, while observations are limited to the visible domain. While a broad-brush approach simply associates each galaxy with a dark matter halo, whose properties and arrangement are presumably unaffected by the existence of the galaxy, in detail the small baryonic effects on the distribution, and the possibility for multiple galaxies in large halos, will modify spatial statistics on small scales.

The information involved in linking the galaxies to dark matter halos runs two ways: the baryonic processes can be marginalised to better understand the pure-gravity aspect of structure formation (and thereby explore the nature of dark matter), or the structure formation processes can be marginalised to better understand galaxy formation and evolution. In this thesis we shall focus on the former.

Linking galaxies to dark matter can be broken into two effects: firstly the effect of the baryons on the internal properties of the halos themselves which results in modified halo mass functions and density profiles. Secondly, we must account for the way that the galaxies are spatially related to the underlying field. This latter effect is called the galaxy *bias*, and will be the focus of the second half of this thesis.

In simple terms, we do not expect that the galaxies will uniformly trace the dark matter. According to the halo model, they will preferentially be located inside high-density regions, defined as halos. The halos themselves will be preferentially situated in high-density areas, with this effect being more dramatic for high-mass halos. At the same time, high-density regions are more highly correlated, and so the correlation function of halos will depend on the masses of halos under consideration. As it turns out, this excess preponderance of halos can be modelled with an extension of the arguments of PS, which leaves us to determine the excess bias of galaxies with respect to the halos.

A popular means of modelling the galaxy-halo bias is to treat the problem entirely statistically. One simply proposes a formula (with free parameters) for the average number of galaxies in halos of a given mass, and assuming the distribution of these galaxies follows the distribution of mass *within* halos, the full galaxy bias can be completely reconstructed within the context of the halo model. This method is referred to as the *halo occupation distribution* (HOD) method, and is invaluable for rapidly generating the statistical properties of the galaxy field so that a range of parameters may be explored.

Armed with these tools, and given data which resolves the appropriate small scales, it would seem that it is only a matter of time before the nature of dark matter is uncovered.

### 1.1.3 On the brink

The coming decade offers both tantalizing possibilities, and overwhelming challenges. Headline surveys such as the Square Kilometre Array (SKA; Staveley-Smith and Oosterloo, 2015) and Euclid (Laureijs et al., 2011) will generate petabytes of data, mapping tens of millions of galaxies in the latter 3/4 of the history of the Universe (i.e. out to a redshift of two). Not only this, but smaller targeted surveys such as WAVES (Driver et al., 2015) will probe the halo mass function below typically nonlinear scales ( $\sim 10^{12}h^{-1}M_{\odot}$ ), generating a statistically coherent picture of the scales pertinent to the effects of dark matter for the first time.

These datasets have the potential to test and constrain our theories of structure formation to unprecedented accuracy, and almost certainly transform them in the process. But to do so requires a concerted effort. With an ever-growing volume of data, the necessity of cohesive, well-understood theoretical frameworks is becoming more apparent. Even more so, in the coming decade, these frameworks must precipitate in efficient, robust, usable and flexible software implementations.

The primary goal of this thesis is to present a set of such tools – software implementations and statistical methods – purposed to aid the interpretation of next-generation massive galaxy and cluster surveys. Using the time-tested models of PS and its various extensions, including the halo model and HOD, we synthesise the vast volumes of work undertaken in refining the various model components over the past two decades, and present unified and extendible tools for its continued development and application.

We envision that the tools and their statistical applications as presented in this thesis will be of benefit to the studies of galaxy formation and evolution, cluster cosmology, dark matter modelling and even large-scale analyses of the effects of dark energy. However, where possible, we focus primarily on the fundamental question of the nature of dark matter.

In the remainder of this chapter, we review the basic concepts of structure formation, especially as formulated in the analytic models of PS (§1.2.4), which forms the central cord of the applications in this thesis. Lastly, we summarise the observational probes, and the past, current and future surveys dedicated to exploring them, most pertinent to applications of our tools.

## 1.2 STRUCTURE FORMATION

We present the details of the so-called *halo model* of large-scale structure throughout this thesis. In particular, chapters 2 and 5 present the theory and implementation of the halo mass function (HMF) and analytic halo model respectively. In this section, we aim to introduce the subject, in preparation for the more formal descriptions of those chapters.

### 1.2.1 The Cosmological Expansion

The fundamental assumption of cosmology, called the *cosmological principle*, is that the Universe is both homogeneous and isotropic in its matter distribution. While these properties are generally assumed (and when initially adopted were little more than a conjecture), they have now also been shown to be valid on the required scales to a high degree of accuracy (eg. Goodman, 1995; Wu, Lahav, and Rees, 1999; Scrimgeour et al., 2012).

The cosmological principle gives rise to a unique solution of the metric describing space-time (Walker, 1935; Robertson, 1935; Robertson, 1936a; Robertson, 1936b), often called the FLRW (Friedmann-Lemaître-Robertson-Walker) metric, which can be written

$$ds^2 = dt^2 - a^2(t) \left( \frac{dR^2}{1 - kR^2} + R^2(d\theta^2 + \sin^2\theta d\phi^2) \right). \quad (1.1)$$

Here the factor  $k$  describes the curvature of space-time, and is zero if the Universe is flat (which we tacitly assume unless otherwise specified). This metric describes an expanding or contracting space, where stationary objects at position  $(R, \theta, \phi)$  recede from each other as the scalefactor  $a(t)$  evolves (clearly if  $a$  is shrinking, then the objects draw closer, but the evidence points to an expansion). The co-ordinates  $(R, \theta, \phi)$  are termed *comoving*, since they expand with the cosmos.



One of the great successes of this model is that it accounts for the fact that nearby stationary (with respect to the comoving co-ordinates) observers will witness each other receding at a rate proportional to their physical distance

$$v = \frac{\dot{a}}{a}D \equiv H(t)D, \quad (1.2)$$

where  $H$  is termed the Hubble constant. This result was the famous discovery of Hubble in 1929, using extra-galactic nebulae (i.e. galaxies) measured to be receding via their redshifts at a rate proportional to their distance.

The particular form for  $a(t)$  is derived via Einstein's general theory of relativity, which maintains that the space-time structure is determined by the mass-energy density, supplemented again by the cosmological principle (Friedmann, 1922; Friedmann, 1924; Lemaître, 1931). A convenient form of these equations, often called the Friedmann equation, is

$$H^2(t) = \frac{8\pi G}{3}\rho - \frac{k}{a(t)^2}, \quad (1.3)$$

where  $\rho$  is the mass-energy density and  $G$  is Newton's gravitational constant. We can define a particularly important density, called the *critical* density, as that which yields a flat Universe, such that  $k = 0$ :

$$\rho_c(t) = \frac{3H^2(t)}{8\pi G}. \quad (1.4)$$

It is also convenient to specify the densities of the constituents of the Universe as ratios to this critical density,

$$\Omega_x = \frac{\rho_x}{\rho_c}. \quad (1.5)$$

Typically, four density components are required to fully specify the composition of the total density: the matter density (comprising both 'dark matter' and normal baryonic matter)  $\Omega_m$ , the radiation density  $\Omega_r$ , the vacuum-energy density (often called the 'dark energy' density)  $\Omega_\Lambda$ , and the curvature density  $\Omega_k$ . In this thesis, unless otherwise specified, the quantities denoted thus will refer to their values *today*, and in addition  $H_0$  will refer to the Hubble constant today. We note also that  $a$  is typically normalised to unity today (with  $a \rightarrow 0$  at early times).

With these definitions, the Friedmann equation becomes

$$\frac{H(t)}{H_0} = \sqrt{\frac{\Omega_m}{a^3} + \Omega_\Lambda + \frac{\Omega_r}{a^4} + \frac{\Omega_k}{a^2}}, \quad (1.6)$$

where for a flat Universe,  $\Omega_k \equiv 1 - \Omega_m - \Omega_\Lambda - \Omega_r = 0$ , and at late times the contribution from radiation is negligible, so that

$$\frac{H(t)}{H_0} \approx \sqrt{\frac{\Omega_m}{a^3} + \Omega_\Lambda}. \quad (1.7)$$

The dependence of each term on the scalefactor introduces a series of epochs in the history of the Universe in which different components dominated. Assuming a flat Universe, the initially dominant component was radiation, which transitioned to matter, and then gradually to the vacuum energy<sup>1</sup>. We find ourselves now in a period with  $\Omega_m \approx 0.3$  and  $\Omega_\Lambda \approx 0.7$ .

A convenient physical tracer of the scalefactor is *redshift*, which measures the stretching of radiation in transit by the expansion of space, and can be measured directly by the (red) shift of known spectral lines. It is given by

$$z = \frac{1}{a} - 1. \quad (1.8)$$

### 1.2.2 The seeds of structure

Clearly, if the Universe was *perfectly* homogeneous, no structures would arise – the density field would be perfectly smooth to the present day. The study of structure formation is principally concerned with determining how inhomogeneities arose, and how they came to be as we observe them today.

The question of how they arose has roots in quantum theory and cosmological inflation. The standard paradigm maintains that in the very early epochs after the big bang, quantum interactions caused density perturbations which were subsequently rapidly stretched by the exponential inflation of the cosmos. The perturbations then existed on a large spectrum of spatial scales, with those that were more stretched (i.e. occupying larger scales) tending to create less dramatic overdensities. For example, the fluctuations measured in the CMB at  $z \sim 1100$  are as small as roughly 1 part in  $10^5$ . From this barely perceptible spatial spectrum of a distribution of over- and under-densities, the runaway effect of gravity grew the structures of today.

A mathematical description of the evolution of the inhomogeneities involves treating the matter as a pressureless fluid, at rest in the comoving co-ordinates, but expanding with the cosmos. The over- and under-densities induced by the inflation-processed quantum fluctuations are modelled as perturbations on this otherwise perfectly homogeneous fluid:

$$\delta(\vec{x}) = \frac{\rho(\vec{x})}{\bar{\rho}} - 1. \quad (1.9)$$

There are two statistical properties of this field of density perturbations which are important: their 1-point PDF (*probability density function*), and their spatial distribution. Both of these are predicted (to some extent) by the inflationary paradigm.

The PDF of density perturbations is predicted to be (very close to) Gaussian, with mean zero (and this is consistent with measurements of the CMB to this date). Their

<sup>1</sup> For given density parameters, the evolution of the scalefactor and  $H(t)$  can be generated numerically, and we refer the reader to [cosmocalc.icrar.org](http://cosmocalc.icrar.org) for a particularly convenient interface for such calculations.



spatial distribution is typically described in Fourier space, by the isotropic two-point function, which for a Gaussian distribution completely specifies the field:

$$P(k) \propto \langle |\delta_{\vec{k}}|^2 \rangle \propto k^n, \quad (1.10)$$

where  $k = 2\pi/\lambda$  and  $n \sim 1$  the *spectral index*. The amplitude of  $P(k)$  is not predicted by the physical theory, but can be set by measurements of the CMB or (as is more common), by large-scale local structure, as we shall see.

The power spectrum is modified in the transition from radiation domination to matter-domination (at  $z \sim 3400$ ), by the Meszaros effect (Meszaros, 1974), which suppresses small-scale structure, leaving a characteristic ‘bend’ in the power spectrum at a scale close to the sound horizon at the time,  $\sim c/H(z \approx 10^4)$ , and a small-scale logarithmic slope of  $n - 4$ . Various other effects throughout these early times, especially those of baryons, have a less pronounced impact on the power spectrum (Eisenstein and Hu, 1999), and as long as the perturbations remain linear ( $\delta \ll 1$ ), they can be fully accounted for by solving the Boltzmann equations for each particle species (Lewis, Challinor, and Lasenby, 2000).

### 1.2.3 Linear Growth

When the perturbations are still small, their evolution can be determined by linearizing the equations of the phase space fluid, which yield a solution in which the perturbations remain stationary, but grow in amplitude in proportion to the growth factor

$$D^+(z) \propto \frac{H(z)}{H_0} \int_z^\infty \frac{(1+z')dz'}{[H(z')/H_0]^3}. \quad (1.11)$$

Thus, at early times, or even late times at large scales, the power spectrum evolves as

$$P(k, a) \propto D^{+2}(a) P(k, a_0). \quad (1.12)$$

To describe the PDF of perturbations, we must have a formula for the variance about the mean of zero. However, the variance is diverging, since perturbations exist on all scales. In reality, we are more interested in the variance at certain scales, corresponding to objects. This requires smoothing the density field at the given scale with some smoothing kernel (also known as filter functions, or window functions),  $W(\vec{x})$ , such that

$$\int W(\vec{x}) d^3\vec{x} = 1. \quad (1.13)$$

A window function will generally (but not always) have a corresponding volume for a given scale. For the most common choice, a top-hat in real space, this volume is given by the usual  $V(R) = 4\pi R^3/3$ , and so the smoothing scale can be related to an enclosed mass,  $m(R) = V(R)\bar{\rho}$ .

With a window function defined, the variance is specified as

$$\sigma^2(m(R)) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) \hat{W}^2(kR) dk, \quad (1.14)$$

where  $\hat{W}$  is the Fourier transform of the window function. Clearly, the evolution of  $\sigma(m)$  with redshift (in the linear domain) is merely given by the evolution of the power spectrum, again specified by the growth factor.

In particular, a convenient parameter for setting the amplitude of the power spectrum is given by the variance in spheres of a given size in the density field linearly extrapolated to today. For historical reasons, the standard scale at which this is set is  $8h^{-1}\text{Mpc}$  in comoving co-ordinates, giving rise to the parameter  $\sigma_8$  which is of order unity according to studies of the CMB and local structure. The precise value of this parameter is important in determining the abundance of massive objects in the Universe.

Thus the homogeneous and isotropic universe yields a simple set of equations precisely describing the statistics of the growth of structures in the regime where nonlinearities are not important. Beyond this point, however, the analytic machinery breaks down. To continue to follow the evolution of the growth of structures in detail requires computationally expensive  $N$ -body simulations, which effectively solve for the evolution of the coarse-grained phase space fluid, with each particle corresponding to an element of this fluid.

However, with certain approximations, there is an alternate route to modelling the growth of structure beyond this point that is far more efficient.

#### 1.2.4 Press-Schechter Formalism

PS considered the case of overdense spheres of collisionless dust in an expanding cosmos, with radius  $R$  and overdensity  $\delta_i \ll 1$  at some very early time. In this case, according to Newton's theorems (and Birkhoff's theorem in the context of general relativity), the evolution of the field inside the spherical region proceeds independently of the density field outside the region, and vice versa. Indeed, the sphere behaves as an independent universe with density  $\Omega = (\Omega_0 + \delta)$ .

The evolution of such a universe is governed by Newtonian gravity, and has well-studied solutions, which have a distinctive expansion and collapse structure. The sphere initially expands with the cosmos, as its overdensity is low. As the overdensity grows with the linear growth factor, gravity begins to dominate and the sphere reaches a maximum expansion, before symmetrically collapsing (theoretically to a point, but realistically to a virialized structure).

Three derived values are of immediate interest (each is derived in terms of an Einstein-de Sitter universe, i.e.  $\Omega_m = 1$ ,  $\Omega_\Lambda = 0$ , but at least for the first two, the cosmology dependence is weak). First of all, considering the purely *linear* growth of the perturbation, the value at the turnaround is (Liddle and Lyth, 2000)

$$\delta_{\text{lin}}^{\leftarrow} \simeq 1.06. \quad (1.15)$$

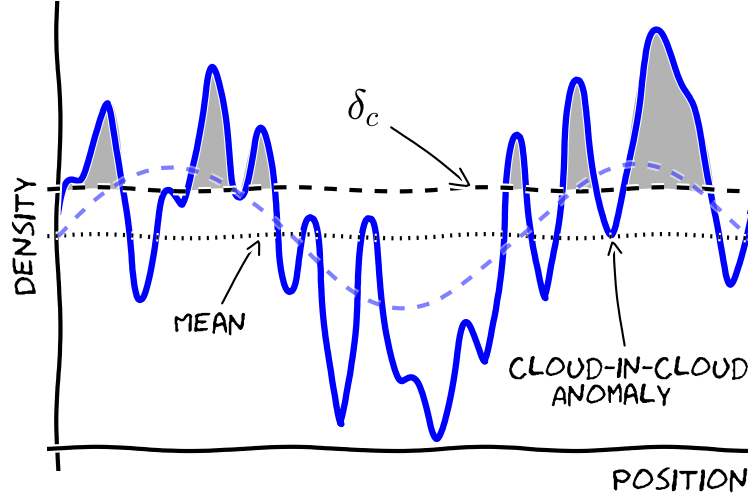


Figure 1.2: Schematic plot of the argument of Press and Schechter (1974). Peaks in the smoothed density field above the threshold  $\delta_c$  collapse, rendering a fraction of the total volume inside collapsed structures. Increasing the smoothing scale inherently decreases the height of peaks, lowering this fraction. A long-wavelength mode is shown as the light-blue dashed curve. Parts of the density field on the peaks of this curve which dip below the critical threshold may be overwhelmed by surrounding overdensity and nevertheless collapse, which is illustrative of the “cloud-in-cloud” problem. One notices that these regions would be above the threshold if smoothed on a larger scale.

Secondly, the value when the region has “collapsed” (i.e. to a theoretical point) is

$$\delta_c \simeq 1.686, \quad (1.16)$$

and finally, the fully *nonlinear* overdensity at collapse is

$$1 + \delta_{\text{nl}}^{\text{vir}} \equiv \Delta_{\text{vir}} \simeq 178. \quad (1.17)$$

This latter value is commonly used to define the extent of virialized halos in simulations.

The important point, at least in terms of the model of PS, is that in the approximation of spherical perturbations, regions irreversibly collapse when they exceed the threshold overdensity of  $\delta_c$  (though it would seem that  $\delta_c$  should be recalculated precisely for a given cosmology, Jenkins et al. (2001) showed that a constant value gives a useful universality in the abundance of halos). In qualitative terms, the situation is like a range of underwater mountain peaks, where those that exceed some threshold height are observed (cf. figure 1.2).

Recalling that the distribution of smoothed overdensities at some scale  $m$  (or equivalently  $R$ ) is a Gaussian, with a well-defined variance  $\sigma^2(m)$ , the fraction of the total spatial volume that is collapsed on any given scale is simply given by the integral over the Gaussian upper tail (cf. figure 1.3):

$$f(> m, z) = \int_{\delta_c}^{\infty} \frac{1}{\sqrt{2\pi}\sigma(m, z)} \exp\left(-\frac{\delta_m^2}{2\sigma^2(m, z)}\right) = \frac{1}{2} \text{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma(m, z)}\right). \quad (1.18)$$

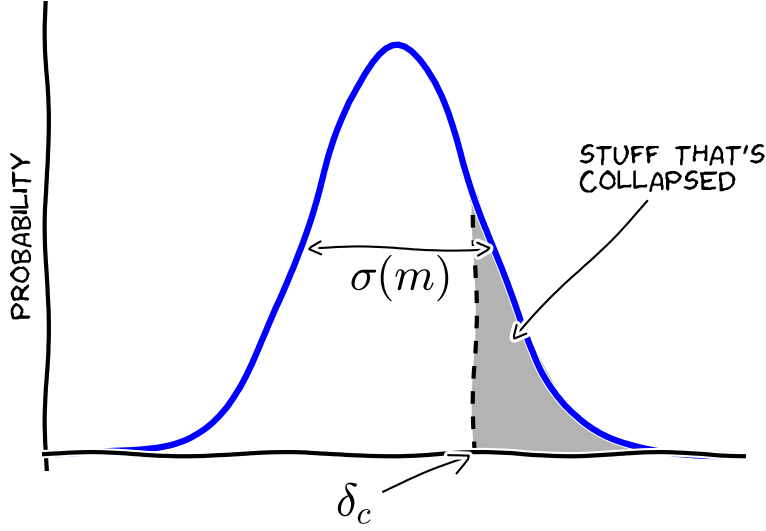


Figure 1.3: Illustration of the integral to determine the fractional volume of collapsed matter in PS theory.

PS noted that at  $m = 0$ ,  $f(> m) \rightarrow 1/2$ , and so only half of the universal mass is here accounted for. They argued that the result needs to be multiplied by a factor of 2, otherwise only positive fluctuations are included. Though this ad hoc change had no clear physical motivation, it succeeded in generating roughly accurate results, and has since been placed on firmer theoretical grounds (Peacock and Heavens, 1990; Bond et al., 1991).

In brief, the factor of 2 came to be known as the “cloud-in-cloud” problem, which conceptually states that while a region of a particular size might be below the collapse density  $\delta_c$ , it may be embedded in a larger region which has  $\delta > \delta_c$ . We would expect in this case that the overdensity of the larger region would cause the entire region to collapse (cf. 1.2 and its caption). The theoretical solution to this problem came in the form of what is called “excursion set theory”. This theory relates small increases of the variance (i.e. decreases in mass/radius) to random steps, chosen from a Gaussian distribution. The result is a random walk, or rather a set of random walks, constituting an “excursion set”. The solution to the cloud-in-cloud problem is to evaluate the probability of a region first exceeding  $\delta_c$  at a given radius. Under the excursion set paradigm, this is equivalent to the probability that the first upcrossing of a random walk occurs at a given  $\sigma^2$ , and this can be solved (under the assumption of a Gaussian density distribution and uncorrelated steps) using the theory of stochastic processes (Zentner, 2007). The solution is identical to the Gaussian case of PS, but includes the necessary factor of 2.

The power of the PS formalism is in recognising that by increasing the filtering scale,  $m$ , the field will be smoother, and therefore the fraction of collapsed matter will decrease. By associating the total decrease in collapsed density  $\Delta\rho = df/dm$  with a number density of halos between  $m$  and  $m + dm$ , we can solve for the number density:

$$\frac{dn(m, z)}{dm} = -\sqrt{\frac{2}{\pi}} \frac{\rho_0}{m} v \frac{d \ln \sigma(m, z)}{dm} \exp(-v^2/2), \quad (1.19)$$

where  $\nu = \nu(m, z) = \delta_c / \sigma$  is often called the ‘peak-height’.

Without specifying the exact form of  $\sigma(m)$ , we expect this function to have power-law-like behaviour at low mass, with an index of  $\sim -2$ , and an exponential cutoff at high mass. Stunningly,  $N$ -body simulations decades later confirmed this basic structure of the mass distribution of halos, showing that the analytic arguments of PS were a remarkably good description of the nonlinear evolution of structure.

### 1.2.5 Extended PS

The primary objection to the PS formalism is the ad hoc factor of 2 necessary to account for all the mass. This objection was formally answered by Peacock and Heavens (1990) and Bond et al. (1991), who extended the simple PS formalism to firmer theoretical grounds (their extension is commonly termed EPS). We will not pursue the details of this extension here, rather referring the reader to the excellent review of Zentner (2007), but we will outline the key ideas.

The central argument is that when smoothed on some scale  $m_1$ , some regions may fall below the necessary threshold, but when smoothed on a larger scale  $m_2$ , they exceed it. This can be thought of as a cloud of moderate overdensity (or even underdensity) within a cloud of larger overdensity, and is often called the “cloud-in-cloud” problem. Clearly we would consider that the smaller region is within a collapsed region, but the basic argument of PS does not treat it so, unless multiplied by the ad hoc factor.

To rigorously describe this process requires the *excursion set* formalism, which follows the trajectories (in overdensity space) of regions as the filtering scale is modified. Given certain assumptions (such as a spherical boundary, and uncorrelated fluctuation modes), the statistical description of the collection of trajectories is analytic (characterised by a Markov process), and neatly gives the same formula as the basic PS formalism – including the correction factor.

The real merit of this extended approach is that it opens the way for more realistic models and other derived quantities. Thus, for instance, Sheth, Mo, and Tormen (2001) considered, in the place of a spherical boundary, a boundary motivated by ellipsoidal collapse. Their solution closely mimics the PS formula, but generalises it, including 2 parameters that could be fit to simulations. This form is still standard within theoretical modelling today.

Furthermore, the generality of EPS makes it possible not only to model halo *abundance* but also the spatial distribution of halos with respect to the linear matter density field. Qualitatively, this proceeds by considering the relative peak-height required to collapse in a large-scale overdensity. Since the relative overdensity for collapse is reduced in such an environment, we expect the number density of halos to increase in dense environments, and thus for higher mass halos to be more clustered. In the context of EPS, this can be formally specified as

$$P_{halo}(k) \approx b^2(m)P_m(k), \quad (1.20)$$

with

$$b(m, z) = 1 + \frac{\partial}{\partial \delta_c} \ln \frac{dn(m, z)}{dm}. \quad (1.21)$$

For the mass function of PS, this results in (Mo and White, 1996)

$$b(m) = 1 + \frac{v-1}{\delta_c}. \quad (1.22)$$

Again, the very approximate arguments of spherical collapse provide a surprisingly good description of the results from  $N$ -body simulations. Summarily, the EPS formalism is a beautiful analytic means of making physical predictions in regimes where the only other option is to run very expensive simulations. The continued development of both methods will continue to be of great benefit in modelling structure formation in the years to come.

#### 1.2.6 WDM

We have thus far focussed on a dark matter which is not only collisionless, but also “cold”, i.e. its velocity at early times is negligible. This results in our power spectrum with large-scale slope of  $n \sim 1$  and small-scale slope of  $n - 4$ . We recall that the small-scale slope is suppressed due primarily to the Meszaros effect, in which the expansion of the Universe due to radiation density damps the growth of structure. In this model, the matter itself is assumed stationary, following the expansion of the cosmos.

If, however, the dark matter particle were “warm” (WDM), we expect that its kinetic energy would drive it out of small-scale perturbations, erasing them completely. Its exact temperature (or equivalently mass, typically measured in keV) would dictate the largest perturbation it could erase. Thus we would expect that the power spectrum would, in addition to the Meszaros damping, be truncated at some characteristic scale given by the temperature.

In pure CDM, growth of structure is expected to proceed hierarchically, or “bottom-up” – that is, small dense halos are the first to form, with larger halos forming by mergers of smaller ones. In WDM, this picture must be altered, as the smallest halos cannot form purely by spherical collapse, due to the suppression of small-scale fluctuations. In this case we get a mixed picture – “medium”-sized halos form first (with a corresponding delay compared to CDM), and proceed to grow hierarchically. However, there is also some “top-down” formation, in which large halos fragment into smaller ones, which begins to populate the smaller scales below the original cutoff.

The overall effect of WDM is thus to suppress small-scale structure at the present time – this gives rise to a dearth of halos in voids, less substructure within large halos and other properties which seem to relieve some of the tensions we mentioned concerning CDM at small scales. It has become clear that although WDM does have these effects, it cannot be the sole contributor to the solution of these problems, as it would have to be *too* warm to be of real use (Schneider et al., 2014). However, these attractive properties have made it an interesting proposition in the past decade and a half, and it still may be

that the final solution is a combination of WDM and as-yet poorly understood baryonic processes.

The attention on WDM models has recently increased significantly due to the tentative discovery of a 3.5 keV X-ray line (Bulbul et al., 2014; Boyarsky et al., 2014), which has been proposed to be the result of the decay of a sterile neutrino (the leading candidate particle for WDM) with a mass of 7.1 keV. Whether or not this discovery turns out to be verified, exploring the possibility of WDM through structure formation is clearly a valuable endeavour.

Unfortunately, WDM has proven notoriously difficult both to model and constrain with observation. On the modelling front, the truncation in the initial power spectrum unearths a deep problem with the setup of  $N$ -body simulations, in which anisotropic collapse and finite resolution effects cause the collapse of spurious halos (Hobbs et al. 2015; Power and Robotham *in prep.*) muddying any attempted analysis at low mass. This also effectively thwarts attempts to refine the PS formalism in the context of WDM. Progress has been made in this regard, but we are far from a perfect solution.

WDM will also be challenging to detect for various reasons, for example, an observed dearth of void galaxies may be due to them lying below the brightness threshold of a survey, or alternatively due to poorly understood physical processes which prohibit the formation of visible galaxies in the existing halos. To unravel these effects from those of WDM is difficult. The most successful constraints have come from the characterisation of the Lyman- $\alpha$  forest, which places lower-bounds on the mass of the particle at  $\sim 3.3$  keV (Viel et al., 2013).

We return to the question of WDM throughout this thesis, but most heavily in chapter 6.

### 1.3 OBSERVATIONS

There are several probes of the large-scale structure which may be used to test our theories of structure formation and thus its constituents – the dark matter and energy, background cosmology and inflationary paradigm. These span the observable age of the Universe, from the CMB to present-day structures, over several orders of magnitude in scale. Impressively, the standard  $\Lambda$ CDM model, with few marginal exceptions, is able to account for them all to within current precision.

In this section we give an overview of two of these probes – cluster surveys and generic galaxy redshift surveys – which are particularly pertinent to the tools presented in this thesis<sup>2</sup>. For each we note the particular benefits and challenges of the approach, and list some key surveys – past, present and future.

#### 1.3.1 Cluster surveys

According to the hierarchical growth scenario, the latest objects to form are also the largest: clusters of galaxies. These objects, with typical masses  $> 10^{13}h^{-1}M_{\odot}$ , provide

<sup>2</sup> We note that a third probe which is particularly suited to halo model calculations, but not discussed in detail in this thesis, is galaxy-galaxy lensing (e.g. Mandelbaum et al., 2005).



a veritable treasure-trove of cosmological information. Their immense size prevents the ejection of material via supernovae, and thus renders them a “fair sample” of the Universe’s components, allowing direct measurements of the cosmic baryon fraction, and insights into key galaxy formation processes (Voit, 2005). However, for our purposes, perhaps the most important tests arise from their abundance.

Recall that the halo mass function in the PS formalism has an exponential cutoff proportional to  $\exp(-\nu^2/2)$  at high masses. This high mass regime is then *exponentially* sensitive to the amplitude of fluctuations, parameterised by  $\sigma_8$ . Along with this important test, the evolution of the abundance of clusters is an excellent tracer of the growth history of the Universe, and provides some of the most stringent constraints on the matter density and extensions to the  $\Lambda$ CDM paradigm concerning the growth history, such as the equation of state of dark energy,  $w_0$  and  $w_a$ , non-gaussianity in the initial conditions and even deviations from general relativity (Allen, Evrard, and Mantz, 2011; Sartoris et al., 2015). Notably, the cluster-scale HMF is insensitive to the nature of dark matter, viable models of which only affect the halo abundance at much lower masses.

The primary challenge of cluster cosmology is the specification of the mass-observable relation, i.e. the mapping from the observation to the mass of the host halo. Understanding the systematics involved in these mappings is crucial to precision measurements of the cluster mass (Noh and Cohn, 2012). A further challenge is the inherent rarity of clusters, which means that very large surveys are necessary to produce satisfactory statistics.

Cluster cosmology is a truly panchromatic field – traditionally, clusters have been in the jurisdiction of X-rays, due to the hot gas constituting their intra-cluster medium. However, they are also commonly observed in optical/IR bands through galaxy kinematics and luminosity. More recently, observations in sub-millimetre bands via the Sunyaev-Zeldovich (SZ) effect (Sunyaev and Zeldovich, 1972) have also become a useful tool.

X-ray surveys have been in operation since the early 1970’s (Giacconi et al., 1972), though the number of detected clusters first entered the hundreds with the ROSAT all-sky survey (RASS, Vikhlinin et al. 1998). Since then, the major X-ray surveys have been performed with the XMM-Newton telescope, including the XMM Cluster Survey (XCS, Mehrrens et al. 2012) which has identified over 500 clusters out to  $z \sim 1$ , and the XMM-XXL survey, which identified about 600 clusters out to  $z \sim 2$ . The aggregate of X-ray measured clusters is now just over 1500, as presented in the meta X-ray cluster survey (MXCS; Piffaretti et al., 2011).

Larger samples are derived from optical/IR observations. Indeed, cluster catalogues measured from SDSS data are larger than 100,000 in number (Wen, Han, and Liu, 2012).

Samples identified via the SZ effect are currently the smallest, but are able to homogeneously sample the rarest halos. Telescopes such as the Atacama Cosmology Telescope (ACT, Hasselfield et al. 2013), the South Pole Telescope (SPT, Reichardt et al. 2013) and Planck (Planck Collaboration, 2014b) have detected over 1000 clusters out to  $z \sim 1$ .

The numbers of detected clusters should increase dramatically in coming years due to current and future missions. Currently operational in the X-ray regime, eROSITA (Pillepich, Porciani, and Reiprich, 2012) expects the detection of almost  $10^5$  clusters of mass  $m > 5 \times 10^{13} h^{-1} M_\odot$  out to  $z \sim 1$ . In the optical bands, the Dark Energy Survey



is expecting  $\sim 1.7 \times 10^5$  clusters out to  $z \sim 1.5$ , with a limiting mass of  $5 \times 10^{13} h^{-1} M_{\odot}$ . The increase in redshift depth is extremely important for tracing the growth history, and this aspect will be pushed even further by the future Euclid mission (Sartoris et al., 2015), which expects up to  $2 \times 10^6$  clusters out to  $z \sim 2$ . The volume and precision of incoming data will rival and perhaps exceed that of the current best models, which motivates a renewed effort to robustly characterise the mass function (Murray, Power, and Robotham, 2013b; Paranjape, 2014).

### 1.3.2 Galaxy redshift surveys

The last few decades have been the golden age of the galaxy survey – large scale efforts to homogeneously map the three-dimensional positions of galaxies over defined portions of the sky. Early galaxy surveys in the 1980s, such as the CfA, were instrumental in shaping the field of structure formation, by conclusively showing that the dark matter could not be “hot”. Since that time, the number of known galaxies has increased by almost three orders of magnitude, with notable large surveys such as the 2dFGRS (Two-degree field Galaxy Redshift Survey; Colless et al. 2001) the SDSS (Sloan Digital Sky Survey; SDSS Collaboration 2009) and GAMA (Galaxy and Mass Assembly; Driver et al. 2011) collecting hundreds of thousands of spectroscopic redshifts, and significantly contributing to wide acceptance of the concordance  $\Lambda$ CDM model. While we have seen that these surveys offer a wide range of scientific revenue, in terms of structure formation, the primary workhorse has been the galaxy clustering.

On large scales, a feature known as the baryon acoustic oscillation, imprinted as excess correlation during the early Universe, provides the most robust cosmological tests (Eisenstein et al., 2005). In effect, its scale is a standard ruler, tracing out the expansion history of the Universe,  $a(t)$ , and remaining independent of other probes such as supernovae.

On smaller scales, the galaxy clustering, interpreted using frameworks such as the halo model with HOD, provides an excellent example of the complex relationship between galaxies and their host halos. Fitting models to the clustering over scales of  $0.01 \lesssim r \lesssim 40 h^{-1} \text{Mpc}$  simultaneously tests the underlying models of structure formation, and constrains the galaxy-halo connection (Zehavi et al., 2004; Zehavi et al., 2005; Zehavi et al., 2011; Beutler et al., 2013).

In addition to the major surveys already mentioned, many other more targeted surveys have been exploited using halo model predictions of clustering, including 6dFGS (Six-degree Field Galaxy Survey; Beutler et al. 2013) and GAMA (Christodoulou et al., 2012) at low redshift, and UKIDSS UDS (UKIRT Infrared Deep Sky Survey; Quadri et al. 2008), NMBS (NEWFIRM Medium Band Survey; Wake et al. 2011), DEEP2 (Mostek et al., 2013), PRIMUS (PRIsm MULTI-object Spectrograph; Skibba et al. 2014) and UltraVISTA (McCracken et al., 2015) at high redshift.

Future surveys will be deeper, and therefore probe smaller halos, where the nature of dark matter is more effective. At the same time, the mean redshift increases with survey depth, which naturally leads to a better chance of delineating dark matter models.

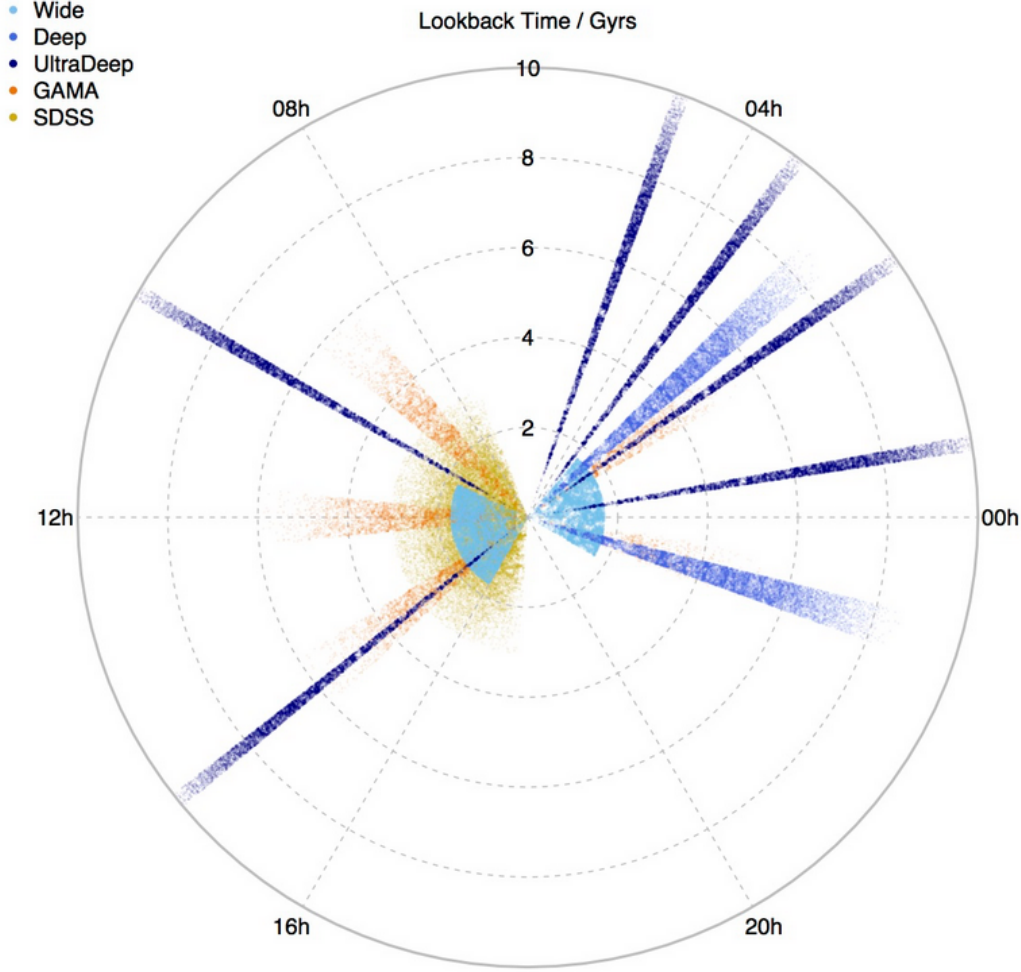


Figure 1.4: Depth (in lookback time) of two completed representative surveys (SDSS and GAMA) alongside the planned surveys of WAVES. Figure from <http://www.wavesurvey.org/WAVES.php>.

In particular WAVES-UltraDeep (Wide Area Vista Extragalactic Survey; Driver et al. 2015), due to commence in 2020, will detect  $\sim 5 \times 10^5$  galaxies down to below the present typical nonlinear mass,  $\sim 10^{12} h^{-1} M_{\odot}$ , out to  $z = 1.5$  (cf. figure 1.4). Furthermore the Prime Focus Spectrograph Galaxy Evolution Survey (PFS GES, Takada et al. 2013) will generate a similar number of galaxies out to  $z = 2$  and VLT-MOONS (Multi Object Optical and Near-infrared Spectrograph for the Very Large Telescope; Cirasuolo et al. 2014) will perform an SDSS-like survey at  $z = 1 - 1.5$ .

These will be precursors to yet-larger surveys: the Dark Energy Spectroscopic Instrument (DESI) survey will obtain optical spectra for tens of millions of galaxies and quasars out to 10 billion light years, the 4-metre Multi-Object Spectroscopic Telescope (4MOST) will collect approximately 25 million spectra over 14000 square degrees, Euclid (Laureijs et al., 2011) will map of order 50 million galaxies out to very high redshifts, and in conjunction with a proposed MSE (Maunakea Spectroscopic Explorer, Simons et al. 2014) instrument, these will make complete samples of galaxies over all relevant mass scales within cosmologically significant volumes out to  $z = 5$ . This will be extended

by the WFIRST telescope, expected to launch close to 2030 (Greene et al., 2012), and complemented by the Square Kilometre Array (SKA) survey at radio frequencies.

Certainly, our view of the growth of structures is set to be completely redefined.

#### 1.4 NEXT-GENERATION TOOLS

We are about to enter into a new era in which the large surveys of the past decade are overwhelmed by surveys that probe both very low mass and extended time-scales. With this new data, many of our questions concerning the nature of dark matter could be answered, if our models are precise enough to fully exploit the data.

Unfortunately, at least in the context of the halo mass function and halo modelling, the software implementations are considerably behind. Though the implementation of the mass function is necessary for many applications, including the analysis of cluster surveys, the status quo in the field is to write one's own code. The implementation is relatively straight-forward, however without a standard to compare to, it is all too easy to introduce small errors. In any case, it is not a productive use of the time of researchers to be re-writing the same code many times, and furthermore it prohibits reproducibility. Despite this, it is very difficult to find *any* publicly available implementations, and typically the ones that are easiest to locate (with the exception of our new code, `hmf`) are difficult to use and extend.

Similarly, though the halo model is evidently a very popular theoretical framework, there has not been a community-wide effort to produce a standardised, efficient software framework<sup>3</sup>. This is particularly concerning in the case of the halo model, as the various integrations and transformations required are non-trivial and ensuring accuracy in the code can be quite time-consuming. Without an appropriate standard to compare to, this task becomes even more difficult.

It is crucial for the success of the field with the onset of the data deluge that simple, well-documented, flexible, extendible and efficient tools are made available for interpretation and analysis.

We address this situation in this thesis, by presenting two new codes, one for calculating halo mass functions via the PS formalism (chapter 2), and the other for calculating halo model quantities (chapter 5). For each, we present an extended worked example that uses the code to answer statistical questions. Using the mass function code, `hmf`, we answer the question "How well do we know the halo mass function?" by examining the uncertainties in the theoretically predicted mass function arising from uncertainties in our knowledge of the cosmological parameters (chapter 3). Secondly, using the halo model code, `halomod`, we ask if there are any discernible signatures of WDM on the correlation function of galaxies (chapter 6). In addition, we present a new statistical description of the mass function, that is highly efficient, and based on a parameterisation directly in terms of the mass (chapter 4). This description, which we call the MRP, will

<sup>3</sup> The Python package at <https://code.google.com/p/chomp/> may be an exception, though it does not have the same level of flexibility we develop in this thesis. Note also that the `HALOTOOLS` package was published after initial submission of this thesis, however it is primarily aimed at populating halo catalogues rather than the analytic formalism.

be useful for fitting mass function models to observed data with arbitrary per-object mass uncertainties. Finally, we summarise and conclude in chapter 7.

---

## HMFCALC : AN ONLINE TOOL FOR CALCULATING DARK MATTER HALO MASS FUNCTIONS

---

### ABSTRACT

The dark matter halo mass function (HMF) is a characteristic property of cosmological structure formation models, quantifying the number density of dark matter haloes per unit mass in the Universe. A key goal of current and planned large galaxy surveys is to measure the HMF and to use it to test theories of dark matter and dark energy. We present a new web application for calculating the HMF – the frontend `HMFCalc` and the engine `hmf`. `HMFCalc` has been designed to be flexible, efficient and easy to use, providing observational and theoretical astronomers alike with the means to explore standard functional forms of the HMF or to tailor their own. We outline the theoretical background needed to compute the HMF, we show how it has been implemented in `hmf`, and finally we provide worked examples that illustrate `HMFCalc`’s versatility as an analysis tool.

### 2.1 INTRODUCTION

There is now a wealth of compelling observational evidence that we live in a Universe whose matter content is predominantly dark ( $\sim 84\%$ ; cf. Planck Collaboration, 2014a) and non-baryonic in nature (cf. Bergström, 2000). Our theories of cosmological structure formation predict that this dark matter clusters into massive gravitationally bound structures called haloes. The dark matter halo mass function (hereafter HMF) quantifies the number of these haloes per unit comoving volume of the Universe as a function of their mass. The HMF is sensitive to the cosmological parameters, primarily the mass-energy density of dark matter  $\Omega_c$  and dark energy  $\Omega_\Lambda$  (e.g. Murray, Power, and Robotham, 2013b), but it also depends on the nature of the dark matter. The standard Cold Dark Matter (CDM) model predicts an HMF in which the number of haloes increases with decreasing halo mass  $M$  approximately as  $M^{-1.8}$  (e.g. Lukić et al., 2007; Bhattacharya et al., 2011), whereas viable Warm Dark Matter (WDM) models predict fewer haloes than the CDM model at low masses (e.g. Schneider, Smith, and Reed, 2013; Pacucci, Mesinger, and Haiman, 2013). The potential of the HMF as a probe of dark matter and dark energy is widely recognised (e.g. Tinker et al., 2008; Vikhlinin et al., 2009) and is one of the key science drivers of current and planned future galaxy surveys (Driver et al., 2011; Pierre et al., 2011).

Cosmological  $N$ -body simulations are now established as the tool for studying the HMF (cf. the recent review by Knebe et al., 2013), but the information contained in a simulation is usually distilled and recast in a more compact form. Usually this is the comoving number density of haloes per unit logarithm of the halo mass  $M$ ,

$$\frac{dn}{d \ln M} = M \cdot \frac{\rho_0}{M^2} f(\sigma) \left| \frac{d \ln \sigma}{d \ln M} \right|; \quad (2.1)$$

here  $\sigma$  and  $\rho_0$  are the cosmology-dependent mass variance and mean density and  $f(\sigma)$  represents the functional form that defines a particular HMF fit. Eq 2.1 is not difficult to compute, but neither is it straightforward. We have developed an easy-to-use and flexible web-application to compute Eq 2.1, which we call `hmf`, and its frontend `HMFcalc` which can be accessed at <http://hmf.icrar.org>. `HMFcalc` can be used in a number of ways, including as

- a standard against which to check one's own code;
- an easy-to-use interface to generate HMFs against which to check observational/simulations data; and
- a visually intuitive way to explore the effects of cosmology on the HMF.

The objective of this chapter is to present a detailed overview of `hmf` and `HMFcalc`, describing its implementation and the underlying philosophy for this approach, as well as providing some worked examples that illustrate its usefulness and versatility.

The chapter is structured as follows. In §2.2 we provide the theoretical background necessary to compute the HMF, setting out a compilation of HMF fitting functions drawn from the literature and demonstrating how the HMF differs in CDM and WDM models. In §2.3 we describe our implementation of `hmf` and `HMFcalc` and discuss the algorithms and methods used. In §2.4 we present some worked examples using `HMFcalc` and finally in §2.5 we summarise our plans for extending `HMFcalc`'s functionality and usability.

## 2.2 THE HALO MASS FUNCTION (HMF)

The HMF quantifies the number of dark matter haloes per unit mass per unit comoving volume of the Universe,

$$\frac{dn}{d \ln M} = M \cdot \frac{\rho_0}{M^2} f(\sigma) \left| \frac{d \ln \sigma}{d \ln M} \right|$$

where  $f(\sigma)$  is the fitting function that we shall return to shortly,  $\rho_0$  is the mean density of the Universe and  $\sigma$  is the rms variance of mass within a sphere of radius  $R$  that contains mass  $M$ ,

$$M = \frac{4\pi\rho_0}{3} R^3. \quad (2.2)$$

The mass variance is calculated via the integral,

$$\sigma^2(R) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) W^2(kR) dk \quad (2.3)$$

where  $k = 2\pi/\lambda$  is the wavenumber corresponding to the wavelength  $\lambda$  of the fluctuation,  $P(k)$  is the linear power spectrum (see below) and  $W(kR)$  is the Fourier transform of the top-hat window function<sup>1</sup> (a sphere with sharp edges in real space), i.e.

$$W(kR) = \frac{3 [\sin(kR) - kR \cos(kR)]}{(kR)^3}. \quad (2.4)$$

Using Eq 2.2 to relate  $M$  to  $R$ , the right-most factor of Eq 2.1 can be written

$$\frac{d \ln \sigma}{d \ln m} \equiv \frac{1}{6\pi^2 \sigma^2} \int k R W(kR) \frac{dW(kR)}{d(kR)} P(k) k^2 dk, \quad (2.5)$$

where

$$\frac{dW(kR)}{d(kR)} = 3 \left( \frac{\sin(kR)}{(kR)^2} - \frac{W(kR)}{kR} \right). \quad (2.6)$$

The linear power spectrum  $P(k)$  characterises the distribution of matter density perturbations as a function of wavenumber  $k$ ; formally it is computed by averaging over  $|\hat{\delta}(\vec{k})|$ , the square of the Fourier-transformed real-space density fluctuations  $\delta(\vec{x}) = \rho(\vec{x})/\bar{\rho} - 1$ , where  $\vec{x}$  and  $\vec{k}$  are the spatial coordinate and wave vector respectively. We express it as

$$P(k) = A k^n T^2(k), \quad (2.7)$$

where  $T(k)$  is the transfer function,  $A$  is the normalisation constant and  $n$  is the spectral index. We follow convention and use the cosmological parameter  $\sigma_8$ , which measures the mass variance on a scale of  $8h^{-1}\text{Mpc}$ , to calculate  $A$  (via Eq. 2.3). The primordial power spectrum, imprinted during the epoch of inflation during the first moments after the Big Bang, is expected to have a form  $P(k) \propto k^n$ . The transfer function quantifies how this primordial form is modified on different scales, and it is particularly sensitive to the nature of the dark matter and the baryon density parameter  $\Omega_b$ .

### 2.2.1 Fitting Functions

As noted in the introduction, it is the fitting function  $f(\sigma)$  that defines a particular HMF fit. Early analytical work by Press and Schechter (1974, hereafter PS) and Bond et al.

<sup>1</sup> Note that there is freedom in the choice of window function; we follow the convention in which the window function is a top-hat in real space, but we could adopt a Gaussian in real space, or a top-hat in Fourier space. The dependence of  $\sigma$  on the choice of window function is well known and the implications have been explored in studies such as Schneider, Smith, and Reed (2013). We shall return to this point in §2.5.



(1991) established a simple form for  $f(\sigma)$  by assuming that haloes form by spherical collapse, finding

$$f(\sigma) = \sqrt{\frac{2}{\pi}} \frac{\delta_c}{\sigma} \exp\left(-\frac{\delta_c^2}{2\sigma^2}\right), \quad (2.8)$$

where  $\delta_c \simeq 1.686$  is the critical overdensity for spherical collapse.

However,  $N$ -body simulations of cosmological structure formation have revealed that the PS form underestimates the abundance of higher mass haloes and overestimates the abundance of lower mass haloes (e.g. Sheth, Mo, and Tormen, 2001; White, 2002; Lukić et al., 2007). Sheth, Mo, and Tormen (2001, hereafter SMT) explored an extension to the PS formalism by considering ellipsoidal rather than spherical collapse and obtained a form for the mass function that is identical to Eq 2.1 but with a modified  $f(\sigma)$ . Their form for  $f(\sigma)$  contains parameters that are constrained by fitting to simulation results, which improves the fit to simulation data but at the expense of a purely analytical description. Subsequent studies (cf. Table 2.3) have followed SMT in assuming that the HMF can be expressed in the form of Eq 2.1 with  $f(\sigma)$  fitted to simulation data to characterise the HMF.

The mass assigned to a halo in a numerical simulation depends on the assumed working definition for what constitutes a halo. This has been covered in comprehensive detail by Knebe et al. (2013) and references therein, but we make a few remarks. Algorithmically, haloes are identified by locating enhancements in the matter density field, and two methods are commonplace;

1. the Friends-of-Friends (FOF) method (Davis et al., 1985) identifies haloes as groups of particles that lie within a linking length of  $b$  times the mean interparticle separation of each other, resulting in groups of arbitrary shape; and
2. the Spherical Overdensity (SO) method (Lacey and Cole, 1994) identifies haloes as spherical groups of particles centred on enhancements in the density field whose radius is fixed by requiring the enclosed mass density is a pre-defined multiple of the background density.

How closely these working definitions match the true underlying halo population depends on a variety of factors, including algorithm parameters (e.g. the choice of  $b$  in the FOF method) and algorithm assumptions (e.g. explicit assumption of sphericity in the SO method), as discussed in Knebe et al. (2013), as well as the more complicated issue of how to define a halo (e.g. Prada et al., 2006; Cuesta et al., 2008; Anderhalden and Diemand, 2011; Firmani and Avila-Reese, 2013; Diemer, More, and Kravtsov, 2013). A proper survey of this topic is beyond the scope of this chapter and so, instead, we focus on fitting functions as they have been presented in the literature, which we summarise in Table 2.3. This provides a concise summary of the forms for  $f(\sigma)$  that have appeared in the literature and which are included in HMFcalc, and we list also the cosmology, mass and redshift ranges and halo definition for each fit. HMFcalc aims to be as general as possible and as such we have included all fitting functions discovered in the literature. Each of these fitting functions are plotted in Fig 2.1 for comparison.



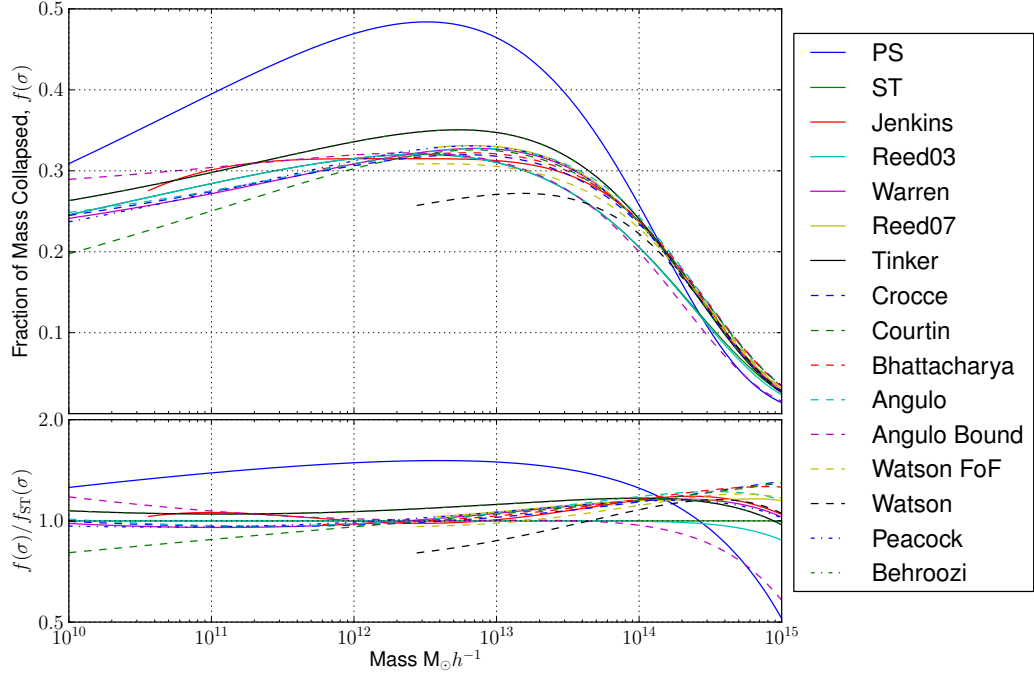


Figure 2.1: Upper: All fitting functions at redshift zero over a large mass range (limits placed as appropriate on each function). The Press-Schechter form over-predicts low-to-intermediate mass haloes and under-predicts high-mass haloes when compared to the results of  $N$ -body simulations. Broadly the fitting functions are in good agreement for intermediate mass haloes  $\sim 10^{12} M_{\odot} h^{-1}$  but are divergent for low and high mass haloes. All functions tend towards zero at high mass, corresponding to the rarity of these massive objects. Lower: All functions normalised by the Sheth-Tormen form.

Note that we use  $\Delta_{\text{vir}}$  to refer specifically to the value of the cosmology-dependent overdensity with respect to the mean or critical density that corresponds to a virialized halo, according to the usual definition based on spherical collapse (cf. Eke, Cole, and Frenk, 1996; Bryan and Norman, 1998). In contrast,  $\Delta_h$  defines a threshold overdensity for a structure to be identified as a halo, but one that is more general insofar as it can assume any value (within reason). For example, it can be set such that  $\Delta_h = \Delta_{\text{vir}} \simeq 100$  for a fiducial  $\Lambda$ CDM model, or fixed at  $\Delta_h = 200$  as in e.g., Jenkins et al. (2001) or Behroozi, Wechsler, and Conroy (2013), or as large as  $\Delta_h = 2400$ , which is one of the values adopted in Tinker et al. (2008); the precise value can be treated as a free parameter. We note that Tinker et al. (2008) adopts parameters  $A, a, b, c$  whose values are dependent on the value of  $\Delta_h$ , which we have omitted in Table 2.3 for brevity.

### 2.2.2 Redshift Dependence

Foundational studies assumed that the form of the fitting function  $f(\sigma)$  was insensitive to redshift (eg. Jenkins et al., 2001), so that  $\sigma(M, z)$  completely captured redshift dependence. It has since been demonstrated that this is a limited approximation (eg. Tinker et al., 2008), however the mass variance remains the primary source of evolution, governed by the growth factor (cf. Eq. 1.11).

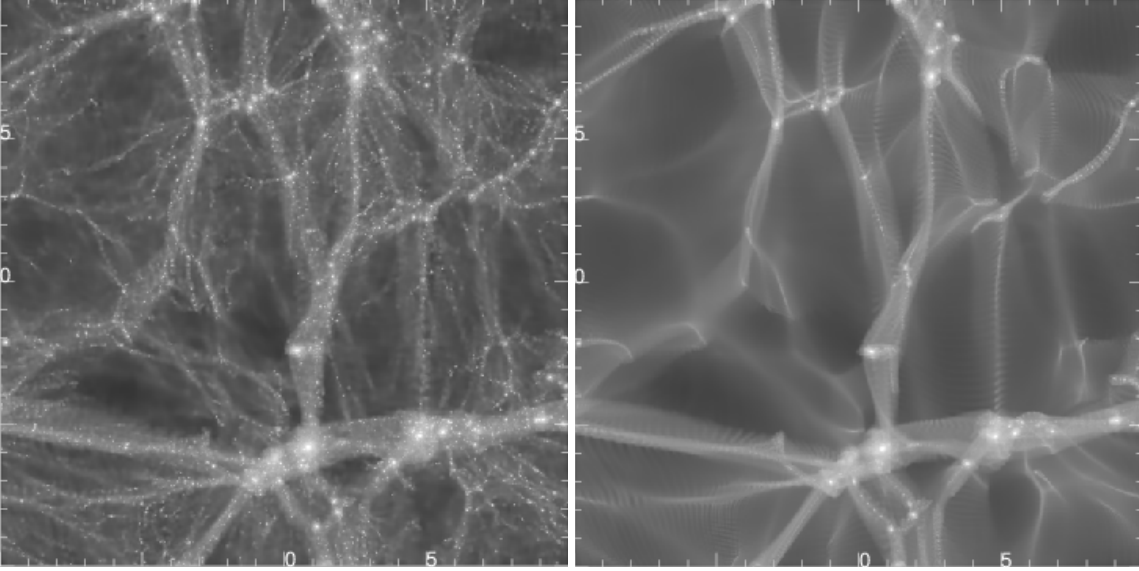


Figure 2.2: Visual impression of the projected dark matter density in a cosmological  $N$ -body simulations of a  $20 h^{-1}\text{Mpc}$  box, modelling the growth of structure in a fiducial CDM model (left panel) and its WDM counterpart (right panel). For the WDM model we assume a particle mass of  $m_X = 0.5 \text{ keV}/c^2$ . Note the absence of small-scale structure (i.e. low mass dark matter haloes) in the WDM run compared to the CDM run.

Modern fitting functions (post 2007) generally incorporate an explicit redshift dependence to account for the effects of evolution beyond this simple approximation.

### 2.2.3 Warm Dark Matter Models

The default in HMFcalc is to calculate the HMF assuming an underlying Cold Dark Matter (CDM) model, but we have included the option to calculate the HMF in Warm Dark Matter (WDM) alternatives. Small-scale power is suppressed in a WDM model compared to its CDM counterpart, reflecting the free-streaming of WDM particles in the early Universe, dispersing small-scale density perturbations that would otherwise collapse via gravitational instability to form low-mass dark matter haloes (e.g. Power, 2013). The influence of the dark matter model on structure formation is readily apparent in Fig 2.2, which shows the projected dark matter density at  $z = 0$  in a  $20 h^{-1}\text{Mpc}$  patch in a fiducial CDM model (left panel) and in a WDM model assuming a WDM particle mass of  $m_X = 0.5 \text{ keV}/c^2$ . Although the large-scale features (filaments and voids, positions of the most massive haloes) are the same in both runs, a wealth of small-scale structure – low mass dark matter haloes – is evident in the CDM run but not in the WDM run (e.g. Power, 2013).

The standard approach to capturing the effect of WDM on the power spectrum is to filter the corresponding CDM power spectrum on small scales by an additional transfer function, i.e.  $P(k) = P_{\text{CDM}} T_k^{X^2}$ . The transfer function of Bode, Ostriker, and Turok (2001) is widely used, and it is the one we adopt in HMFcalc; here

$$T_k^X = (1 + (\alpha k)^{2\mu})^{-5/\mu}, \quad (2.9)$$

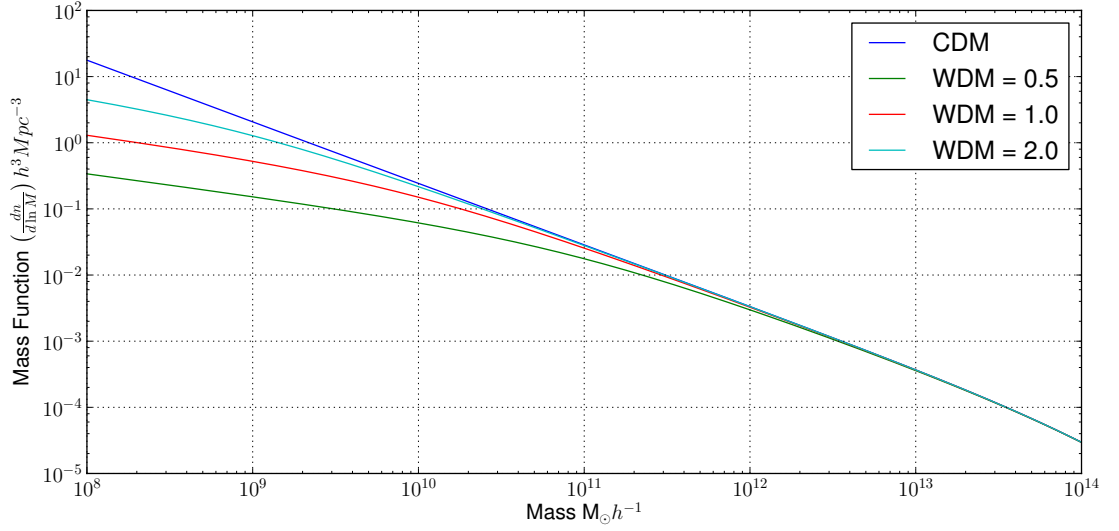


Figure 2.3: The HMF calculated assuming a fiducial CDM model and three different WDM models with particle masses of  $m_X=0.5, 1$  and  $2 \text{ keV}/c^2$ . Note the suppression of the number density of lower mass haloes.

with  $\mu = 1.2$ , and

$$\alpha = 0.048 \left( \frac{\Omega_X}{0.4} \right)^{.15} \left( \frac{h}{.65} \right)^{1/3} \left( \frac{1}{m_X} \right)^{1.15} \left( \frac{1.5}{g_X} \right)^{.29} \quad (2.10)$$

with  $\Omega_X$  the current fractional density of the WDM particle (this can be taken as equivalent to the CDM density  $\Omega_c$  in a single-species WDM model),  $m_X$  is the particle mass in keV, and  $g_X$  controls the abundance of the species relative to photons and has the fiducial value of 1.5 for a light neutrino.

By default in `HMFcalc`, we assume that  $\nu$  and  $g_X$  are set to their fiducial values and allow only a single-species model; the only free parameter that we allow is  $m_X$ . Fig 2.3 shows a fiducial CDM and WDM HMFs for  $m_X=0.5, 1$  and  $2 \text{ keV}/c^2$  – as the WDM particle mass decreases, so too does the number density of lower mass haloes, as we would expect. However, as we note in § 2.5, there have been recent developments in the modelling of Warm Dark Matter models (e.g. Benson et al., 2012; Schneider, Smith, and Reed, 2013; Pacucci, Mesinger, and Haiman, 2013) that we will incorporate in future versions of `HMFcalc`.

### 2.3 IMPLEMENTATION

`HMFcalc` has been designed to be visually intuitive, easy to learn and use, flexible and extensible. These requirements guided our implementation;

- The code is written in Python, an extremely versatile language that is widely used in the astronomical community.
- The user interface is a web-app frontend, which is designed to work across multiple platforms and for its ease of use.

- The web-framework is provided by Django. It is flexible, easy to use and being written in Python, it integrates easily with the hmf backend.
- The code has been made open-source at <https://github.com/steven-murray/HMF>. This way the code can be examined and updated, and also understood by those wishing to learn about the HMF.

We have also chosen to take an object-oriented approach, which makes sense because many of the HMF calculations share attributes. Class structures<sup>2</sup> provide a way to define common attributes and make them available throughout all methods of the class, which in our case enables faster updating of parameters.

In the following subsections we describe in detail the backend hmf (§ 2.3.1), before introducing the web-application (hereafter web-app) frontend HMFcalc.

### 2.3.1 HMFcalc's engine, hmf

The calculation of the HMF is performed using our Python package hmf. Algorithm 1 summarises the steps necessary for the calculation of the differential HMF in hmf. We expand upon the key concepts below.

---

#### Algorithm 1 Calculate HMF

---

```

if No Transfer File then
     $k, T \leftarrow \text{CAMB}()$ 
else
     $k, T \leftarrow \text{ImportTransferFile}()$ 
end if
 $T_f() \leftarrow \text{Interpolate } T$ 
 $k \leftarrow \text{Re-grid between user-set bounds}$ 
 $T \leftarrow T_f(k)$ 
 $P_0(k) = k^n T^2$  ▷ Eq 2.7
 $\sigma_{8,0} \leftarrow \sigma(R = 8, z = 0)$  ▷ Eq 2.3
 $P(k) \leftarrow \left( \frac{\sigma_8}{\sigma_{8,0}} \right)^2 P_0(k)$  ▷ Normalise
if WDM then
    Find  $T_{\text{WDM}}(k)$  ▷ Eq 2.9
     $P(k) \leftarrow P(k) * T_{\text{WDM}}^2$  ▷ Eq 2.9
end if
 $\sigma_0(M) \leftarrow \sigma(M, z = 0)$  ▷ Eq 2.3
 $K \leftarrow \frac{d \ln(\sigma)}{d \ln(M)}$  ▷ Eq 2.1
Find  $d(z)$  ▷ Eq 1.11
 $P(k) \leftarrow P_0(k) d^2(z)$  ▷ Linear Growth
 $\sigma(M, z) \leftarrow \sigma_0(M) d(z)$  ▷ Linear Growth
Calculate  $f(\sigma)$  with appropriate form
HMF  $\leftarrow \frac{\rho_0}{M} * f(\sigma) * K$  ▷ Eq 2.1

```

---

<sup>2</sup> Objects to which attributes and methods may be attached.

### 2.3.1.1 User-Interface

Interaction with `hmf` is provided through the `Perturbations` class, which contains all attributes needed to calculate the HMF (the power spectrum, limits in  $k$ , the redshift, cosmology and mass variance etc.) and all the methods needed to calculate these quantities.

The constructor for the `Perturbations` class initialises an ‘empty’ object with a range of cosmological and numerical parameters set. The various quantities associated with the object are calculated only as necessary, and accessed as class properties, rather than methods. These include the power spectrum (`power`), mass variance (`sigma`), growth factor (`growth`), effective spectral index (`n_eff`), fitting function (`fsigma`) and the differential and cumulative mass functions (`dndlnm` and `ngtm`).

While each fitting function is defined as a single method within a dedicated class, they hide behind a simple user interface, in which the user specifies a string identifier for the required fit (such as ‘ST’ for Sheth, Mo, and Tormen (2001) or ‘Reed03’ for Reed et al. (2003)). Extending the range of fitting functions is as straightforward as writing a method containing the function definition, and registering it in a dictionary with a unique string identifier. However, users also have an option to pass a function in the format of a simple string that the class will parse and use as the fitting function.

### 2.3.1.2 Transfer Function

The transfer function is handled in one of two ways. If a filename is provided, the relevant file will be read and used as the transfer function. If no filename is provided, we use the public *Code for Anisotropies in the Microwave Background* (CAMB) (Lewis, Challinor, and Lasenby, 2000) to compute our transfer functions. CAMB is called via `pycamb`<sup>3</sup>, its open-source Python wrapper. This enables efficient shared-memory multi-core calculation of the transfer function for a range of input parameters. We directly use the transfer function rather than the power spectrum, as this allows for more flexibility in modifying certain parameters (eg.  $n_s$  and  $\sigma_8$ ).

### 2.3.1.3 Calculating the Mass Variance

The most involved component of the mass function calculation is the mass variance integral. Eq. 2.3 must be integrated numerically, with the following considerations.

The window function and its derivative are functions of the product  $kR$ , but we evaluate Eqs 2.3 by integrating over  $k$ . For this reason care must be taken when solving the integrals to ensure that the results converge. We demonstrate why in Fig 2.4, where we plot  $\int_0^{kR} W^2(x)dx$  and  $\int_{kR}^{\infty} W^2(x)dx$ . The integral  $\int_0^{kR} W^2(x)dx$  allows us to identify an upper limit on the minimum  $kR$  required for convergence; we want the range of  $kR$  for any  $R$  to have a minimum that bounds the non-zero parts of the function. As a rule of thumb, we identify this as the value of  $kR$  for which 95% of  $\int_0^{\infty} W^2(x)dx$  is included; from Fig 2.4, this corresponds to  $kR = 0.1$  so we have the condition,

$$k_{\min} R_{\max} < 0.1 \tag{2.11}$$

<sup>3</sup> <https://github.com/joezuntz/pycamb>

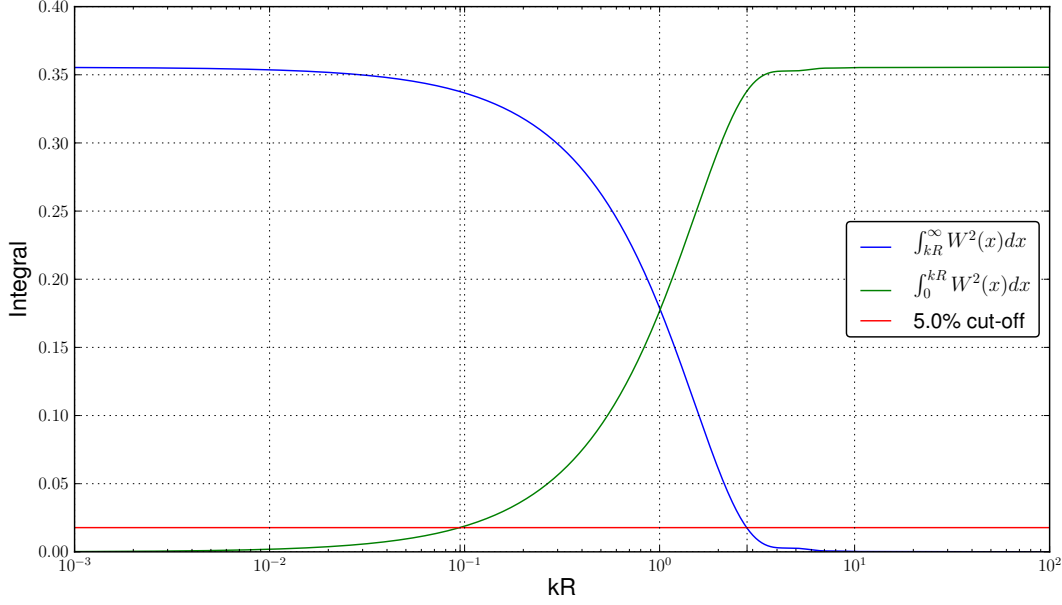


Figure 2.4: The integral of the window function gives an upper limit of 0.1 on the minimum of the  $kR$  range, and a lower limit of 3 on the maximum of the  $kR$  range, with a 95% contribution of the total integral.

Repeating this exercise for  $\int_{kR}^{\infty} W^2(x)dx$ , we deduce

$$k_{\max}R_{\min} > 3. \quad (2.12)$$

These bounds on  $kR$  are enforced in hmf by default, with a warning issued if they are not observed. Usually the limits on  $k$  to which the power spectrum is extrapolated far exceed these rules for any sensible mass range, but instances can arise where more stringent limits on  $k$  are required (cf. §2.4.2).

The second consideration is the integration scheme, which was chosen for a balance of performance and accuracy. In order to choose a suitable metric for these quantities, we note the following. The HMF is exponentially sensitive to the mass variance at high mass, so we choose to base our maximum uncertainty requirement for  $\sigma$  on its impact on the HMF here. Specifically, we require a maximum of 0.1% error due to  $\sigma$  in the differential mass function at  $10^{15}M_{\odot}h^{-1}$ . This limit ensures that uncertainty due to numerical integration is negligible compared to that of choice of fitting function or precision of the fit. Furthermore, the uncertainty in the cumulative function (which is more often used in clustering studies) is dominated by the uncertainty at the same mass of the differential function, so this limit also represents an approximate precision for the cumulative mass function.

The dependence of the mass function on fractional changes,  $\Delta$ , in  $\sigma$  is dominated by  $f(\sigma)$ :

$$\frac{n(M)}{n_{\text{true}}(M)} - 1 \simeq \frac{f(\sigma_{\text{true}} + \Delta\sigma_{\text{true}})}{f(\sigma_{\text{true}})} - 1. \quad (2.13)$$

Evaluation of Eq. 2.13 with  $\sigma_{\text{true}}(10^{15}M_{\odot}h^{-1}) = 0.5$  for several common fitting functions reveals  $\Delta \leq 10^{-4}$  is sufficient for the desired error in the mass function.

We tested several integration schemes, with the condition that their relative accuracy in calculating  $\sigma(10^{15}M_{\odot}h^{-1})$  be  $< 10^{-4}$ . For each scheme, the transfer function from CAMB is first interpolated with linear splines (higher order splines tend to perform poorly upon extrapolation) in log-log space. Each scheme was run for a range of output resolutions from CAMB, along with an extremely high-resolution run, which we take as the ‘true’ value. The most efficient scheme was selected as that which produced the required accuracy in the shortest total time, where total time was measured as  $t_{\text{CAMB}} + 1400t_{\text{INT}}$ . This time takes into account the competing effects of CAMB run time and integration time, where it is assumed that on average a user will have a length 700 vector of masses.

We find that a combination of CAMB resolution parameters of  $k_{\text{max}} = 0.25$  and 11 steps per logarithmic interval, with a simple trapezoid rule integrator using 250 steps, is the most efficient scheme. This is especially so if one wishes to re-calculate the mass variance for parameters that do not require the transfer function to be re-calculated.

#### 2.3.1.4 *Calculating the Cumulative Mass Function*

Calculation of the cumulative HMF,  $n(> M)$ , has several pitfalls. To remain as general as possible, HMF allows the user to pass any vector of masses, so long as it is equispaced in log-space. Calculation of the cumulative function however generally must be performed up to  $M = 10^{18}M_{\odot}h^{-1}$  to ensure convergence. To complicate this, most of the fitting functions are only defined in a finite mass range (see Table 2.3), which does not extend to the requisite mass.

In `hmf`, we implement a tiered approach to resolving this issue. It first determines whether the input mass range extends to  $10^{18}M_{\odot}h^{-1}$ . If not, it attempts to calculate the differential function as far as possible above the input mass range given the intrinsic mass range of the fitting function. Above this mass it assumes a power law. This ensures accuracy for any input mass vector.

The integration scheme used here is likewise a simple trapezoid rule in log-log space. This is well justified as the mass function is close to linear in this space, and the trapezoid rule can produce a cumulative integral for free, rather than separately integrating above each mass.

#### 2.3.1.5 *Re-calculation for varying parameters*

Users may want to re-calculate the HMF for different cosmological parameters, redshifts or fitting functions. To optimise this process, `hmf` tracks dependencies of quantities through chaining immediate links. For example, algorithm 1 demonstrates that the unnormalised power spectrum at the current epoch is dependent on the transfer function (and therefore all cosmological and numerical parameters involved in that calculation), and also the spectral index, but it does not depend on redshift, normalisation, or overdensity (either virial or critical). Any variable can be updated by calling the `update()` method with relevant keyword arguments; quantities that depend *directly* on the updated variables are deleted, which triggers deletion of quantities that depend directly on these and so forth. The next time any of these quantities are required, they are automatically re-calculated, using the updated variable. This system ensures that quantities are only re-calculated as necessary.



This dependency tracking allows for the optimal calculation of many HMFs for varying parameters in an efficient and error-free manner. We have used it in Murray, Power, and Robotham (2013b), generating 20,000 varying HMF realisations for each of 11 different fitting functions. On average this method reduces processing time by 50% in tests where each available parameter is changed one time, although the size of reduction depends on which parameters are most often modified (e.g. changing only redshift reduces computations on average by more than 94%).

### 2.3.2 *HMFcalc's web interface and the Django Framework*

Django provides a web-framework tuned to fast development and performance. Although developed with online newsrooms in mind, Django is easily adapted to scientific applications. A scientific Django web-app will generally be built around four parts – urls, views, templates and calculation – and Fig 2.5 depicts the flow of information within this framework<sup>4</sup>. The user initiates the information flow by pointing their browser at a particular URL. This URL is transformed to an address and contacts the Apache server on which HMFcalc is operating. The server translates the URL for input to the app's own `urls.py` module, which parses it and uses the information to choose a view function from the `views.py` module. This function may also be passed extra arguments from the input URL, based on query strings. The view chosen does all the logic, and dictates *what* should be seen on the final page. In the case of HMFcalc, one of the views accesses a form within the `forms.py` module for display. This form is passed through to a *template*, which is an enhanced html file. The template controls the presentation logic, putting the form fields down in appropriate positions on the page (with the help of CSS). The final template is then passed through to the server to be displayed by the browser.

We utilise a standard approach to the Django implementation. Nevertheless, we note some aspects of the design. Most of the views are defined as *generic class-based views*. This enables simpler construction and more flexibility in function. It also supports the creation of dynamic tabs for the user's navigation, via the plugin *django-tabination*. This implementation of the views is most important in the view that displays the main input form.

The view that displays the form is non-standard because there are two variants of it – one is for the initial creation of data, the other is to add more data. The forms for these differ in that the first lets the user choose a mass range (and grid), while the second does not – it uses the mass range specified previously. To create this functionality in the simplest way, we define a base form view, which implements the same logic for each form (i.e. what to do with the data after it has been successfully entered). We then create two subclasses, one for creation and one for addition of data. Each of these passes different arguments to the form class, which dynamically adds or removes the mass range inputs. This construction also allows for each of these subclasses to define their own tab for user navigation.

<sup>4</sup> Note that the 'Calculation Module' node is a black-box for the actual calculations using the hmf backend, which has been described already.



Upon successful submission of the form, the base form view performs the main analysis. To do this, it takes the parsed input arguments and passes them to an external function that wraps around the hmf backend. For each combination of the input parameters, an HMF is produced and saved as an entry in a Python dictionary with a key unique to the combination of parameters. This process is greatly simplified by using the `update()` method of the hmf package.

To enable later retrieval of the data for plotting and/or addition of extra data, the input parameters and calculated HMFs are saved to the session object, which runs continuously throughout the course of a user's session. It can be accessed by any view and provides a convenient way to transfer data. This is used as sparingly as possible, however, because it increases overall memory usage.

In terms of presentation, we utilise Django's templating system, combined with the *bootstrap*<sup>5</sup> CSS framework. This system allows for *template inheritance* – base templates can be written that define certain blocks that sub-templates may alter. This is convenient because we define a single base template for the whole site, which provides the general layout, including the navigation bar. The template for each page needs only to specify the unique content, which simplifies the design greatly. We use the *bootstrap* framework as it provides a simple means to create an aesthetically pleasing design.

Output and presentation of the calculations is done by means of the creation of figures and generation of downloadable ascii data. For both of these options, all files and resources are served in memory – nothing is written to disk on the server. This enhances performance and increases simplicity server-side. We have chosen to use `matplotlib` (Hunter, 2007), the standard scientific plotting library of Python, to drive the plotting for the application. This is used to produce PNG images for dynamic display on the web, and PDF images for optional download. Currently the plotting process is quite static – the data itself is the only dynamic aspect. This may result in plots that are not drawn in the best possible way for a given set of parameters. However, it is expected that those who wish to plot many different combinations will likely download the data file and come to their own plotting solution.

### 2.3.3 Usage

Using `HMFcalc` is straightforward – generation of a HMF is as easy as pressing "Calculate!" because the web form is populated with default parameters. Submission of the form generates a page with a plot of the calculated HMF with relevant legend, as well as a drop-down menu for choosing the particular additional functions to be displayed. These include the mass variance, the fitting functions, the cumulative HMF and the power spectrum. Furthermore, an additional drop-down menu allows the user to download the currently viewed plot, all calculated plots, or data files from the calculation. Below the plot additional cosmographic quantities that have been calculated are presented (e.g. age of Universe at redshift  $z$ , distances, growth factors, etc...). In the following subsections we provide a glossary of parameters used in `HMFcalc`.

---

<sup>5</sup> <http://twitter.github.io/bootstrap/>

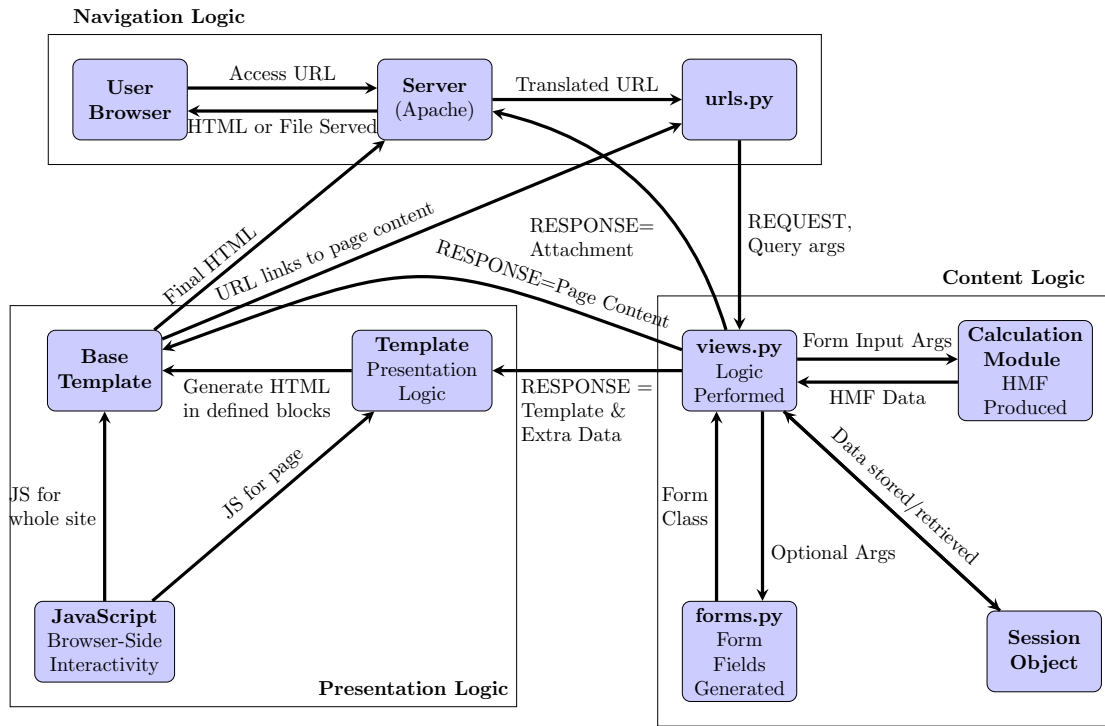


Figure 2.5: Flowchart depicting the several components of this django application and how they interact

### 2.3.3.1 Parameters

The first category of parameters are the ‘run parameters’, which control the calculation (cf. Table 2.1). Where multiple parameters are possible, all combinations of parameters will be used in the calculation. This means that the number of HMFs generated can rise rapidly and so we place a limit on the total number of possible combinations.

The second category of parameters are the ‘cosmological parameters’ (cf. Table 2.2). The key parameter is the transfer function, which is invoked in one of three modes;

1. Pre-calculated (default); there are several transfer functions pre-calculated for commonly used cosmologies, each produced at high resolution by CAMB.
2. Custom; the user may upload their own transfer function, assumed to have been produced either using CAMB or in CAMB format.
3. On-the-fly; CAMB is run at invocation to produce the required transfer function.

If a pre-calculated transfer function is chosen, all cosmological parameters are set as read-only on the form; if a custom transfer function is used, the user must ensure that the cosmological parameters entered in the web-app match those used to generate the transfer function.

In Table 2.2, we use T, H and B to indicate whether or not a parameter affects the transfer function (T), the HMF (H) or both (B). Note that those that affect both need special care if the user uploads their own transfer function.

Table 2.1: Description of Run Parameters

PARAMETER	MULT?	DESCRIPTION
Redshift	✓	Epoch at which the HMF is calculated; default is $z=0$ . Affects HMF through growth parameter (cf. §2.2.2), quick to update. Most recent fits contain explicit redshift dependence.
Halo Overdensity $\Delta_h$	✓	Defines the extent of a halo (cf. §2.2.1; default is 200. Note that at present, only the Tinker et al. (2008) and Watson et al. (2013) HMFs depend on $\Delta_h$ .
$\Delta_h$ with respect to	✗	Defines whether $\Delta_h$ is with respect to mean or critical density; default mean density.
WDM Mass, $m_X$	✓	WDM particle mass in single-species models, in units of $\text{keV}/c^2$ ; default is 0. This is used to calculate the WDM transfer function and consequently the power spectrum (cf. Eq 2.9).
Fitting Function	✓	Defines which fit from the literature to use. See §2.2.1 for more details.
Custom Fitting Function	✗	User-defined fitting function, written in Python syntax as a string, with independent variable named $x$ corresponding to the mass variance; default is empty. To add more complex functions, the user will need to edit the source code.
Min & Max $k$	✓	Minimum and maximum wavenumber ( $k$ ) for the extrapolated power spectrum; defaults are $[10^{-8}, 2000.0] h\text{Mpc}^{-1}$ for accurate prediction (cf. §2.2)
Min & Max Mass	✗	Defines minimum and maximum halo mass bounds for plotted HMF; defaults are 8 and 15 in units of $\log_{10} h^{-1} M_\odot$ .
Mass Bin Width	✗	Defines logarithmic mass bin width; default is 0.05 dex.
Restrict Mass Range?	✗	Option to restrict returned functions to mass range of fit (cf. Table 2.3). Default True.

Table 2.2: Description of Cosmological Parameters

PARAMETER	AFFECTS	DESCRIPTION
Transfer Function	–	Choose from drop-down menu. Either pre-calculated and based on popular cosmologies or customised based on the available parameters.
HMF Labels	–	Must be unique – one for each parameter set plotted. Appears in plots as legends and in downloadable data as column headers.
$\delta_c$	H	Critical overdensity for collapse. Although cosmology-dependent, we assume the Einstein-de Sitter value of 1.686 as default.
$n_s$	H	Spectral index of power spectrum – see §2.2).
$\sigma_8$	H	Mass variance on scale of $8 h^{-1}\text{Mpc}$ at $z=0$ used to normalise power spectrum – (see §2.2).
$H_0$	T	Hubble Parameter at $z=0$ .
$\Omega_b$	B	Baryon density at $z=0$ .
$\Omega_c$	B	Dark matter density at $z=0$ .
$\Omega_\Lambda$	B	Dark energy density at $z=0$ .

### 2.3.3.2 An Important Caveat

HMFcalc will calculate a HMF given any set of cosmological parameters and fitting function. However, most HMFs are empirical fits to the results of cosmological  $N$ -body simulations and are tied to specific cosmologies (but see, e.g. Jenkins et al., 2001; Courtin et al., 2010); whether or not the HMF is universal remains a topic of active research (see discussion in Tinker et al. 2008, Bhattacharya et al. 2011 and Murray, Power, and Robotham 2013b). For this reason we caution the careful user to compare the results of HMFcalc against  $N$ -body simulations of the appropriate cosmology.

### 2.3.4 Comparison To Other Codes

A small number of public codes are available for calculating the HMF – eg. genmf<sup>6</sup>, MF\_Code<sup>7</sup>, and the HMF calculator in the yt project<sup>8</sup>. We have tested hmf against these codes and find excellent consistency. In fact, bundled with the hmf package is a test-suite which ensures consistency to the per-cent level with the commonly used genmf. However, we note that HMFcalc has a number of advantages over these codes; it is inherently visually interactive and cross-platform, it allows for dynamic accurate calculation of the transfer function with CAMB and efficient and self-consistent parameter updates, and it can be easily extended by the user with bespoke fitting functions.

<sup>6</sup> [http://icc.dur.ac.uk/Research/PublicDownloads/genmf\\_v1.01.tar](http://icc.dur.ac.uk/Research/PublicDownloads/genmf_v1.01.tar)

<sup>7</sup> [http://cosmo.nyu.edu/~tinker/massfunction/MF\\_code.tar](http://cosmo.nyu.edu/~tinker/massfunction/MF_code.tar)

<sup>8</sup> <http://yt-project.org/>

## 2.4 EXAMPLE APPLICATIONS

We have designed `HMFCalc` to be a useful tool for observers and theorists alike, and we expect the primary use of `HMFCalc` to be the quick and straightforward generation of HMFs for different fitting functions and as a function of cosmological parameters and cosmic epoch. However, `HMFCalc` should be a particularly useful aid to those who work with cosmological simulations, and we demonstrate how with two examples.

### 2.4.1 Box Size for One Halo of Mass $M$

A cosmological simulation follows the growth of structure in a periodic cubic box of volume  $L^3 \text{Mpc}^3 h^{-3}$  containing  $N_p$  particles. In principle  $L$  can be arbitrarily large, but normally it's set by requiring that the smallest resolved haloes contain a minimum number of particles, or equivalently the particle mass  $m_p \propto L^3 / N_p$ . Given this restriction on  $L$  enforced by  $m_p$ , it is interesting to ask how large the periodic box needs to be to contain at least one halo of mass  $> M$ . Clearly, the answer is probabilistic, and so to be more precise, we may ask what size the box must be to induce a probability  $P$  of including at least one halo of mass  $> M$ .

There are two contributions to the total variance of the number densities within a box – Poisson shot-noise and sample variance. Since the first is assumed to be Poisson, its variance is equal to the mean:  $\sigma_{\text{SN}}^2 = n(> M)L^3$ . The sample variance requires a knowledge of the *halo bias*, which is not built in to `HMF` (but see chapter 5 for its inclusion into `HALOMOD`). Thus, at first we assume negligible sample variance, and determine the probability based only on shot-noise, also showing how such a calculation is included in the `HMF` library. Following this, we will present the full result including sample variance.

#### 2.4.1.1 Shot-noise

The shot-noise contribution is the upper CDF of the Poisson distribution with expectation  $n(> M)L^3$ :

$$P(n \geq 1) = \Gamma(2, n(> M)L^3) = P, \quad (2.14)$$

where  $\Gamma$  is the upper-incomplete gamma function. In this case the identity  $\Gamma(2, x) = 1 - e^{-x}(1 + x)$  holds, and after inversion (involving the product-log function), yields the result  $x = 4.74386$  for  $P = 0.95$ .

Thus we require

$$n(> M)L^3 = 0.366362 \rightarrow L = \left( \frac{0.366362}{n(> M)} \right)^{1/3}. \quad (2.15)$$

We have chosen to create a method in `Perturbations` (from the `hmf` backend) to derive Eq 2.15. We add the following code to the class:

```
@property
def how_big(self):
    try:
```

```

        return self.__how_big
    except:
        return (0.366362 / self.ngtm)**(1./3.)

@how_big.deleter
def how_big(self):
    try:
        del self.__how_big
    except:
        pass

```

along with

```
del self.how_big
```

under the `ngtm` deleter method. This is all that is needed to use `hmf` to calculate Eq 2.15. Note the `@property` decorator ensures that the value of `how_big` will change as the value of `ngtm` is modified. We extend the web interface to output this information by adding the following to the function `hmf_driver()` in `utils.py` towards the end of the loops;

```

if 'get_L' in extra_plots:
    mass_data["L(N=1)_" + getname(labels)] = pert.how_big

```

Here, the `extra_plots` variable is a list that contains the names of the extra plots to be calculated (defined in the form later). The function is saved to the existing `mass_data` array, under a label beginning with `L(N=1)_`, which describes the “box size needed for one halo”. The `Perturbations()` object has been previously initialized as `pert`, and thus all we must do is access its `how_big` property.

As we have noted, we must implement the choice of calculating this function on the form. To do this we add the entry

```
('get_L', 'Box Size for One Halo'),
```

to the `optional_plots` list variable in the `HMFInput()` class in `forms.py`. Its information will be passed to the form template, which will render it automatically in the ‘Optional Extra Plots’ fieldset.

We also need to create a conditional entry for the function in the drop-down menu of the results page, so that we can view the plot. To do this we add the following code to the `__init__()` method of the `PlotChoices()` class of `forms.py`:

```

if 'get_L' in session_plots:
    extra_plots.append(('L',
                        'Box Size for One Halo'))

```

Here we add the choice only if `get_L` was checked on the initial input form. The first argument is the name of the choice internally, while the second is the label that the user sees.

The only remaining task is to create the plot definition for the function, in `views.py`. We add the following code to produce the required axis labels and extract the  $L(N=1)$  function from the whole dataset:

```

elif plottype == 'L':
    keep=[string for string in mass_data \

```

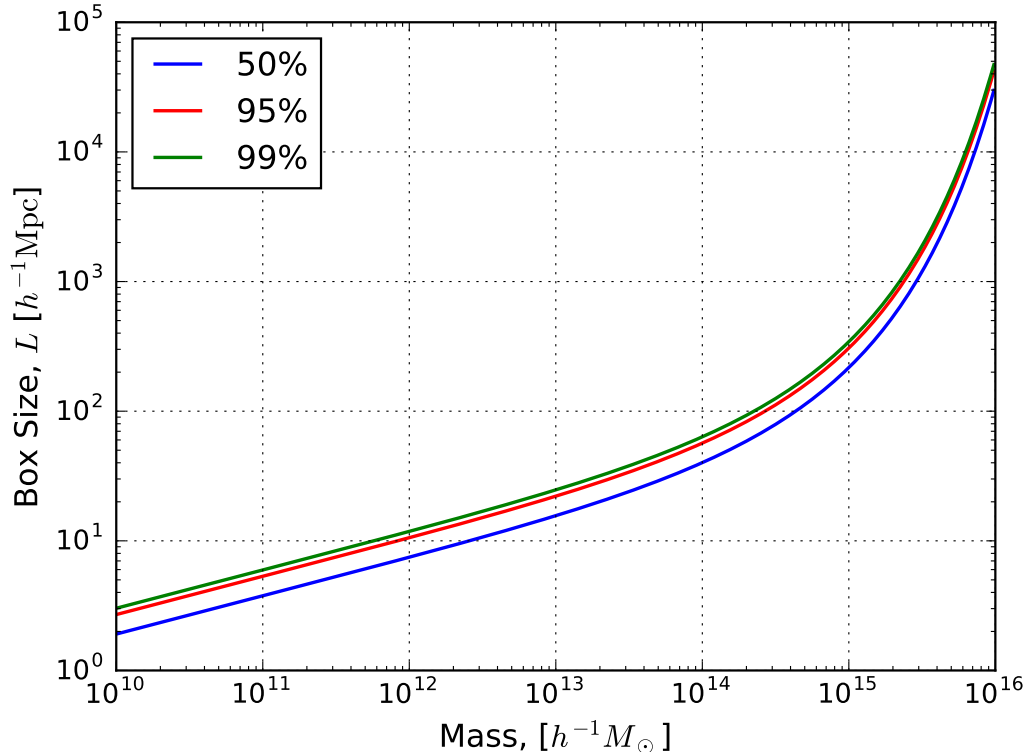


Figure 2.6: Box size required to expect at least one halo of mass  $> M$  with varying probability. This shows the results for a WMAP7 cosmology at  $z = 0$ .

```

if string.startswith('L(N=1)-')
title="Box Size, L, for One Halo"
ylab= "Box Size, L (Mpc/h)"
yscale = 'log'

```

and add the plottype to the first list, `mass_plots`, specifying that the  $x$ -axis will be based on the mass grid. Here the plottype refers to the internal label from the `PlotChoice()` form. The first line chooses all the columns whose labels begin with the specified string, corresponding to our custom function<sup>9</sup>.

The result, for the WMAP7 cosmology at  $z=0$ , is shown in Fig 2.6; it is computed by default by the web-app. We note that current large-scale simulations typically have  $L \sim 10^3 h-1\text{Mpc}$ , so we should not expect haloes of mass greater than  $10^{16} h-1\text{Mpc}$ .

We have also determined how many haloes are expected in a simulation of box size  $L$  (see Fig 2.7). This is not currently implemented in the web-app, but it will be in future versions.

<sup>9</sup> Note that all code samples given here apply to the codes `HMF` and `HMFcalc` at the time of publishing this chapter as an article. Subsequent modifications have retained the general conceptual framework however

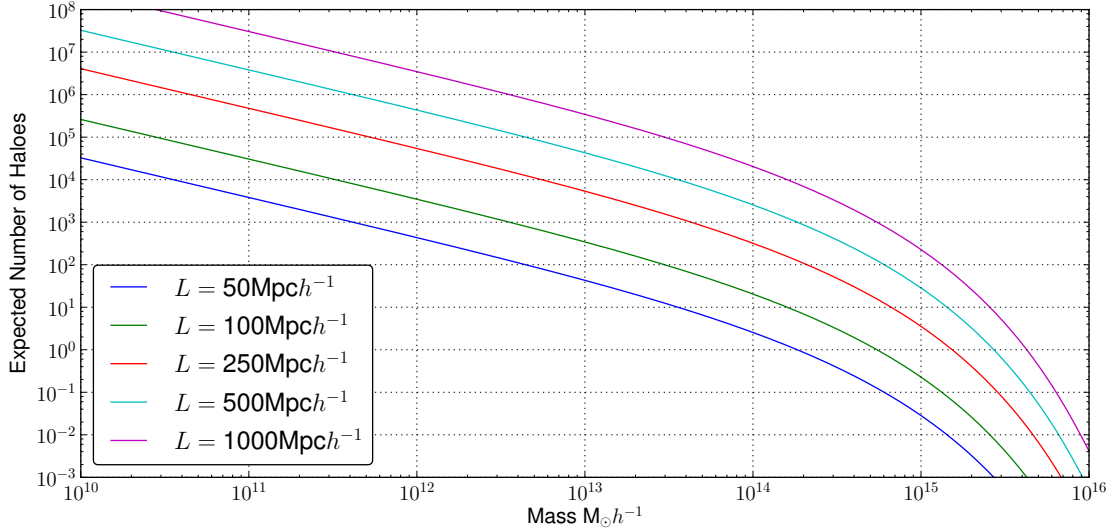


Figure 2.7: How many haloes do we expect of mass  $> M$  in a simulation box of size  $L$ ? This shows the result for a WMAP7 cosmology at  $z=0$ .

#### 2.4.1.2 Sample Variance

Sample variance in the number counts for a population of collapsed structures has been extensively studied in Hu and Kravtsov (2003). In general, for a mass bin  $M_i$  at redshift  $z = 0$ , the sample variance on scale  $R$  is

$$\sigma_n^2 = \frac{b^2(M_i)}{(2\pi)^3} \int_0^\infty P(k) W^2(k) d^3k, \quad (2.16)$$

where  $b(M)$  is the bias of a halo of mass  $M$  with respect to the underlying matter field (we use the form of Tinker and Weinberg 2005). Within a spherical region of radius  $R$ , this collapses, via Eq. 2.3, to

$$\sigma_n^2 = b^2(M_i) \sigma^2(R), \quad (2.17)$$

i.e. proportional to the linear rms variance in top-hat spheres. However, in the context of periodic cosmological simulations, three modifications are in order.

Firstly, the periodic boundary conditions of the volume mean that the integral should only be performed from  $k_{\min} = 2\pi/L$  (cf. next subsection) to  $k_{\max} = 2\pi/(L/N^{1/3})$ . For simplicity we set the latter to infinity, as it should not affect the results of this analysis (since we are focussing on large-scale fluctuations).

Secondly the bias given here is at a given mass, whereas our calculation is cumulative in mass. Thus, we replace this with the *effective bias* (cf. 5.24):

$$b_{\text{eff}} = \frac{1}{n(> M)} \int_M \frac{dn}{dm} b(m) dm. \quad (2.18)$$

Finally, the simulation box is not a spherical region, but cubic. While one can use a boxcar function for  $W(k)$  in Eq. 2.16, for the sake of simplicity we assume that the variance within a box of side  $L$  is equivalent to the variance within a sphere of equal



volume, i.e. with  $R = L(3/4\pi)^{1/3}$ . We expect the effect of this approximation to be minimal. Thus we arrive at (after addition of the shot-noise):

$$\sigma_{n(>M)}^2(L) = n(>M)L^3 + \frac{b_{\text{eff}}^2(M)}{2\pi^2} \int_{2\pi/L}^{\infty} k^2 P(k) W\left(kL \left(\frac{4\pi}{3}\right)^{1/3}\right) dk. \quad (2.19)$$

If we make the simplification that the distribution of counts is Gaussian (clearly a simplification since it must be an integer distribution, but has the necessary property that the variance be arbitrarily larger than the mean; a more accurate choice for future work might be the negative binomial distribution), then  $L$  can be calculated via the CDF of the Normal distribution:

$$\frac{1 - n(>M)L^3}{\sqrt{n(>M)L^3 + b_{\text{eff}}^2\sigma_L^2}} = \text{erf}^{-1}(2(1 - P) - 1). \quad (2.20)$$

This equation must be numerically solved for each  $M$ .

However, we simply note that for any  $P > 0.5$ , the RHS is negative, which requires at a minimum that  $L > 1/n(>M)^{1/3}$ . For the smallest haloes considered here –  $10^{10}h^{-1}M_{\odot}$  – this requires that  $L \gtrsim 1.6h^{-1}\text{Mpc}$ . For an *infinite* box,  $\sigma_L$  about 2-3 at this scale, but for a box with periodic boundary conditions of this size, some of this variance disappears because for example the mean density is set to be  $\Omega_m$  (this comes about through the integral bounds in Eq. 2.19). Thus in practice, the *largest possible* value of the sample variance is of order  $10^{-3}$ , three orders of magnitude smaller than the effect of shot noise. We can thus safely ignore it, and refer directly to figure 2.6.

#### 2.4.2 The Impact of Finite Box Size

Cosmological simulations model the growth of structure over cosmic time by following the evolution of matter density perturbations in a periodic box. Because the box has a finite size  $L$ , the longest wavelength perturbation that can be reliably modelled will have a wavenumber  $k_{\min}$  of

$$k_{\min} = \frac{2\pi}{L}. \quad (2.21)$$

This means that the effect of perturbations with wavelengths longer than the box size cannot be captured, and implies that the mass variance we would expect (cf. Eq 2.3) will be larger than the one that we measure. This effect, and its influence on the HMF at higher masses, has been noted in previous studies (e.g Bagla and Ray, 2005; Power and Knebe, 2006; Lukić et al., 2007).

We can use `HMFcalc` to quantify this effect. We choose box sizes of size 32, 43, 64 and  $128 h^{-1}\text{Mpc}$ , or  $k_{\min} = 0.1257, 0.0628, 0.02513$  and  $0.01257 h\text{Mpc}^{-1}$ , corresponding to the box-sizes presented in Power and Knebe (2006). We use the inbuilt functionality of `HMFcalc` to enter as input all of these minimum  $k$  bounds at once, plus a baseline of  $k_{\min} = 10^{-8}$  that corresponds to the limit of infinite box size. The result is shown in the left panel of figure 2.8, with the right panel displaying figure 4 from Power and Knebe (2006), which contains the same information.

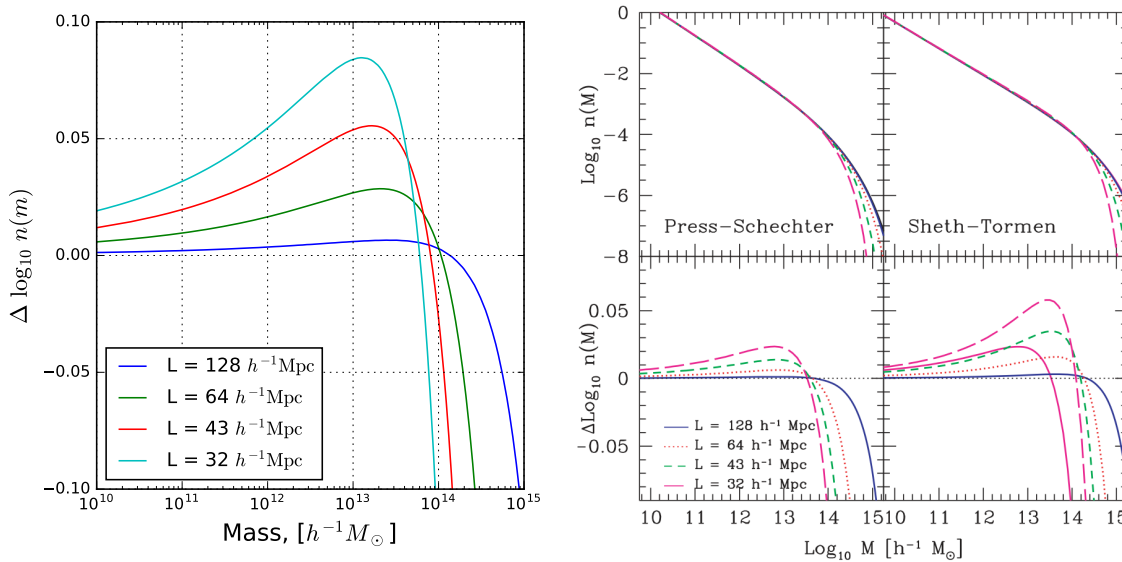


Figure 2.8: The effect of finite box-size on the HMF. Left panel shows results of HMFcalc and right panel shows figure 4 from Power and Knebe (2006). Both panels plot the same box-sizes, and we use the fit of Sheth and Tormen (1999) in HMFcalc for comparison. Results are similar in both cases, though the exact values vary, likely due to small calculation differences.

There are two notable effects of increasing  $k_{\min}$ . The first is that the high-mass haloes are increasingly suppressed with smaller box-size, and the second is that low-mass haloes are increasingly enhanced. In all cases, the very lowest masses are asymptotically unaffected<sup>10</sup>. The physical origin of the high-mass suppression is the artificial lack of large-wavelength modes which are coupled with shorter wavelength modes during nonlinear structure formation. The removal of these modes prevents the formation of the deepest potential wells at late times. These wells become shallower, and migrate down the mass function to provide the enhancement at intermediate scales.

These results have strong implications for the construction of simulations as tests of cluster counts, and especially those that must simultaneously model small-scale structures reliably. In particular, they prescribe a minimum box-size for the reconstruction of halo counts at a given mass for a given accuracy. With a fixed volume and cosmology, the number of particles is inversely proportional to the particle mass, so that to probe a given mass scale more particles are required for a larger volume.

Conversely, the effects should not be present in limited-volume surveys, as they arise in simulations entirely due to the periodic nature of the box, which serves to neutralise all modes longer than the box-size. In observations, while the observed volume is finite, super-sample modes are able to affect the distribution (Hu and Kravtsov, 2003; Takada et al., 2013). In fact, these results and the effect of “super-sample covariance” as described by Takada and Hu (2013) can be seen as difficulties lying on opposite sides of the coin: the deterministic suppression presented here affects simulations but not observations, while the stochastic effects of super-sample models affect observations but

<sup>10</sup> In the paper version of this chapter (Murray, Power, and Robotham, 2013a), the corresponding plot shows a decrement for low-mass haloes. This arose because the normalisation of the power spectrum was calculated using  $\sigma_{R=8}$  derived with the truncated  $k$ -limits. This behaviour has since been rectified in HMFcalc.

not simulations. Both must be accounted for to rigorously compare observations with simulations.

Finally, we comment on the similarity and differences between the left- and right-hand panels of figure 2.8. While the broad characteristics of the two are similar, the precise numbers differ by up to  $\Delta \log_{10} n(m) = 0.02 - 0.03$ , or  $\sim 7\%$  in terms of difference of ratios. In both cases, the same fitting function (that of Sheth and Tormen, 1999) has been used. The cosmology is different, but we have found that setting our cosmology equal to that of Power and Knebe (2006) actually enhances the differences. Beyond this, there is little we can do to ascertain the source of the discrepancy. This shows how important it is to have an open and reliable standard for these kinds of computations.

## 2.5 FUTURE DEVELOPMENT

### 2.5.1 *Extending Range of Dark Matter and Dark Energy Models*

In the current version of `HMFcalc`, we have considered only the fiducial Cold Dark Matter model and a straightforward extension to model Warm Dark Matter, using the framework set out in Bode, Ostriker, and Turok (2001). In forthcoming versions of `HMFcalc`, we shall update the framework to account for more recent developments in the modelling of Warm Dark Matter models (e.g. Schneider, Smith, and Reed, 2013; Pacucci, Mesinger, and Haiman, 2013; Benson et al., 2012), as well as alternative dark energy models such as those explored in, for example, Francis, Lewis, and Linder (2009) and Courtin et al. (2010).

### 2.5.2 *Calibrating Synthetic Galaxy Surveys*

“Halo Occupation Distribution” (HOD) models parameterise how likely galaxies are to occupy dark matter haloes and provide a convenient framework for creating synthetic galaxy surveys. Formally, the HOD defines a probability  $P(N|M)$  of finding  $N$  luminous galaxies in a halo of mass  $M$  (Berlind and Weinberg, 2002), and it allows a population of dark matter haloes to be mapped to a synthetic galaxy population whose abundance and spatial clustering can be tuned to reproduce the observed galaxy population. We will extend `HMFcalc` so that we can take as input a given set of cosmological parameters, a dark matter and dark energy model, and a set of observational constraints (e.g. 2-point correlation function in different wavelength bands) and calibrate HOD parameters to reproduce the properties of the observed galaxy population. With this framework we will develop empirically-motivated synthetic galaxy surveys using both the outputs from both cosmological  $N$ -body simulations (e.g. Springel, White, and Jenkins, 2005) as well as faster, more approximate schemes such as 2<sup>nd</sup> order Lagrangian perturbation theory (e.g. Manera et al., 2012), vital in the coming era of large galaxy surveys.

### 2.5.3 *Dynamic and Adaptable User Interface*

We are developing HMFcalc into a dynamic and adaptable online tool that will give the user finer and more varied control over the generation of data. For example, we would like users to be able to explore how variations in the cosmological parameters influence the HMF and to see the resulting HMF update in real-time. A similar framework will be used to allow users to explore how HOD-derived galaxy population observables vary with changes in cosmological parameters and dark matter/dark energy parameters. We would also like users to be able to explore how, for example, changing from the traditional top-hat window function (cf. Eq 2.4) to sharp- $k$  filter window functions (cf. Schneider, Smith, and Reed, 2013) influence the predicted HMF, especially in WDM models. Such an online toolbox will provide an invaluable community resource for both observers and theorists alike.

## 2.6 SUMMARY

We have presented HMFcalc, a flexible, efficient and easy to use web application for calculating the dark matter halo mass function (HMF), a fundamental property of cosmological structure formation models. HMFcalc and its engine hmf are implemented in Python, which offers a simple yet powerful and extensible cross-platform method for building community software. Observers and theorists alike should find HMFcalc a valuable resource to rapidly explore the effects of cosmological parameters, redshift and fitting function on the predicted HMF, while cosmological simulators will find it a useful tool for understanding, for example, the biases introduced by finite simulation volume. In future work we will extend HMFcalc to give the user greater flexibility in how they interact with their data; we will broaden the range of dark matter and dark energy models considered; and finally, we will use the framework provided by HMFcalc to develop the infrastructure for fast approximate methods for generating bespoke synthetic galaxy surveys.

Table 2.3: Compilation of Fitting Functions: Note that we define  $\nu = \delta_c / \sigma$ .

REFERENCE	FITTING FUNCTION $f(\sigma)$	MASS RANGE	REDSHIFT RANGE	COSMOLOGY	HALO DEF. (b and/or $\Delta_h$ )
Press and Schechter (1974)	$f_{\text{PS}}(\sigma) = \sqrt{\frac{2}{\pi}} \nu \exp[-nu^2/2]$	-	-	-	-
Sheth, Mo, and Tormen (2001)	$f_{\text{SMT}}(\sigma) = A \sqrt{\frac{2a}{\pi}} \left[ 1 + (\nu^2/a)^p \right] \nu \exp[-a\nu^2/2],$ $A = 0.3222, a = 0.707, p = 0.3.$	$0.5 < \nu^2 < 10$	0 – 2	$\Lambda\text{CDM, OCDM, SCDM}$	$b = 0.2$
Jenkins et al. (2001)	$f_{\text{J}}(\sigma) = 0.315 \exp[- \ln \sigma^{-1} + 0.61 ^{3.8}]$	$-1.2 < \ln \sigma^{-1} < 1.05$	0 – 5	$\tau\text{CDM, } \Lambda\text{CDM}$	$b = 0.2, 0.164, \Delta_h = 200, 32$
Reed et al. (2003)	$f_{\text{R03}}(\sigma) = f_{\text{SMT}}(\sigma) \exp\left[\frac{-0.7}{\sigma \cosh(2\sigma)^5}\right]$	$-1.7 < \ln \sigma^{-1} < 0.9$	0 – 15	$\Omega_M = 0.3, \Omega_\Lambda = 0.7$	$b = 0.2, \Delta_h = \Delta_{\text{vir}}$
Warren et al. (2006)	$f_{\text{W}}(\sigma) = 0.7234 (\sigma^{-1.625} + 0.2538) \exp[-1.1982\sigma^{-2}]$	$10^{10} \text{M}_\odot < M < 10^{15} \text{M}_\odot$	0	$\Lambda\text{CDM: WMAP1}$	$b = 0.2$
Reed et al. (2007)	$f_{\text{R07}}(\sigma) = \nu \exp\left[-\frac{cav^2}{2} - \frac{0.03(\nu)^{0.6}}{(n_{\text{eff}}+3)^2}\right]$ $\times A \sqrt{\frac{2a}{\pi}} \left[ 1 + (\nu^2 a)^{-p} + 0.6G_1(\sigma) + 0.4G_2(\sigma) \right]$ $n_{\text{eff}} = 6 \frac{d \log \sigma^{-1}}{d \log M} - 3,$ $G_1(\sigma) = \exp[-(\ln \sigma^{-1} - 0.4)^2/0.72],$ $G_2(\sigma) = \exp[-(\ln \sigma^{-1} - 0.75)^2/0.08]$	$-1.7 < \ln \sigma^{-1} < 0.9$	0 – 30	$\Lambda\text{CDM: WMAP1}$	$b = 0.2$
Peacock (2007)	$f_{\text{P}}(\sigma) = \frac{\nu \exp(-c\nu^2)}{(1+av^b)^2} [bav^{b-1} + 2cv(1+av^b)], a = 1.529,$ $b = 0.704, c = 0.412$	$10^{10} \text{M}_\odot < M < 10^{15} \text{M}_\odot$	0	$\Lambda\text{CDM: WMAP1}$	$b = 0.2$

Table 2.3: continued...

REF.	FITTING FUNCTION $f(\sigma)$	MASS RANGE	REDSHIFT RANGE	COSMOLOGY	HALO DEF. $(b$ and/or $\Delta_h)$
Tinker et al. (2008)	$f_T(\sigma, z) = A \left( \left( \frac{b}{\sigma} \right)^a + 1 \right) \exp[-c\sigma^{-2}]$ , $A = 0.186 (1+z)^{-0.14}$ , $a = 1.47 (1+z)^{-0.06}$ , $b = 2.57 (1+z)^{-\alpha}$ , $c = 1.19$ , $\alpha = \exp \left[ - \left( \frac{0.75}{\log_{10}(\Delta_h/75)} \right)^{1.2} \right]$	$-0.6 < \log_{10} \sigma^{-1} < 0.4$	0 – 2.5	$\Lambda$ CDM: WMAP1, WMAP3+	$\Delta_h = 200, 300, 400,$ 600, 800, 1200, 1600, 2400
Crocce et al. (2010)	$f_{Cr}(\sigma) = A (\sigma^{-a} + b) \exp[-c\sigma^{-2}]$ , $A = 0.58 (1+z)^{-0.13}$ , $a = 1.37 (1+z)^{-0.15}$ , $b = 0.3 (1+z)^{-0.084}$ , $c = 1.036 (1+z)^{-0.024}$	$10^{10.5} M_\odot < M < 10^{15.5} M_\odot$	0 – 2	$(\Omega_M, \Omega_\Lambda, n, h, \sigma_8) = (0.25, 0.75, 0.95, 0.7, 0.8)$	$b = 0.2, 0.164$
Courtin et al. (2010)	$f_{Co}(\sigma) = f_{ST}(\sigma)$ , $A = 0.348$ , $a = 0.695$ , $p = 0.1$	$-0.8 < \ln \sigma^{-1} < 0.7$	0	$\Lambda$ CDM: WMAP5	$b = 0.2$
Bhattacharya et al. (2011)	$f_B(\sigma, z) = A \sqrt{\frac{2}{\pi}} \exp[-av^2/2] \left[ 1 + (av^2)^{-p} \right] (v^2 \sqrt{a})^q$ , $A = 0.333 (1+z)^{-0.11}$ , $a = 0.788 (1+z)^{-0.01}$ , $p = 0.807$ , $q = 1.795$	$10^{11.8} M_\odot < M < 10^{15.5} M_\odot$	0 – 2	$w$ CDM+	$b = 0.2$
Angulo et al. (2012)	$f_A(\sigma) = A \left[ \left( \frac{b}{\sigma} \right)^a + 1 \right] \exp[-c\sigma^{-2}]$ , $(A, a, b, c) = (0.201, 1.7, 2.08, 1.172)$ or $(A, a, b, c)_{SUB} = (0.265, 1.9, 1.675, 1.4)$	$10^8 M_\odot < M < 10^{16} M_\odot$	0	$\Lambda$ CDM: WMAP1	$b = 0.2$
Watson et al. (2013)	$f_W^{FOF}(\sigma, z) = f_T(\sigma, z)$ , $A = 0.282$ , $a = 2.163$ , $b = 1.406$ , $c = 1.21$	$-0.55 < \ln \sigma^{-1} < 1.31$	0 – 30	$\Lambda$ CDM: WMAP5	$b = 0.2$

Table 2.3: continued...

REF.	FITTING FUNCTION $f(\sigma)$	MASS RANGE	REDSHIFT RANGE	COSMOLOGY	HALO DEF. and/or $\Delta_h$	( $b$ )
Watson et al. (2013)	$f_W^{\text{SO}}(\sigma, z) = \Gamma(\Delta, \sigma, z) f_{\text{T}}(\sigma, z),$ $(A, a, b, c)_{z=0} = (0.194, 2.267, 1.805, 1.287),$ $(A, a, b, c)_{z>6} = (0.563, 0.874, 3.810, 1.453),$ $(A, a, b, c)_{0<z<6} = \Omega_M(z) \times (1.907(1+z)^{-3.216} + 0.074,$ $3.136(1+z)^{-3.058} + 2.349, 5.907 \times (1+z)^{-3.599} + 2.344,$ $1.318),$ $\Gamma(\Delta, \sigma, z) = C(\Delta) \left(\frac{\Delta}{178}\right)^{d(z)} \exp \left[ p \left(1 - \frac{\Delta}{178}\right) \sigma^{-q} \right],$ $C(\Delta) = \exp \left[ 0.023 \left( \frac{\Delta}{178} - 1 \right) \right],$ $d(z) = -0.456 \Omega_M(z) - 0.139, p = 0.072, q = 2.130.$	$-0.55 < \ln \sigma^{-1} < 1.05$ $(z = 0),$ $-0.06 < \ln \sigma^{-1} < 1.024$ $(z > 0)$	0 – 30	$\Lambda$ CDM: WMAP5	$\Delta_h = 178,$ (100–1600)	
Behroozi, Wechsler, and Conroy (2013)	$n_B(> M) = f(a) \left( M/10^{11.5} M_{\odot} \right)^{g(a)} + n_{\text{T}}(> M),$ $f(a) = \frac{0.144}{1 + \exp[14.79(a - 0.213)]}, g(a) = 0.5(1 + \exp(6.5a))^{-1},$ $a = 1/(1+z)$	$-0.6 < \log_{10} \sigma^{-1} < 0.4$	0 – 8	$(\Omega_M, \Omega_{\Lambda}, n, h,$ $\sigma_8) = (0.27, 0.73,$ $0.95, 0.7, 0.82)$	$\Delta_h = 200$	





---

HOW WELL DO WE KNOW THE HALO MASS FUNCTION?

---

## ABSTRACT

The parameters governing the standard  $\Lambda$  Cold Dark Matter cosmological model have been constrained with unprecedented accuracy by precise measurements of the cosmic microwave background by the Wilkinson Microwave Anisotropy Probe (WMAP) and Planck satellites. Each new data release has refined further our knowledge of quantities – such as the matter density parameter  $\Omega_M$  – that are imprinted on the dark matter halo mass function (HMF), a powerful probe of dark matter and dark energy models. In this letter we trace how changes in the cosmological parameters over the last decade have influenced uncertainty in our knowledge of the HMF. We show that this uncertainty has reduced significantly since the 3<sup>rd</sup> WMAP data release, but the rate of this reduction is slowing. This is limited by uncertainty in the normalisation  $\sigma_8$ , whose influence is most pronounced at the high mass end of the mass function. Interestingly, we find that the accuracy with which we can constrain the HMF in terms of the cosmological parameters has now reached the point at which it is comparable to the scatter in HMF fitting functions. This suggests that the power of the HMF as a precision probe of dark matter and dark energy hinges on more accurate determination of the theoretical HMF. Finally, we assess prospects of using the HMF to differentiate between Cold and Warm Dark Matter models based on ongoing improvements in measurements of  $\Omega_M$ , and we comment briefly on optimal survey strategies for constraining dark matter and dark energy models using the HMF.

## 3.1 INTRODUCTION

The halo mass function (hereafter HMF), which encodes the comoving number density of dark matter haloes in the Universe at a given epoch as a function of their mass, is a powerful probe of cosmology, dark matter and dark energy (Press and Schechter, 1974; Jenkins et al., 2001; Tinker et al., 2008; Vikhlinin et al., 2009). For example, the amplitude of the HMF on the scale of galaxy clusters at the present epoch may be used to deduce limits on the combination of the power spectrum normalisation  $\sigma_8$  and the matter density parameter  $\Omega_M$  (Vikhlinin et al., 2009; Allen, Evrard, and Mantz, 2011). Similarly, the evolution of this amplitude over cosmic time may be used to characterise the dark energy equation of state  $w_0$  (Vikhlinin et al., 2009; Allen, Evrard, and Mantz, 2011).

Over the last decade, precise measurements of the cosmic microwave background (CMB) by the Wilkinson Microwave Anisotropy Probe (hereafter WMAP; cf. Spergel et al., 2003) and Planck (Planck Collaboration, 2014a) satellites have led to increasingly accurate estimates of key cosmological parameters such as  $\Omega_M$ , by factors of  $\sim 4$  in most cases. It is therefore interesting to ask how our knowledge of the HMF has evolved over the same period.

In this Letter, we estimate the uncertainty in the HMF assuming “best-bet” WMAP and Planck cosmological parameters, and we determine the independent parameters that are the primary sources of this uncertainty. This is of crucial importance because quantifying the significance of any deviation between an observationally derived HMF and one predicted within the standard cosmological framework is difficult without understanding the framework’s intrinsic uncertainties. Specifically, we determine uncertainties in the predicted HMF for a suite of flat  $\Lambda$  Cold Dark Matter (hereafter  $\Lambda$ CDM) cosmologies, adopting a range of HMF fitting functions drawn from the literature. We present the 68% error on the amplitude and slope of the HMF, given the reported errors on a number of input parameters; a comparison of errors due to uncertainty in cosmology with errors in the chosen fitting function; and an analysis of the parameters that provide the primary sources of uncertainty. Finally, we consider the sensitivity of the HMF to the assumed dark matter model by exploring the range of Warm Dark Matter (WDM) particle masses for which the HMF can be used to differentiate between  $\Lambda$ CDM and its  $\Lambda$ WDM alternatives.

### 3.2 METHODOLOGY

We calculate the HMF using the formalism of Press and Schechter (1974) and Bond et al. (1991). This defines the HMF as the differential number density of haloes in logarithmic mass bins,

$$\frac{dn}{d \log M} = \frac{\rho_0}{M} f(\sigma) \left| \frac{d \ln \sigma}{d \ln M} \right|, \quad (3.1)$$

where  $\rho_0$  is the mean matter density of the universe;  $\sigma$  is the mass variance at mass scale  $M$ ; and the function  $f(\sigma)$  differentiates between fitting functions.

Many forms for  $f(\sigma)$  have been proposed in the literature. The original form proposed by Press and Schechter (1974) is the only one derived completely analytically, but makes the simplifying assumption of collisionless spherical collapse. Sheth, Mo, and Tormen (2001) proposed a form motivated by the more general assumption of ellipsoidal collapse in which three parameters were set by fitting to simulation data. To date, no other analytical form for  $f(\sigma)$  has been derived; instead, its form is empirically derived by fitting to the abundance of haloes measured by halo finders in cosmological simulations. In this Letter we adopt many of these empirical fitting functions from the literature in addition to the forms of Press and Schechter (1974) and Sheth, Mo, and Tormen (2001). For clarity, we focus on the form of Sheth-Tormen unless otherwise stated; we note that results are qualitatively similar for all of the fitting functions we have considered.

We calculate HMFs using the `hmf` code (cf. [github.com/steven-murray/hmf](https://github.com/steven-murray/hmf)), which is the backend of the `HMFcalc` web-application (cf. [hmf.icrar.org](https://hmf.icrar.org)); further details can be found in Chapter 2. Our code uses `CAMB` (cf. [camb.info](https://camb.info); details in Lewis, Challinor, and Lasenby, 2000) to produce a transfer function for a given input cosmology, and interpolates and integrates this function to compute the mass variance. Importantly, our code is optimized to quickly and easily update parameters of the calculation, which allows us to generate the large number of realisations necessary for this study.

We use data spanning the last 7 years of cosmological parameters derived from the CMB – from WMAP3 (Spergel et al., 2007), WMAP5 (Komatsu et al., 2009), WMAP9 (Hinshaw et al., 2013) and Planck (Planck Collaboration, 2014a). We choose parameter sets that derive from isolated fits to the CMB, not using any extra data (such as BAOs or lensing) to ensure consistency across samples. In all cases we assume a flat  $\Lambda$ CDM geometry, consistent with the theoretical base model of the Planck results, and consider five parameters – the baryon and CDM densities combined with the hubble constant  $\Omega_b h^2$  and  $\Omega_c h^2$ , the spectral index  $n_s$ , the Hubble parameter  $H_0$  and the normalization  $\sigma_8$ . We constrain  $H_0$  by adopting Eq. 11 of Planck Collaboration (2014a); this is possible because we can calculate directly the angular size of the sound horizon at last scattering, to better than 0.3% accuracy.

In order to sample from the parameter distributions from each base cosmology, we use available Monte Carlo Markov chains (MCMC), randomly choosing 5000 realizations from each chain. This allows for robust sampling of the covariance between each parameter. We calculate the HMF for each realization, finding the resulting 68% uncertainty about the median for both the amplitude and slope. Note that we have checked our results for consistency with more basic approaches based on (i) variance in the parameters, in which parameter uncertainties are assumed to be uncoupled, and (ii) covariance between the parameters, in which parameter uncertainties are coupled but assumed to be Gaussian-distributed. Averaging over masses, we find that the maximum difference between the variance and MCMC approaches is 6% (for WMAP3), while the maximum difference between the covariance and MCMC approaches is 2% (for Planck). If we do not average over masses, we find that the variance approach underestimates the uncertainty in the HMF by  $\sim 25\%$  ( $10\%$ ) for all WMAP (Planck) datasets at high masses ( $\sim 10^{15} h^{-1} M_\odot$ ), whereas it is  $\sim 2\%$  at most if we adopt the covariance approach.

The amplitude uncertainty range is simply the  $16^{th}$  –  $84^{th}$  quantile of the value of the HMF. Calculation of this quantity eliminates any information associated with the gradient of the mass function, which may be an important aspect in constraining cosmology. Therefore we also calculate the slope, as the arc-tangent of the discrete gradient (via the method of central differences) at each mass bin. For this analysis we use a mass bin width of 0.05 dex.

Note that we perform our analysis assuming a “universal” form for the mass function, in which changes in cosmology are captured by changes in the mass variance  $\sigma$ . Recent work suggests that this assumption does not hold in detail and can lead to  $\sim 10\%$  uncertainties (cf. Tinker et al., 2008; Bhattacharya et al., 2011). We do not explicitly account for this source of uncertainty in our analysis and we suggest that the reader allow for an additional 10% uncertainty in our quoted results to account for this.

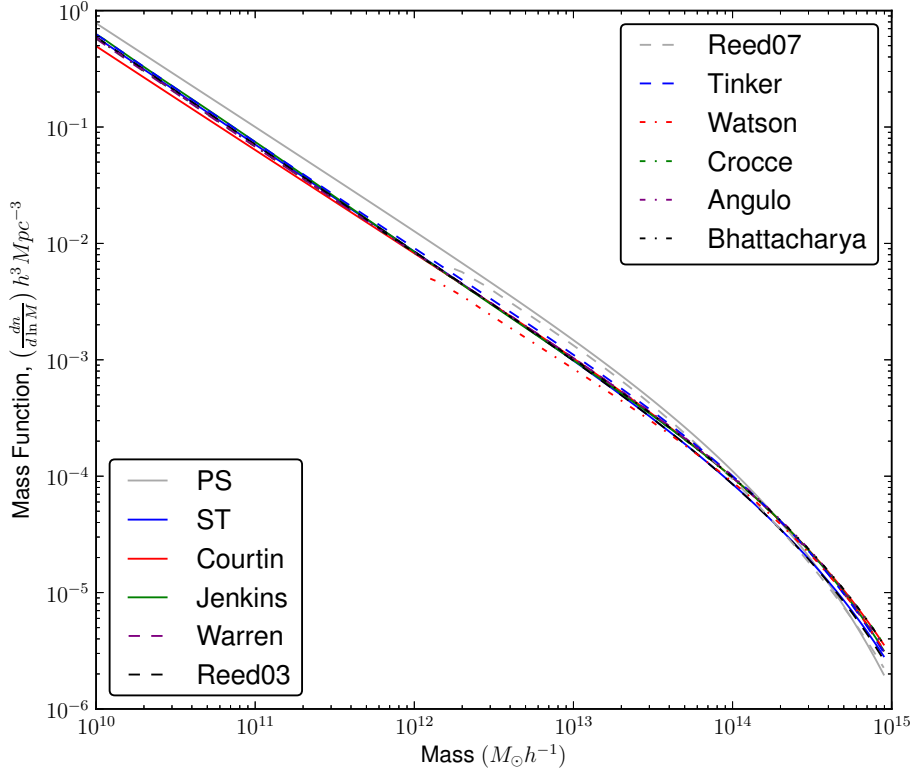


Figure 3.1: The HMF for several fitting functions drawn from the literature, using a Planck (Planck Collaboration, 2014a) cosmology; see text for further details.

### 3.3 RESULTS

**EFFECT OF CHOICE OF FITTING FUNCTION.** In Figure 3.1 we show the variety of HMF fitting functions drawn from the literature – Press and Schechter (1974), Sheth, Mo, and Tormen (2001), Jenkins et al. (2001), Reed et al. (2003), Warren et al. (2006), Reed et al. (2007), Tinker et al. (2008), Crocce et al. (2010), Courtin et al. (2010), Bhattacharya et al. (2011), and Angulo et al. (2012) and Watson et al. (2013) – assuming a Planck cosmology. The scatter between fitting functions is noticeable and has been remarked upon already in the comprehensive comparison of halo-finders presented in Knebe et al. (2011). Given this scatter, we wish to investigate two effects; (i) whether or not the fitting functions behave similarly under changes of cosmology, and (ii) whether or not the scatter in the fitting functions is the dominant source of uncertainty in the HMF for contemporary cosmological parameter sets. That is, whether the variation between fitting functions for a cosmology with no variance is greater than the variation between HMF’s of a single fitting function for a cosmology with non-zero variance.

In Figure 3.2 we show the fractional error intrinsic to each parameter set; the Sheth-Tormen fitting function is shown by the solid lines, while the fractional error attributable to the choice of fitting function, estimated at the median value of the HMF for the given cosmology, is shown by the dashed lines. It is noticeable that the variance in earlier parameter sets (e.g. WMAP3) is dominated by uncertainty in the cosmological parameters; in contrast, the variance in the most recent parameter sets is comparable to the variation between fitting functions across the mass range considered. This implies

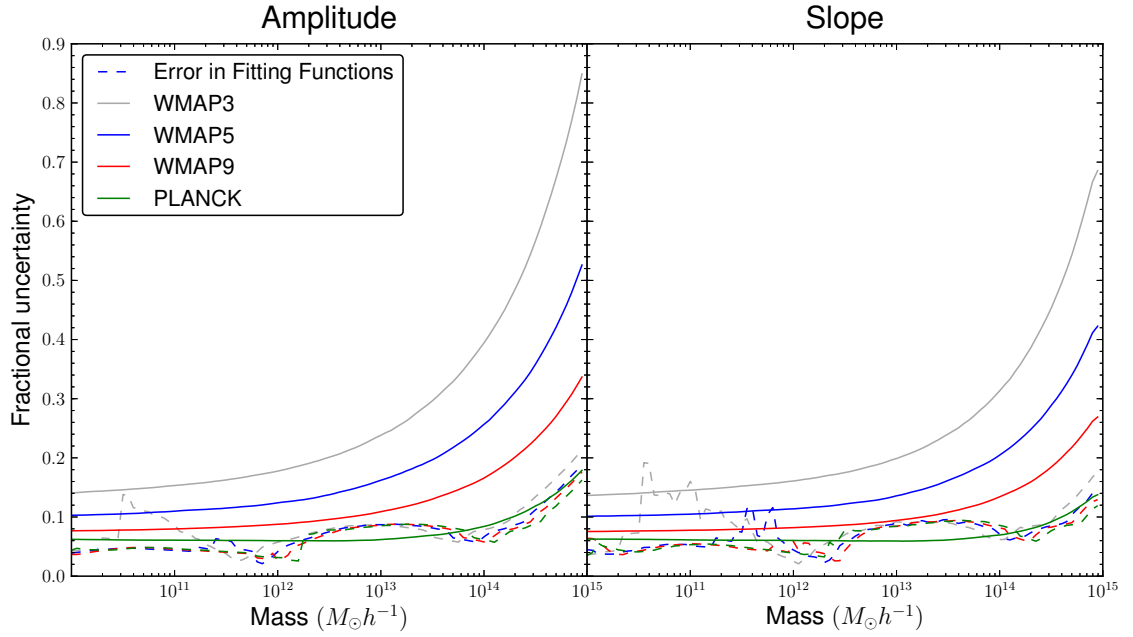


Figure 3.2: Errors on the amplitude (left) and slope (right) of the HMF for a range of cosmological parameter sets, as a fraction of the HMF. Solid lines indicate intrinsic parameter error, while dashed lines indicate error due to the choice of fitting function.

that future constraints on cosmology using the HMF will be limited by uncertainties in our knowledge of the HMF fitting function itself, unless more robust non-parametric means are used.

More encouraging is the similarity of the several dashed curves in Figure 3.2, which imply that a change of cosmology (within reasonable ranges) does not greatly affect the behaviour of the general class of fitting functions<sup>1</sup>; that is, each fitting function displays a similar trend with the change of cosmology. This derives from the fitting functions depending on cosmology only through its influence on the mass variance (cf. Eq. 3.1), and justifies our use of a single fitting function on which to base a general analysis of the cosmological parameter errors. Figure 3.3 confirms this; it shows the normalized HMF for several cosmologies for three fitting functions. While there are noticeable differences between cosmologies, there is consistency of behaviour between the fitting functions.

**AMPLITUDE OF ERRORS IN STANDARD COSMOLOGIES.** Figure 3.2 shows that at low masses, the variation asymptotes to a constant value, while at high masses the error rises seemingly exponentially. Indeed, for early parameter sets, cluster-mass haloes had associated uncertainties of almost 70%. This increase in uncertainty at high masses reflects the sensitivity of the abundance of clusters to the value of  $\sigma_8$ , which has been noted previously (Vikhlinin et al., 2009; Allen, Evrard, and Mantz, 2011) and will be discussed further. From the latest Planck measurements, we deduce a ‘rule-of-thumb’ of  $\sim 6\%$  uncertainty in both the amplitude and slope of the HMF on scales less than  $10^{13} M_\odot h^{-1}$ , rising to  $\sim 20\%$  for the amplitude and  $\sim 15\%$  for the slope at scales of  $10^{15} M_\odot h^{-1}$ .

Furthermore, although we observe a decrease in uncertainty over the consecutive WMAP and Planck measurements, the rate of decrease is becoming smaller, especially

<sup>1</sup> Even allowing for the 10% uncertainty arising from non-universality – see note in §3.2.

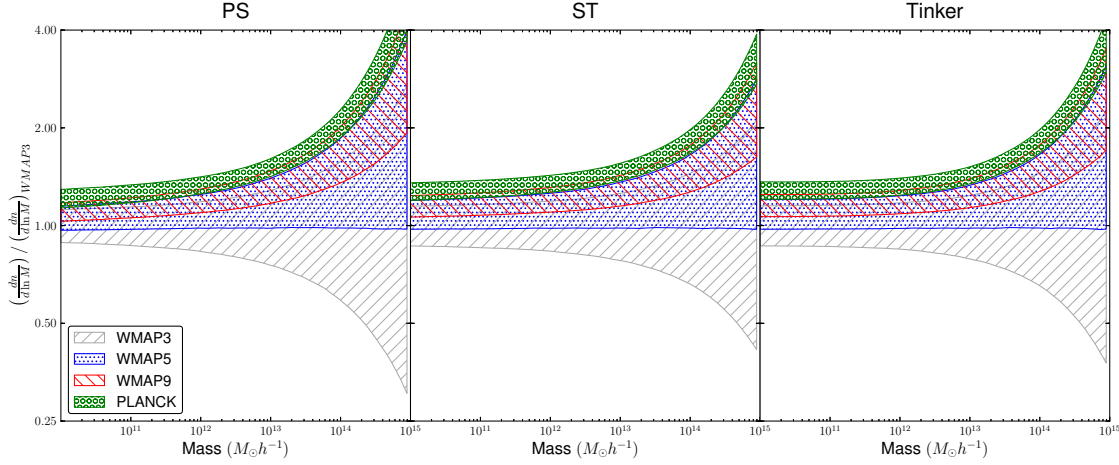


Figure 3.3: HMFs, with shaded regions indicating 68% errors, normalized by the median WMAP<sub>3</sub> HMF.

at high masses. This occurs because of the dominance of  $\sigma_8$ , which is the primary contributor to the uncertainty at high masses (cf. Figure 3.5). Therefore, a significantly more constrained HMF will depend strongly on the  $\sigma_8$  constraint. Maximum likelihood techniques applied to observed clusters may be the best way to constrain  $\sigma_8$ , given the strong sensitivity in this regime. Such a constraint could be obtained from clusters drawn from the XMM XXL X-ray survey, which is designed to identify galaxy clusters out to  $z \sim 2$  with masses of  $M_{\text{vir}} \gtrsim 10^{14} M_{\odot}$  in 50 square degrees of sky (cf. Pierre et al., 2011).

**CONSISTENCY OF PARAMETER SETS.** Figure 3.3 shows the HMF, with error ranges (given by the 16<sup>th</sup> and 84<sup>th</sup> quantiles) for several parameter sets, normalized by the HMF of the mean parameters from WMAP<sub>3</sub>. This gives a visual clue as to the relative amplitude of the HMFs and their error, which reveals whether they are statistically consistent. We see that while Planck results are marginally consistent with WMAP<sub>9</sub>, they do not overlap with WMAP<sub>5</sub> over the entire range of masses considered. This is driven by the inconsistency of  $\Omega_c h^2$  between these parameter sets.

**CONSISTENCY OF FITTING FUNCTIONS.** One notices in figure 3.2 that while uncertainties in the HMF vary widely between parameter sets, variance between different fitting functions for a single cosmology are remarkably similar between cosmologies. One may ask from where this apparent consistency across cosmology derives. What drives the scatter between fitting functions? Since each function is a fit to different numerical data, we may reason that the scatter arises due to differences in numerical choices for each of the datasets – force softening, halo definition, halo finder, initial redshift etc. – which are uncorrelated with cosmology. Thus the final scatter would be independent of cosmology, which is modelled directly through  $\sigma$ , equivalently between fits.

To explore this, we show in figure 3.4 the variation between fitting functions colour-coded by various simulation parameters, both numerical and cosmological. We note first that this comparison is simplistic – most of the fits use a suite of simulations, several of which employ different choices for the parameters between boxes. We choose to display the *maximum* box size, *minimum* softening length, *minimum*  $z_{\text{start}}$ , and *mean* for  $\sigma_8$  and



$\Omega_m$ . We note also that the comparison is incomplete – many of the fits lack proper description of the processes used in the fit, and some of the details of the simulation construction.

Despite this, we are afforded some limited insight. Firstly, halo definition is shown to be vital, with spherical-overdensity haloes having a 10-20% variation with respect to friends-of-friends haloes at  $10^{12}h^{-1}M_\odot$ . Secondly, the two largest deviants at high mass are both characterised by relatively low starting redshift using the Zeldovich approximation for initial conditions. Unfortunately, the predominance of fits have used one of a few choices of cosmological parameters (typically WMAP1 or WMAP7), which makes the variation due to cosmology difficult to ascertain.

Ultimately, the variance could be dominated by differences in the final fitting process, eg. choice of mass bin width, whether mass corrections have been applied or objective function employed. These details are often not mention in the literature, and neither are the final uncertainties on the fitted parameters. With these unknown sources of uncertainty, it is impossible to say much more about the reason for consistency or otherwise of the fits. A more robust study would consider the contribution of each of these sources in a controlled experiment.

**EFFECTS OF SINGLE PARAMETERS.** To better understand the parameters that are most important in the variance of the HMF, we marginalise over all but one of the parameters at a time, calculating the uncertainty, and plot the fractional uncertainty of each parameter with respect to the total uncertainty (summed in quadrature) in Figure 3.5. These results are for the Planck parameter set.

The dark matter density parameter  $\Omega_c h^2$  is the dominant source of uncertainty in both the amplitude and slope of the HMF except at high masses.  $\Omega_c h^2$  has the intuitive effect of increasing (decreasing) the amplitude of the HMF as it increases (decreases). At high masses, the effect of  $\sigma_8$  is even more important than  $\Omega_c h^2$  because it regulates the amplitude of density perturbations. Most fitting functions contain an explicit inverse exponential dependence on the mass variance, which is linear in  $\sigma_8$ , so for large masses (where the mass variance is very small), the HMF is exponentially sensitive to this parameter.

**APPLICATION TO NON-STANDARD DARK MATTER MODELS.** We expect the HMF to be sensitive to the nature of dark matter. Warm Dark Matter (WDM) models, in which the abundance of low-mass haloes is suppressed, provide a popular alternative to the CDM model and there has been much interest in them recently (e.g. Smith and Marković, 2011; Schneider, Smith, and Reed, 2013). The signature of WDM should be evident in the HMF on galaxy-mass scales and below, whereas most work on the HMF has focussed on group- to cluster-mass scales. This raises the question – what is the maximum mass scale for a given WDM model which can theoretically differentiate it from CDM given the current best cosmological parameters? Or conversely, how well constrained must the parameters be to differentiate reasonable WDM models at a certain mass? These questions are important for future efforts to detect a WDM signal in the HMF, such as, for example, the GAMA survey (Driver et al., 2011; Robotham et al., 2011) that aims to

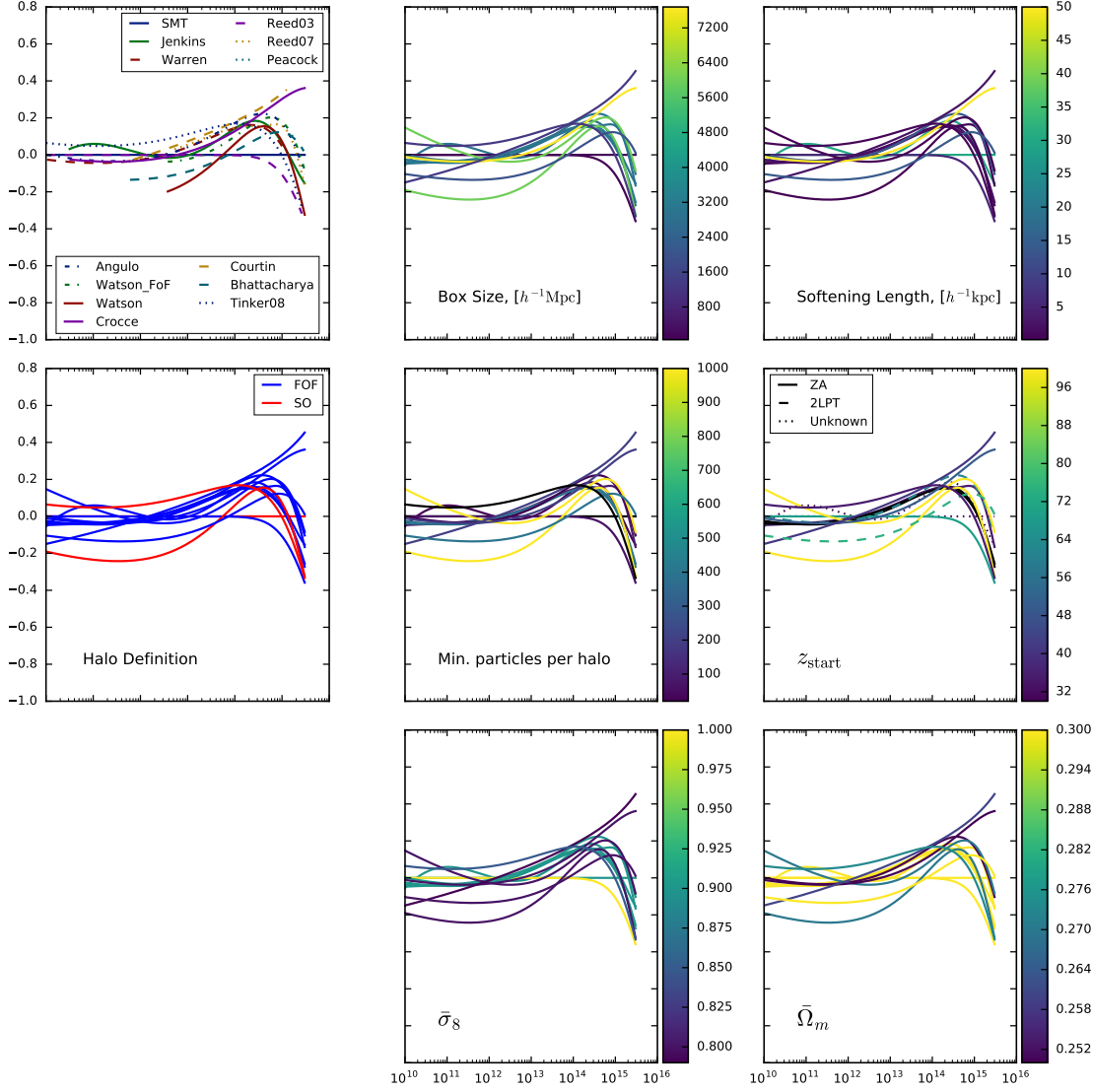


Figure 3.4: Variation of HMF fits colour-coded by various simulation parameters pertinent to each. Each fit is normalised to the Sheth, Mo, and Tormen (2001) form, and all panels show the same curves, coloured by different parameters. In many cases, fits were performed to a suite of simulations which makes description via a single number somewhat arbitrary. We choose to use the *maximum* box size, *minimum* softening length, *minimum*  $z_{\text{start}}$ , and *mean* for  $\sigma_8$  and  $\Omega_m$ . Most clear is the lack of definitive correlation between any single parameter and fit. However, halo definition and  $z_{\text{start}}$  do seem to have an appreciable effect.



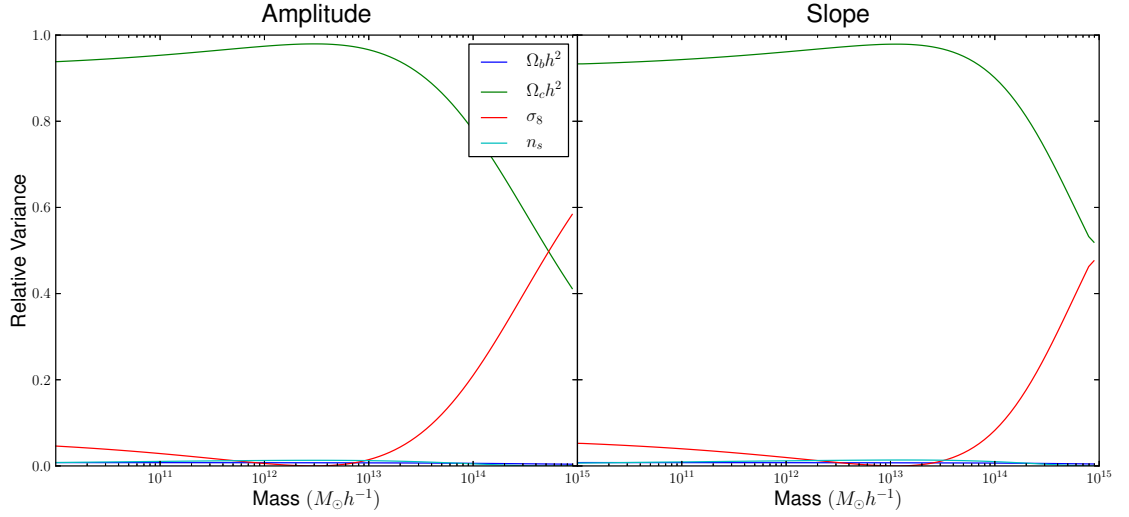


Figure 3.5: Contribution of single parameters to the overall variance. Each curve is the error of a single parameter as a fraction of the error summed in quadrature of all four parameters. Errors are taken from the Planck cosmology.  $\Omega_c h^2$  is the dominant term for most of the mass range, with  $\sigma_8$  becoming extremely important at high mass.

measure the HMF down to scales of  $10^{12} M_\odot h^{-1}$ . To answer them, one needs to know the intrinsic uncertainty in the CDM mass function, and deduce where the particular WDM model becomes significantly different.

Figure 3.6 shows the results of calculating the CDM HMF for the Planck parameters, with error region shaded, compared to equivalent WDM HMFs for candidate particle masses between 1 and 10 keV. Here we assume the transfer function of Bode, Ostriker, and Turok (2001) is applied to the CDM linear matter power spectrum to produce the WDM linear matter power spectrum, required to compute the mass variance  $\sigma$  in Eq. 3.1. For a plausible WDM particle mass of 1 keV – in the sense that it is consistent with observational limits on the WDM particle mass – we find that the maximum mass scale at which the CDM and WDM models are inconsistent is  $\sim 10^{11} M_\odot h^{-1}$ . Increasing the particle mass should mimic a “colder” dark matter model and reduce the discrepancy with the CDM HMF, which is what the results show.

By observing that the dominant source of uncertainty at intermediate mass scales is  $\Omega_M$ , we can calculate error-ranges on the HMF for a given error in this parameter (marginalising over other insignificant parameters). By doing so, we may calculate the necessary uncertainty required to theoretically distinguish a WDM model with given particle mass from the CDM model at a given mass scale. This allows us to make a rough estimate of the time it will take to reduce errors in  $\Omega_M$  sufficiently to differentiate reasonable WDM models. We can already, in principle, distinguish a 1keV WDM model from the CDM model assuming Planck parameters at  $10^{11} M_\odot h^{-1}$ , which is expected from figure 3.6, but it will be 20+ years before we can distinguish between WDM and CDM on group mass scales ( $\sim 10^{13} M_\odot h^{-1}$ ).

**REDSHIFT DEPENDENCE** We have repeated our analysis at  $z=1$ , the redshift at which a number of current and future surveys are operating. Results are qualitatively similar although we note that uncertainties for all base cosmologies are increased by a factor of

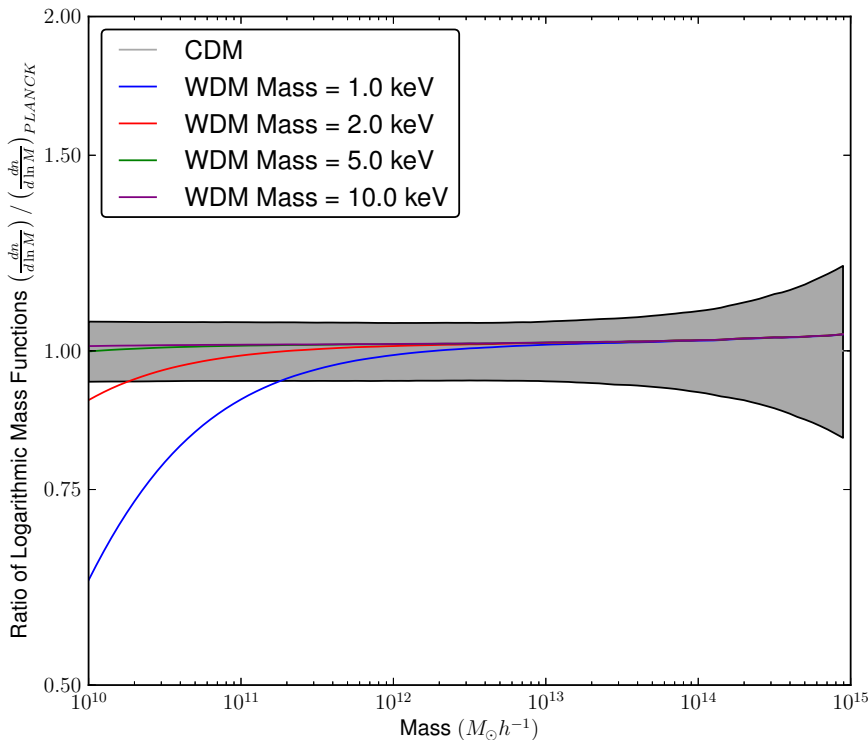


Figure 3.6: Theoretical HMFs for WDM models with specified particle masses in keV. Overplotted is the error region for the Planck cosmology. This gives an indication of the maximum mass which can be used to differentiate a particular WDM model.

$\sim 2$  across the mass range considered. The uncertainty stemming from choice of fitting function also increases, and we find that Planck results have smaller uncertainties than the fitting functions. This highlights the need for a consistent approach to fitting the HMF, and, as argued by Bhattacharya et al. (2011), raises the question of the suitability of simple analytic fits to the mass function for future percent-level analyses.

### 3.4 SUMMARY

We have examined how our knowledge of the HMF has improved over the last decade with the availability of increasingly accurate estimates of the cosmological parameters by the WMAP and Planck satellites. Our analysis reveals the limiting uncertainty in the HMF now comes from the variance between fitting functions, with errors intrinsic to the parameters being constrained to  $\sim 6\%$  at low masses, and  $\sim 20\%$  at high masses. We find that the primary sources of uncertainty in the HMF for a standard flat  $\Lambda$ CDM cosmology are the matter density parameter  $\Omega_M$  and the normalization  $\sigma_8$ . Of these,  $\sigma_8$  is the key parameter which will influence further reductions in the uncertainty in the HMF.

In addition, we have examined prospects for differentiating between the standard CDM model and its WDM variants using the HMF and found that reasonable models only become detectable at scales less than  $\sim 10^{11} M_\odot h^{-1}$  with current parameter uncertainties. We place an upper-limit of 0.8% on the uncertainty in  $\Omega_M$  before theoretical

detection of a WDM model of  $1\text{keV}$  is possible at  $10^{12}M_{\odot}h^{-1}$  – a  $\sim 300\%$  tighter constraint than the Planck results.

We conclude by noting that, while the strongest constraint on the HMF can be made in theory using data drawn from the entire halo mass range, this is prohibitively difficult to do in practice. Assuming a volume limited sample of haloes, the strongest constraint comes from the smallest haloes because of Poisson statistics. However, for an HMF sample constructed with a volume limit scaled optimally to match halo mass (i.e. detection of larger haloes require larger volumes), there will be a volume disconnect between bins, so the halo mass bin to bin variance will dominate HMF shape. This means that we cannot use the entire HMF to constrain, say, both dark matter and dark energy models, because the volume overlap between observable low mass halos and large clusters is virtually nil. Therefore future surveys that seek to use the HMF should target either the most massive haloes – and tests of dark energy – or the hosts of the lowest mass galaxies – and tests of dark matter; they should not do both.



---

## AN EMPIRICAL MASS FUNCTION DISTRIBUTION

---

### ABSTRACT

The halo mass function, encoding the comoving number density of dark matter halos of a given mass, plays a key role in understanding the formation and evolution of galaxies. As such, it is a key goal of current and future deep optical surveys to constrain the mass function down to mass scales which typically host  $L_*$  galaxies. Motivated by the proven accuracy of Press-Schechter-type mass functions, we introduce a related but purely empirical form consistent with standard formulae to better than 4% in the medium-mass regime, that is more relevant to the description of observed data in this regime. In particular, our form consists of 5 parameters, each of which has a simple interpretation, and can be directly related to parameters of the galaxy distribution, such as the turnover mass,  $M_*$ .

Using this form within a hierarchical Bayesian likelihood model, we show how individual mass-measurement errors can be successfully included in a typical analysis whilst accounting for Eddington bias. Furthermore, we utilise the relationship between our form and the double-Schechter stellar mass function to motivate more precise parameterisations of the stellar-mass halo-mass relation.

This distribution is a generalisation of the popular Schechter function, and we give several useful results that apply to both. Open-source Python and R codes to apply our new form are provided at <http://mrpy.readthedocs.org> and <https://cran.r-project.org/web/packages/tggd/index.html> respectively.

### 4.1 INTRODUCTION

Our understanding of galaxy evolution is largely hinged upon the dark scaffolding of the Universe – the dark matter distribution and its network of filaments, halos and voids. In particular, every galaxy is believed to be housed in a dark matter halo, each the product of a turbulent series of mergers over cosmic time. This history of mergers gives rise to an elegant picture of cosmic structure formation: the so-called hierarchical paradigm (White and Rees, 1978). This paradigm has been helpful in allowing predictions of many key characteristics of the halo distribution, including the halo mass function (HMF): the comoving number density of haloes as a function of their mass.

Clearly, the evolution of the halo distribution and that of the hosted galaxies is fundamentally interconnected. For this reason, measurement of the HMF and its connection

to observed galaxy properties are of fundamental importance. Whilst the abundance of cluster-mass halos is tightly related to the underlying cosmology, the galaxy population is predominantly hosted by haloes in the medium-mass regime  $\sim 10^{10} - 10^{13} h^{-1} M_{\odot}$ . Measurement of the HMF in this regime via dynamical mass estimates is thus a key goal for current surveys such as GAMA (Galaxy and Mass Assembly, Driver et al., 2011) and future deep surveys such as WAVES (Wide Area VISTA Extra-galactic Survey, Driver et al., 2015) and MSE (Manukia Spectroscopic Explorer, McConnachie et al., 2016).

The utility of measuring the HMF with dynamical halo masses inferred from groups of galaxies has been shown in Eke et al. (2006), who used the 2PIGG (2dFGRS Percolation-Inferred Galaxy Group) catalogue. Using a simple empirical correction to mitigate Edington bias, they were able to infer the group luminosity function and mean mass-to-light ratio of galaxies within their sample. Additionally, they were able to infer a best-fit value of  $\sigma_8$  by comparing to simulation results. However, similar studies with improved analysis methods and larger surveys have not been forthcoming. Instead, there has been an increased focus on less direct methods of connecting the dark sector to galaxies, such as the halo occupation distribution (HOD; Berlind et al., 2003; Zehavi et al., 2005; Zehavi et al., 2011). One reason for this focus may be an increased reliance on the standard formalism for theoretical HMFs.

The standard model for the HMF arises from physical arguments proposed by Press and Schechter (1974, cf. Chapter 1 & 2), and we shall refer to this framework (with its many specific fits) collectively as the “EPS formalism” (not to be confused with the specific original form of Press and Schechter). The EPS formalism has proven remarkably successful in describing the HMF measured in cosmological simulations, albeit with modifications of  $f(\sigma)$  to better capture the non-spherical nature of halo collapse (eg. Sheth, Mo, and Tormen, 2001). However, though it appears to be universal with respect to cosmology and redshift, it has been shown to exhibit non-universality at the  $\sim 20\%$  level, especially for halos defined by a spherical-overdensity criterion (Tinker et al., 2008; Watson et al., 2013). Furthermore, the best-fit seems to depend not only on underlying physics and halo definition, but on the algorithm used for locating halos within cosmological simulations (Knebe et al., 2011). Though these uncertainties do not prevent the use of the PS formalism in accurate fits to cluster abundance (Haan et al., 2016), they motivate the use of a simpler empirical distribution in the context of observed HMFs at lower masses.

In addition to this motivation, the most common descriptions of ensemble galaxy properties are entirely empirical; both luminosity functions (LFs) and galaxy stellar mass functions (GSMFs) are parameterised by single- or double-Schechter functions (Schechter, 1976), which capture their forms adequately without any first-principles motivation. Thus for example, in the single-Schechter case, the galaxy stellar-mass function (GSMF) is commonly expressed as

$$\frac{dn}{dM} = \phi_{\star} \left( \frac{M}{M_{\star}} \right)^{-\alpha} \exp \left( -\frac{M}{M_{\star}} \right). \quad (4.1)$$

Here the model parameters themselves hold explanatory power:  $M_*$  is a characteristic turn-over mass, and  $\alpha$  is a power-law slope for low-mass galaxies, while  $\phi_*$  is the normalisation of the abundance. Measuring the HMF via the PS formalism obfuscates the relation between parameters and data, rendering the comparison of the dark HMF with the visible galaxy properties less clear. Given the apparently similar forms of the HMF, LF and GSME, it would seem beneficial to construct their parameterisations in a similar fashion so as to facilitate comparison; for example, can we relate the evolution of a typical halo mass  $\mathcal{H}_s$  with that of the typical stellar mass,  $M_*$ , or luminosity  $L_*$ ?

Using a simple empirical form defined directly in terms of mass also bears other advantages. It is precise, in the sense that output depends on input parameters in a simple, reproducible way, without the numerical uncertainties involved in infinite-limit integrals (cf. Eq. 2.3) and all the machinery required for the PS formalism. It is portable, both in the sense that anyone can easily create their own numerical function for its evaluation, and that data described by the form are reduced to a small set of interpretable parameters. It is fast, since it is a direct calculation. And since it is analytic, it affords more extended analyses and methods.

Conversely, the empirical nature of the form means that it is not easily linked with an underlying cosmology. For this reason, it will not be useful for cluster abundance studies which primarily seek to constrain the power spectrum normalisation  $\sigma_8$ , and dark energy parameters. Rather, it is most usefully applied to questions about the connection between dark halos and galaxies.

In this paper, we propose an empirical form that is equivalent to a re-normalised truncated generalised gamma distribution (TGGD). The TGGD is a generalisation of the Schechter function, in which the sharpness of the exponential cut-off is variable. This similarity enables direct comparison between the HMF and galaxy ensemble properties (cf. §4.5). At the same time, we develop an hierarchical Bayesian likelihood model, which takes full advantage of the analytic nature of our form to incorporate arbitrary per-object mass-measurement uncertainties into the analysis, correctly accounting for pervasive Edington bias. We envision that this combination of empirical form and Bayesian model will enable efficient and accurate measurements of the low-mass HMF in future deep surveys.

The paper is set out as follows: §4.2 introduces the model, which we call the MRP throughout, its main features and properties, and shows some initial comparison to standard EPS modelling. §4.3 describes the general method of using MRP to fit halo data with and without per-object uncertainties, applying this method to simulation and mock data. §4.4 gives some relations between the MRP parameters and the main physical parameters that the HMF is sensitive to, and §4.5 uses the analytic nature of MRP to deduce properties of the Stellar-Halo Mass Relation (SMHM), before concluding in §4.6.

To accompany this paper, we have developed an extensive Python code, called `MRPY`<sup>1</sup>, which implements all of the functionality explored throughout the paper. Furthermore, we provide a reduced R version of this code which focuses on the basic statistical applications, called `TGGD`<sup>2</sup>.

<sup>1</sup> Found at <http://mrpy.readthedocs.org>. Note that wherever possible, we provide the exact code to reproduce the figures in this paper as examples at <https://github.com/steven-murray/mrpy/examples>.

<sup>2</sup> Found at <https://github.com/asgr/tggd> and downloadable from CRAN as `tggd`.

## 4.2 THE MRP MODEL

The MRP form is a truncated generalised gamma distribution (GGD) with non-unity normalisation. A truncated distribution is used because the HMF is non-convergent in number counts towards low masses (while converging in mass density), so it is mathematically necessary to impose a lower limit to form a distribution. Halo mass data is usually in the form of all (selected) halos above some threshold, so this truncation is also useful in practice.

The MRP can be parameterised as

$$\text{MRP}(m|\vec{\theta}) = A\beta \left(\frac{m}{\mathcal{H}_s}\right)^\alpha e^{-\left(\frac{m}{\mathcal{H}_s}\right)^\beta} \equiv Ag(m|\vec{\theta}) \quad (4.2)$$

where  $A$  is an arbitrary normalisation, and  $\vec{\theta}$  is the parameter vector  $(\mathcal{H}_s, \alpha, \beta)$ . This is a generalisation of the Schechter function, which has a fixed  $\beta = 1$ . Note that many of the results derived in this paper will be applicable (in modified form) to the Schechter function also.

Usually, the GGD is defined for  $m > 0$ ,  $\mathcal{H}_s > 0$ ,  $\beta > 0$  and  $\alpha > -1$ , where the latter two constraints ensure convergence at the high- and low-mass ends respectively. In the case of the truncated GGD which we adopt, convergence is assured at low mass and we can dispense with the constraint on  $\alpha$ . Indeed, typical HMFs exhibit  $\alpha \sim -2$ , and we note that  $\alpha = -2$  is the limit for convergence in total mass density.

### 4.2.1 Cumulative Distribution

An important quantity is the integral of  $g$  over mass, which determines the normalisation of the distribution. We have

$$\int_{m_{\min}}^{\infty} g(m|\vec{\theta}) dm = \mathcal{H}_s \Gamma(z, x) \equiv q(\vec{\theta}, m_{\min}) \quad (4.3)$$

where

$$z = (\alpha + 1)/\beta \quad (4.4)$$

$$x = \left(\frac{m_{\min}}{\mathcal{H}_s}\right)^\beta, \quad (4.5)$$

and  $\Gamma$  is the Euler (upper-incomplete) Gamma function. We will use  $z$  and  $x$  in this manner throughout this paper, noting that  $x_m$  will refer to  $x$  with  $m_{\min}$  replaced by arbitrary  $m$ .

We note that most numerical libraries (with notable exceptions) only implement  $\Gamma(z, x)$  for  $z \geq 0$ , so that  $\alpha$  is required to be  $> -1$  (consistent with the usual definition of the GGD). Fortunately,  $\Gamma$  admits a stable recurrence relation which can be used to efficiently calculate the function with  $z < 0$ , in terms of only positive values, so long as  $|z|$  is not too large (in our case, it should not exceed 3 very often). We present details of the mathematical derivation and numerical implementation of this relation for our work in [Appendix A.1](#).



PROPERTY	VALUE
$n(> m \vec{\theta})$	$Aq(\vec{\theta}, m)$
$n(> m_l, < m_u \vec{\theta})$	$A(q(\vec{\theta}, m_l) - q(\vec{\theta}, m_u))$
$\rho(> m \vec{\theta})$	$A\mathcal{H}_s^2\Gamma(z + 1/\beta, x)$
Mode	$m_{\min}$
Mean	$\mathcal{H}_s \frac{\Gamma(z+1/\beta, x)}{\Gamma(z, x)}$
Variance	$\mathcal{H}_s^2 \left( \frac{\Gamma(z+2/\beta, x)}{\Gamma(z, x)} - \left( \frac{\Gamma(z+1/\beta, x)}{\Gamma(z, x)} \right)^2 \right)$
Log. Mass Mode, $\mathcal{H}_T$	$\mathcal{H}_s \sqrt[3]{z + 1/\beta}$

Table 4.1: Properties and derived quantities of MRP. The mass mode is the logarithmic mass bin with the highest mass density.

#### 4.2.2 Normalisation

While in principle,  $A$  can be set arbitrarily to fit given data, there are two values of interest. Firstly, to ensure the MRP is a valid statistical distribution defined on the support  $(m_{\min}, \infty)$ , one can set

$$A = A_1 = 1/q(\vec{\theta}, m_{\min}). \quad (4.6)$$

To normalise the MRP so that its magnitude is equivalent to the expected number density of objects,  $\frac{dn}{dm}$ , we may make the (standard halo model) assumption that all CDM density is contained within halos at some scale (Cooray and Sheth, 2002). In this case, we force the integral of the mass-weighted function over all masses to converge to the known matter density,  $\Omega_m \rho_{\text{crit}}$ . We find that

$$A = A_{\rho_c} = \Omega_m \rho_{\text{crit}} / k(\vec{\theta}) \quad (4.7)$$

where

$$k(\vec{\theta}) = \mathcal{H}_s^2 \Gamma(z + 1/\beta), \quad (4.8)$$

which is positive for  $\alpha \geq -2$ , as per previous discussion.

Table 4.1 lists some further properties and derived quantities of MRP.

#### 4.2.3 Re-parameterisations

Re-parameterisations of the MRP are available which can reduce the covariance between the parameters (Lagos Álvarez, Ferreira, and Valenzuela Hube, 2011; Lawless, 1982). We explore these additional parameterisations in more detail in Appendix A.2. We only comment here that while these may be beneficial in attaining parameter fits, we choose to use the parameterisation in Eq. 4.2 because the parameters  $\vec{\theta}_i$  have well defined char-

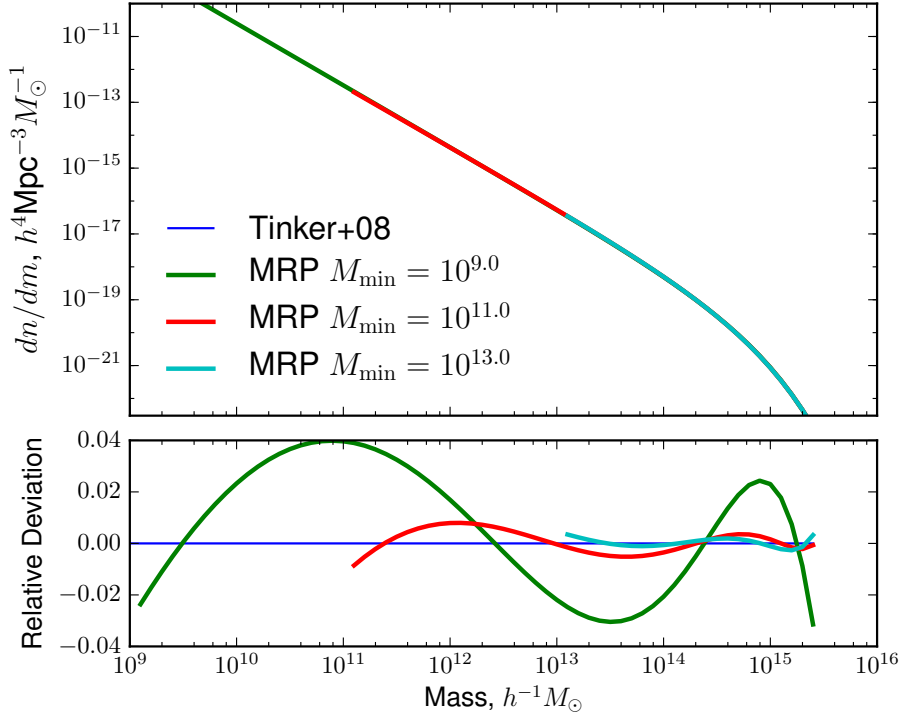


Figure 4.1: Comparison of MRP form to that of To8. MRP is within 4% over a large range of masses when fitted over the range  $10^9 < M < 10^{15.5}$ , and this deviation is dramatically reduced by fitting over smaller ranges.

acteristics in this form:  $\mathcal{H}_s$  is the turnover mass,  $\alpha$  is the low-mass power-law slope, and  $\beta$  controls the steepness of the exponential cut-off at high mass.

We note here that it is possible to formulate the MRP in terms of the logarithmic mass mode  $\mathcal{H}_T$  directly, and that this parameter is closely related to  $\mathcal{H}_s$  (the exact relation is given in Table 4.1). Conceptually,  $\mathcal{H}_T$  is the ‘typical’ mass of a halo for a given distribution, and we expect it to be less dependent on the other parameters than  $\mathcal{H}_s$ . It may be beneficial in future studies to adopt this parameterisation, but we do not here for the sake of simplicity.

We also note that the Schechter function affords the same relation, though since  $\beta = 1$  and  $\alpha$  is generally close to  $-1$ ,  $\mathcal{H}_s$  is very close to  $\mathcal{H}_T$ .

#### 4.2.4 Comparison to EPS fits

Throughout this paper we adopt the fiducial cosmology of Planck Collaboration (2014a) (hereafter P13), so that all plots and data – including illustrative ones – are calculated using it (unless explicitly stated). Likewise, we adopt the fit of Tinker et al. (2008, hereafter To8) as our fiducial mass function.

Figure 4.1 shows a comparison of MRP to To8. The fit was performed using least-squares minimisation in log-log space, with the three parameters  $\vec{\theta}_i$  free, and  $A$  post-fit by equating integrals.

The MRP provides a relatively good fit down to very low mass scales, with deviations not exceeding 4% over the range  $10^{11} < M < 10^{15.5}$  if fit over that range. The high-mass

cut-off is not perfectly modelled by MRP, but it is difficult to assess whether the To8 HMF itself is well fit at these scales, due to the inherent Poisson noise in simulations.

The most marked feature of the deviation is an oscillatory pattern in the power-law region of the distribution. This oscillatory pattern is expanded as the truncation mass is decreased, indicating that it is a model deficiency, and that the best-fit values will be dependent on  $m_{\min}$ . We explore the ramifications and extent of this dependency in Appendix A.3. It is of interest whether this feature is a loss of information in MRP as compared to the true physical distribution, or whether the oscillations are falsely present in the EPS formalism. This question is best answered by observing residuals of fits to simulations, which can be found in §4.3.

We note that the MRP form can be precisely recovered via the EPS formalism if we assume the analytic fit of Press and Schechter (1974) and that  $\sigma(m) \approx (m/m_1)^{-\gamma}$ . This assumption clearly does not hold in detail, but is reasonable to zeroth order. Doing so yields the following equalities:

$$\mathcal{H}_s = m_1 \sqrt{a} \tag{4.9}$$

$$\alpha = \gamma - 2 \tag{4.10}$$

$$\beta = 2\gamma \tag{4.11}$$

$$A = \frac{a}{m_1 \sqrt{\pi}}, \tag{4.12}$$

where  $a = (2/\delta_c^2)^{1/\gamma}$ . In practise, the MRP is not so restrictive –  $\alpha$  and  $\beta$  are free – so it is able to achieve more precise fits than this approximation would suggest.

### 4.3 FITTING MRP

One of the motivating features of the MRP form is that it is numerically efficient, and has three well-defined parameters, making it well suited for fitting observed or simulated data. In addition, since the functional form is analytic, it can be used in efficient Markov-Chain Monte Carlo (MCMC) step methods such as Hamiltonian Monte Carlo (HMC), which require the gradient, usually via automatic differentiation. The utility of such methods for problems with (very) large dimensions makes such a form invaluable for robust model fitting.

In this section we outline the general method for fitting data with MRP. We first present the basic likelihood function, before enriching the framework by addressing some common issues with the posterior. We then cover three applications: i) fitting a suite of DM-only simulations simultaneously, ii) fitting a large suite of hydrodynamic simulations independently, and iii) fitting a set of mock haloes with per-object “measurement” uncertainties.

Along with the requisite method development introduced by each of these applications, we are particularly interested in the behaviour of the residuals in the first two cases. These give insight into the performance of MRP with respect to both the EPS formalism and also intrinsic uncertainties in specifying the HMF.

### 4.3.1 Basic framework

The simplest method is to fit to pre-binned data, or equivalently a theoretical curve. Such fits can be adequately performed via least-squares methods (we generally recommend performing this in log-log space), and we present the relevant formulae for this in Appendix A.4.1. Our focus here however will remain on the more useful fits which utilise all halo information, as well as individual uncertainties.

Such a procedure has at least two advantages: (i) it eliminates the need for choosing artificial bins, which can introduce biasing, and (ii) it allows robust book-keeping of measurement errors in halo masses, where each halo may have independent measurement uncertainty.

The key quantity in Bayesian estimation is the log-likelihood, which may be minimized using down-hill gradient optimization methods, or traced with Markov-Chain Monte Carlo (MCMC). In the simplest case, with a uniform prior and no uncertainties on halo masses, the log-likelihood of given data is

$$\ln \mathcal{L} = \sum_{i=1}^{N_{\text{data}}} \ln f(m_i | \vec{\theta}, m_{\text{min}}), \quad (4.13)$$

where  $f = g/q \equiv A_1 g$ .

We note that this likelihood has no dependence on the normalisation,  $A$ , which must be post-fit. This can be achieved by equating the mass-weighted integral with the total density, so

$$A = \frac{\rho_{\text{tot}}}{\mathcal{H}_s^2 \Gamma(z + 1/\beta, x)}. \quad (4.14)$$

When fitting with downhill-gradient optimization schemes, it is often useful to have an analytic Jacobian and Hessian available. We provide these (assuming uniform priors on parameters) in Appendix A.4.2.

### 4.3.2 Common Issues

#### 4.3.2.1 Numerical Issues

When  $x$  becomes very large,  $q$  asymptotes to zero and errors can arise when taking logarithms. Even when  $x$  is not so large as to give hard errors, the saturation of the value of  $q$  renders the posterior difficult to explore.

To mitigate this, it is advisable to begin iterations of the exploration at reasonable values of  $\vec{\theta}$ . Furthermore, it is useful to set appropriate boundaries on the value of  $\mathcal{H}_s$ , which stop it from exploring values too low. These values may otherwise be explored due to either poor starting values, high covariances between parameters, or the algorithm itself (some downhill methods try both boundaries to begin with).

Finally, if an arbitrary precision library is available, it is useful to perform the calculation of  $\ln \Gamma(z, x)$  with it, so as to retain precision for extremely small values.

#### 4.3.2.2 High Covariance

The level of covariance between parameters is dependent on the data, however in general it is quite high (see eg. figure 4.4). For certain pathological cases (which do not seem to be all that rare), the covariance can be so high that it forms a ridge in the posterior which introduces non-identifiability in the parameters.

This can result in a posterior which “piles up” against the artificial parameter boundary, or entirely unphysical solutions. A less severe side-effect of this behaviour is a tendency to swap the roles of  $\alpha$  and  $\beta$ , while setting  $\mathcal{H}_s < m_{\min}$ . This is clearly not the intended behaviour of the MRP parameters –  $\alpha$  should reflect the power-law slope down to low masses. Data with a relatively high  $m_{\min}$  is more likely to prefer this unphysical situation.

A simple way to mitigate this is to use weakly informative priors. For instance, broad Gaussian priors on all three parameters, centred on some fiducial value, can help to constrain the parameters to reasonable values. We use these as necessary throughout this study.

A second way to alleviate the high covariance is to use alternate parameterisations, as previously discussed. We do not employ this technique here, but remark that it may be beneficial.

A third method lies in upweighting high-mass haloes which otherwise contribute very little to the likelihood. As this predominantly affects the next issue, we shall discuss it in that context.

#### 4.3.2.3 Low influence of high-mass halos

Since the slope is generally quite steep ( $\alpha \sim -1.85$ ), the likelihood is dominated by low-mass halos, and very little information comes from the high-mass cut-off, rendering  $\beta$  difficult to constrain, especially in small samples.

As mentioned, we can balance the contribution of low- and high-mass halos in the final likelihood by up-weighting high-mass halos. A convenient choice of weight function is

$$w(m) = cm^s, \quad (4.15)$$

where  $s$  is some positive scaling factor. A reasonable choice for the normalising coefficient  $c$  is that which leaves the total number density unchanged, ie.

$$c = N / \sum m_i^s. \quad (4.16)$$

This ensures that modifying  $s$  doesn’t artificially increase the perceived “signal” of the data.

In this case, the likelihood becomes

$$\ln \mathcal{L} = \frac{N_{\text{data}}}{\sum m_i^s} \sum_{i=1}^{N_{\text{data}}} \ln f(m_i | \vec{\theta}_s, m_{\min}) m_i^s, \quad (4.17)$$

where  $\vec{\theta}_s = (m_s, \alpha + s, \beta)$ . That is, we are implicitly fitting data that has a slope of  $\alpha + s$ .

Name	$N$	$L$ , [ $h^{-1}\text{Mpc}$ ]	$m_p$ , [ $h^{-1}M_\odot$ ]	# Halos	$M_{\min}$ , [ $h^{-1}M_\odot$ ]	$M_{\max}$ , [ $h^{-1}M_\odot$ ]
$\nu^2\text{GC-M}$	$4096^3$	560	$2.2 \times 10^8$	$5.3 \times 10^7$	$2.2 \times 10^{10}$	$7 \times 10^{14}$
$\nu^2\text{GC-H1}$	$2048^3$	140	$2.75 \times 10^7$	$5.5 \times 10^6$	$2.75 \times 10^9$	$2 \times 10^{13}$

Table 4.2: Details of halo abundances in the New Numerical Galaxy Catalog simulations used in this study.  $M_{\min}$ , as given by the public catalogues, is set by a minimum of 40 particles per halo. For our fits, we prefer to use a minimum of 100 particles per halo.

To achieve a fit which treats all logarithmic scales with the same weight, a scale of  $s \approx 1$  should be used, and we show that this is in fact the best choice of  $s$  for a range of circumstances in Appendix A.5.

An additional benefit of using this weighting arises in numerical evaluation. Since for all physical data,  $\alpha > -2$ , if we ensure  $s \geq 1$ , the likelihood only ever contains positive shape-parameters  $z$  for the Gamma function. While arbitrary  $z$  is calculable via the recurrence relations (cf. Appendix A.1), it is more often more efficient to use only positive values.

The weighting must be used with care – high mass halos contain higher Poisson scatter, which is liable to lead to greater uncertainty on the derived parameters when upweighted. For smaller test datasets ( $N \sim 25000$  halos), we found that while the weighting permitted a meaningful result, it was outperformed in terms of recovered accuracy by the non-weighted scenario. However, when fitting to simulation data throughout this paper, we found that using  $s = 1$  was often necessary to obtain meaningful results.

#### 4.3.3 Fitting to simulations

The simplest realistic halo catalogues arise from DM-only simulations. Such catalogues contain high-resolution, fully nonlinear determinations of the HMF, completely independent of the EPS framework assumptions. In this section we describe the necessary machinery to fit the MRP to a suite of simulations, focussing on the comparison of the resultant residuals to those from the EPS formalism and issues of halo definition.

In this study, we use the publicly available halos from the New Numerical Galaxy Catalog ( $\nu^2\text{GC}$ , Ishiyama et al. 2015; hereafter I15). These simulations were run with the P13 cosmology, in several box-sizes and mass resolutions. Specifically, we use  $\nu^2\text{GC-M}$  and  $\nu^2\text{GC-H1}$ , which when taken together probe a volume of  $560h^{-3}\text{Mpc}^3$  down to halos of mass  $2.75 \times 10^9 h^{-1}M_\odot$ . As such, it spans the halo mass range relevant for group catalogue studies. Details of the simulations are found in Table 4.2. In particular, we note that though the original catalogues contain haloes defined down to 40 particles, we fit only to those with 100 particles or more, to alleviate the effect of Poisson noise on the masses of these highly influential haloes. Furthermore, the box size of  $\nu^2\text{GC-M}$  is small enough to make the highest-mass objects quite rare, and thus dominated by cosmic variance and Poisson noise. We set upper limits of  $7 \times 10^{14} h^{-1}M_\odot$  and  $2 \times 10^{13} h^{-1}M_\odot$  on the mass ranges respectively.

We perform the fit simultaneously by summing the contributions to the log-likelihood from each individual simulation, keeping the appropriate truncation mass  $m_{\min}$  for each

and naturally weighting its contribution by the number of halos it contains. In effect, this highly weights the  $\nu^2\text{GC-M}$  box over most scales, except those which it does not probe.

In this case, because of the high resolution of the data, we use a scale-weighting of  $s = 1$ , and use uniform priors on all parameters. The fit was performed with a routine using the popular `EMCEE` python package for Monte Carlo sampling. The choice to use `EMCEE` rather than `STAN` here was based primarily on convenience and language limitations – synthesising several simulations is most naturally achieved through array slicing or ragged arrays, which are not yet implemented in the `STAN` language (as of v2.7.0).

Figure 4.2 shows the residuals of this fit as the blue markers (differing marker styles represent different boxes). A variation between boxes can be clearly seen, especially between the H2 box and the rest. This is almost certainly due to cosmic variance issues. Nevertheless, over the more robust mid-range, the MRP fit is accurate to within a few percent. This can be compared to the variation between simulated mass functions with different halo finders. The SO mass function is shown with green markers, and deviates from the FOF HMF by a far greater margin over the same range. This indicates that an arbitrary choice of halo definition is a greater loss of information than the approximation introduced by the MRP.

Also plotted in figure 4.2 are the residuals of the best-fit EPS mass function from I15, which was calibrated with the FOF halos. At low masses, the I15 residuals strongly favour the H2 simulation, which we have inherently down-weighted, and in fact the MRP better follows the larger simulations. This results in a 10% lower  $\chi^2$  value for the MRP, calculated on the weighted means of the bins shown in figure 4.2. In summary, the simple and efficient form of the MRP, which contains one less free parameter, is able to model the mass function data as well as the EPS formalism over a broad range of masses.

We also plot the posterior joint-likelihoods of the parameters in 4.3. We stress that this is a high-resolution run, which will be poorly represented by realistic observed datasets. With that in mind, the joint posterior is a very good approximation to a multivariate normal distribution, which has the benefit of allowing for quick analyses by downhill-gradient methods, combined with analytic determination of the distribution around the solution via the Hessian.

The analysis sets very tight constraints on the MRP parameters – better than a hundredth of a percent in all cases except for  $\beta$ , which is a few hundredths of a percent. The mass scale  $\mathcal{H}_s$  is around  $10^{14.5}h^{-1}M_\odot$ , and the mass mode is  $10^{13.4}h^{-1}M_\odot$ . We find a power-law slope of  $-1.9$ , consistent with the halo model assumption. We also find a cut-off parameter  $\beta = 1.19$ , which is marginally sharper than a Schechter function. We do not expect the constraint on  $\beta$  to be particularly robust because of the small box size of the simulation.

#### 4.3.4 MRP in the presence of baryons

There has been a recent surge in studies of the effect of baryons on the HMF (Cui et al., 2012; Cui, Borgani, and Murante, 2014; Martizzi et al., 2014; Bocquet et al., 2015).



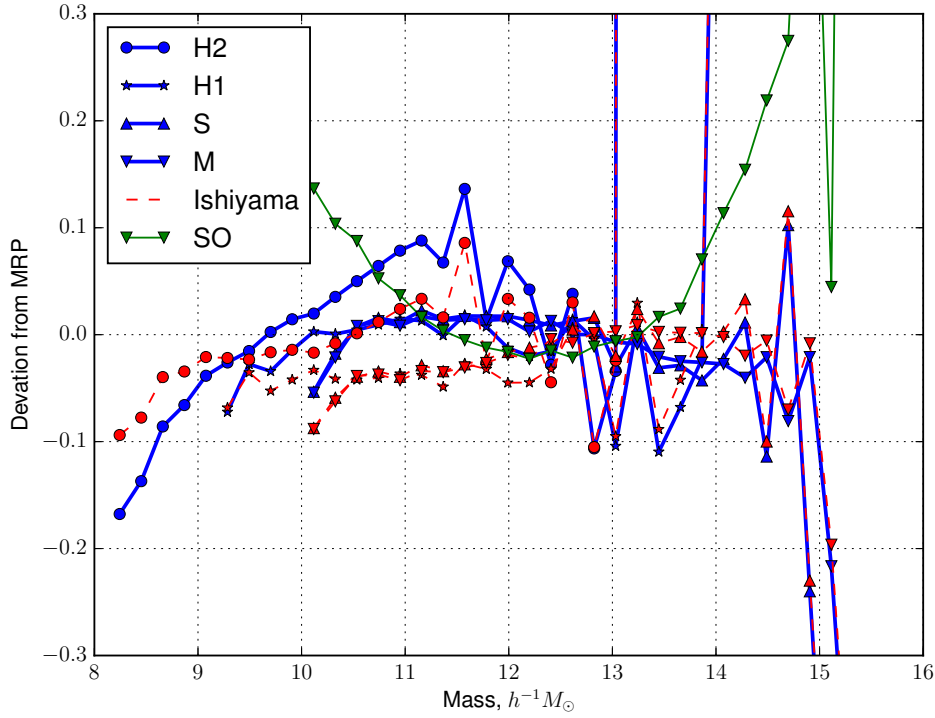


Figure 4.2: Residuals of MRP global fit to  $\nu^2$ GC simulations. The blue markers show residuals from the suite of FOF halos (which are used in I15 to calibrate their fit). The red markers show the I15 fit residuals, while the green markers show the mean residual of the SO halo catalogues (against the FOF-calibrated MRP fit), which illustrates that the FOF residuals are well within the discrepancy between halo finders.

Such studies are vital because they uncover expected deviations in the observed HMF as compared to pure DM-only theory. It is becoming clear that incorrect understanding of the baryonic physics will play a larger role in uncertainty in the HMF than the intrinsic residual of the theoretical model.

To exemplify the flexibility and efficiency of MRP in its application, and also the fact that its information loss is dominated by uncertainties from baryonic processes, we here fit the MRP to mass functions including baryonic effects.

We have chosen to use cosmo-OWLS (McCarthy et al., 2013; Le Brun et al., 2014) for this purpose. These simulations were carried out in boxes of size  $L = 400h^{-1}\text{Mpc}^3$ , with  $N = 1024^3$  particles, with cosmologies corresponding to WMAP7 and Planck, and a range of physics models. Full details can be found in Le Brun et al. (2014). In summary, we use six models: a dark-matter only base run (DMONLY), a run with just UV/X-ray background and no cooling, star formation or supernovae or AGN feedback (NOCOOL), a reference run with everything except AGN feedback (REF), and 3 runs that also include AGN feedback with different heating temperatures. In this analysis we use halos found with an SO algorithm, with the overdensity criterion set to  $200\rho_{\text{mean}}$ .



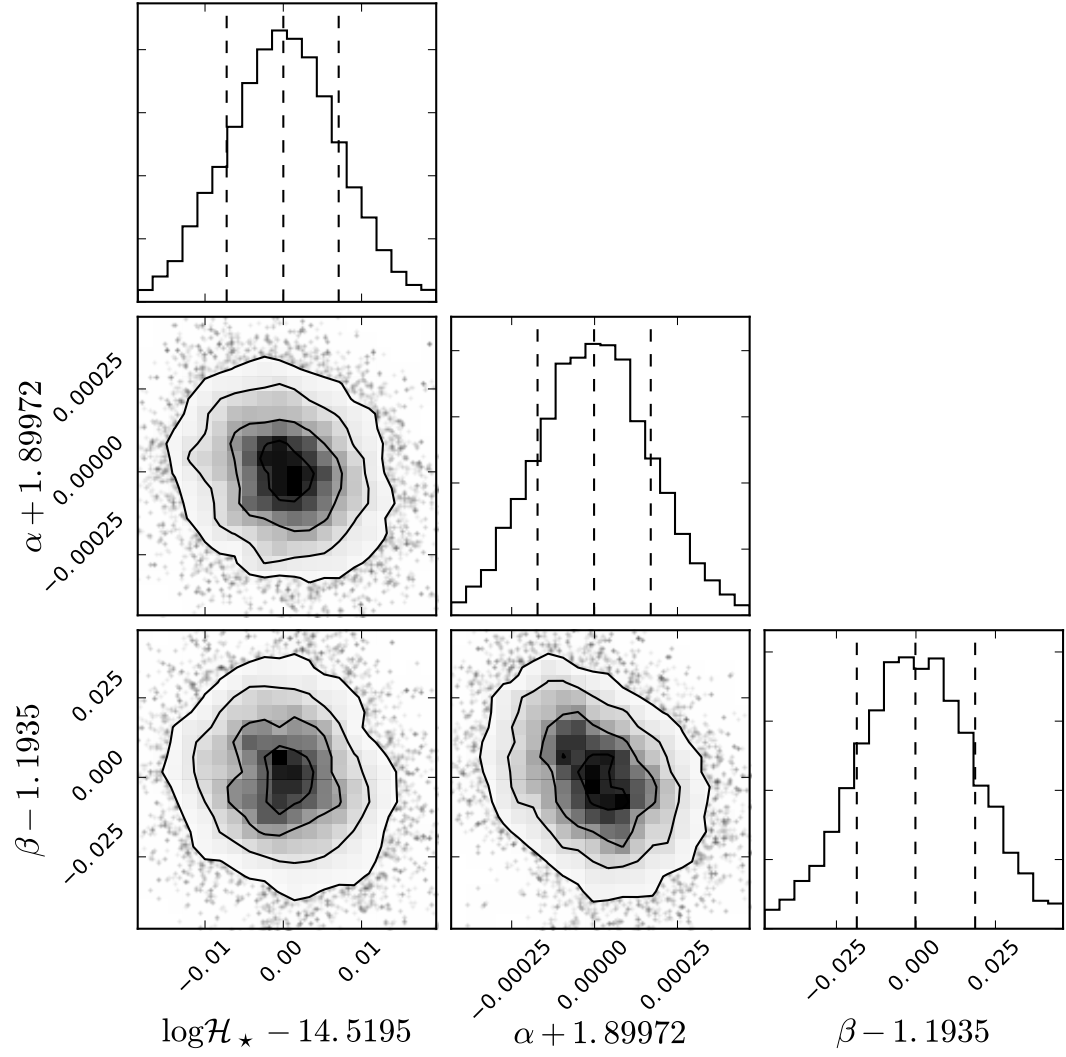


Figure 4.3: Posterior joint-likelihoods of the MRP parameters estimated simultaneously on all four  $\nu^2$ GC simulations. All contours display high degrees of normality.

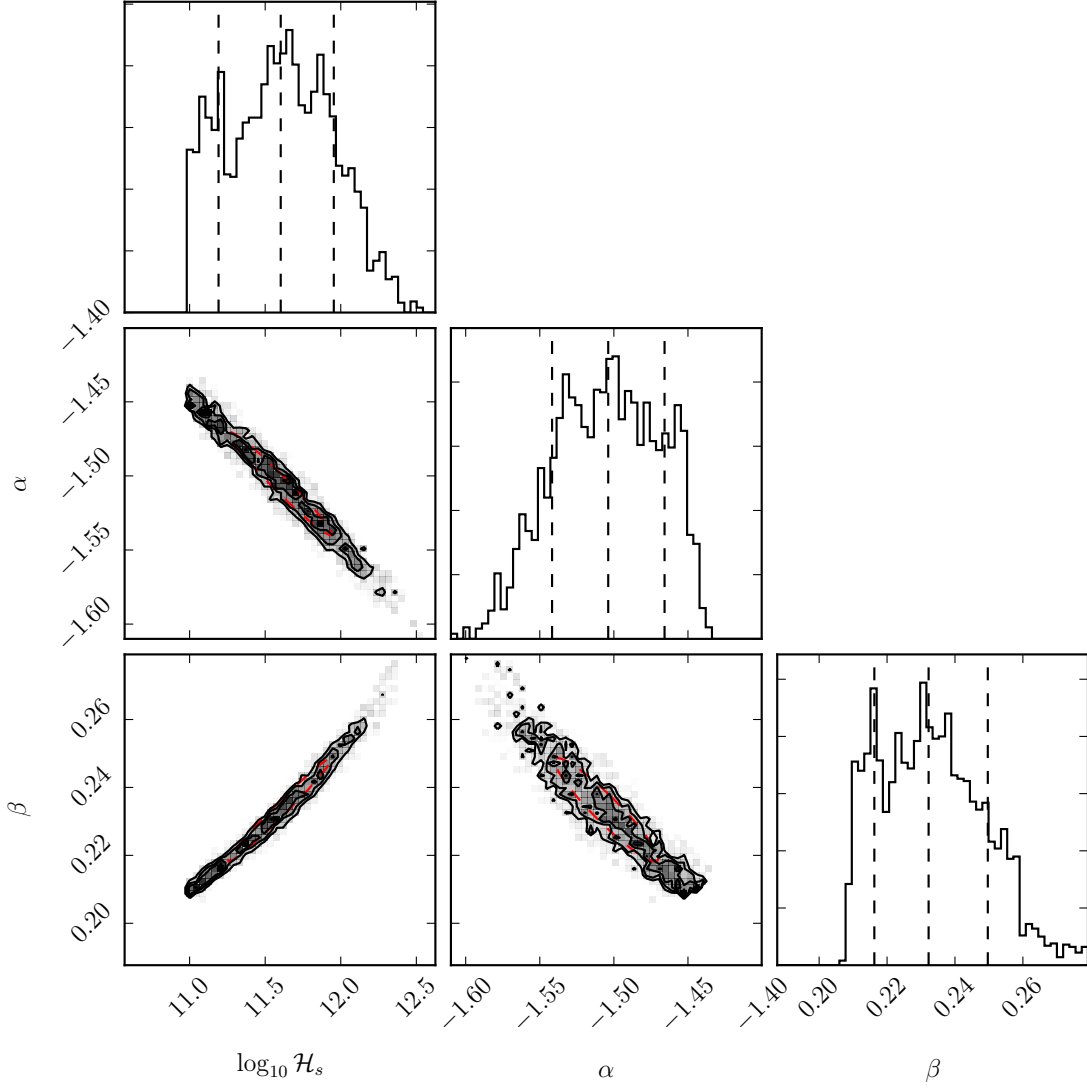


Figure 4.4: Joint posterior distributions for the AGN model with  $T_{\text{heat}} = 10^{8.7}$  at  $z = 0$  in WMAP7, using  $s = 0$ . The extreme level of covariance causes problems for the sampler.

We fit each of the 60 cosmo-OWLS snapshots using the same EMCEE-based routines as were used for  $\nu^2\text{GC}$ . In this case, for extra reliability (due to the smaller numbers of halos in cosmo-OWLS), we use both  $s = 1$  and weakly informative priors:

$$\log_{10} \mathcal{H}_s \sim \mathcal{N}(13, 1) \quad (4.18)$$

$$\alpha \sim \mathcal{N}(-1.8, 0.1) \quad (4.19)$$

$$\beta \sim \mathcal{N}(0.6, 0.2) \quad (4.20)$$

The bounds on each parameter corresponds to  $2\text{-}\sigma$  in terms of its prior, rendering the prior quite weak. We found that using  $s = 0$  for some of the cosmo-OWLS snapshots produced unphysical results, or even numerical exceptions, and the reason for this is exhibited in figure 4.4: the covariance is extreme. When using  $s = 1$ , the covariance is reduced to high, but manageable, levels.

It is possible that the improvement in performance for  $s = 1$  is also affected by the non-standard behaviour of the HMF for baryonic models at low- to mid-range masses. This

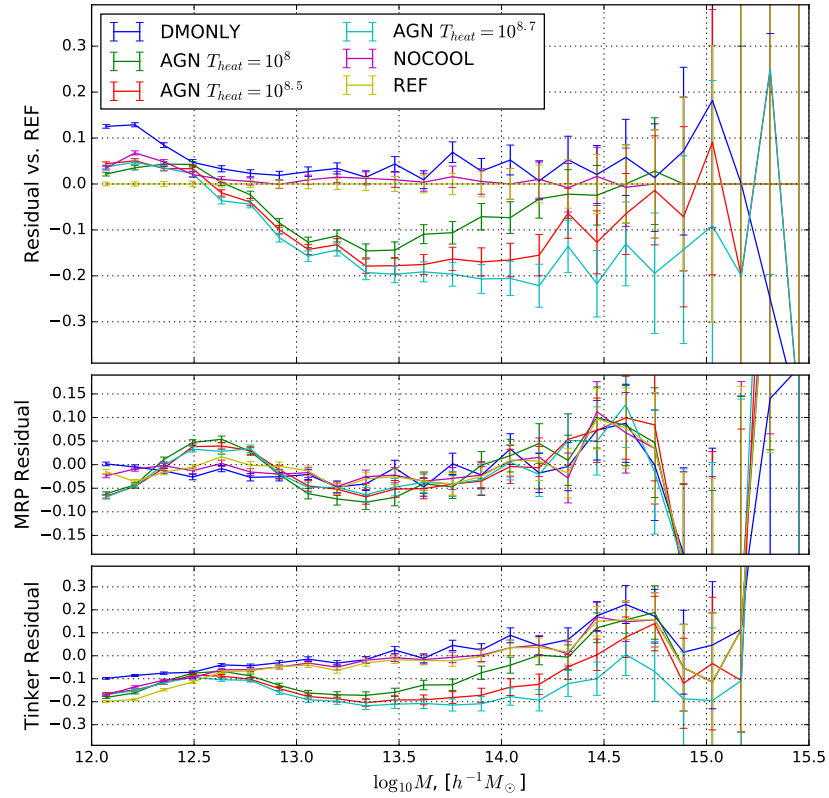


Figure 4.5: *Upper panel*: relative variation between the baryon models’ HMFs in cosmo-OWLS, normalised to the REF run in Planck cosmology. All baryon models tend to suppress the HMF, with AGN heating exacerbating this effect to the level of 10-20% around the mass mode. Error-bars indicate Poisson uncertainty in each bin. *Middle panel*: residuals of the same models against their best-fit MRP characterisation. The residuals are within 5% over the reliable range of masses, with the exception of a feature in the AGN runs at  $M \approx 10^{12.5} h^{-1} M_{\odot}$ . *Lower panel*: residuals of the models against the HMF fit of Tinker et al. (2008). The systematic residuals are of order 20% for baryon models, and 10% for pure dark matter. In particular, the low-mass slope of the Tinker function is at odds with the OWLS data.

can be seen in figure 4.5. The upper panel here shows the residual in the HMF of each model at  $z = 0$  for the Planck cosmology, versus the REF simulation. All baryonic physics models tend to suppress the HMF, especially around the mass mode, with increased AGN heating exacerbating this effect. We note that the uncertainty introduced by the baryon physics is generally not less than  $\sim 5\%$ , indicating that systematics due to fitting with the MRP are quite reasonable.

The middle panel shows the residuals of each of the same models with its best-fit MRP model. Clearly the residuals are within  $\sim 5\%$  over the reliable range of masses for all models, which is within the uncertainty introduced by unconstrained physics models.

We note that the MRP cannot model the AGN runs adequately, and a feature arises in the residuals at  $M \approx 10^{12.5} h^{-1} M_{\odot}$ . This captures a relative enhancement in the HMF due to AGN feedback at these scales. Far from being detrimental, this feature may become useful in identifying the level of AGN activity in observed data.

This feature is the possible reason alluded to previously for the  $s = 0$  analysis failing. When  $s = 0$ , most of the information for the likelihood comes from low masses, which in this case are poorly modelled by MRP. Increasing to  $s = 1$  means that most of the information comes from around  $m = \mathcal{H}_s$ , which behaves more appropriately.

We note that the characteristic oscillatory pattern seen in figure 4.1 is present here to some degree, but it is difficult to determine whether it is at the same level as compared to the direct EPS calculation. The oscillations are of course enhanced by the AGN heating, though due to a completely different mechanism.

To better compare the performance of the MRP and the EPS prescription, we show the residuals of the models with the fit of To8 in the lower panel. This fit is an appropriate choice as it is an SO-based fit at  $z = 0$ . We do note that we use only the best-fit parameters as provided in To8, rather than re-fitting to the data at hand. While this is an appropriate choice, since it is how the fit is used observationally, clearly this fit cannot properly describe each of the models here at once. In particular, the AGN models have well-fitted shapes, but with a decrement in normalisation of  $\sim 15\%$ . Conversely, the DMONLY run, which would naively be well-fit by To8, has a mismatched slope at low masses, causing a systematic offset of  $\sim 10\%$ . It would appear that the MRP and EPS fits possess a similar capability for fitting baryonic models. The difference is that by construction, it is clear that the MRP must be fit on a case-by-case basis, whereas this is easy to neglect for the EPS, which could result in models that systematically deviate by 10%.

To adequately model the full baryonic physics cases, and potentially eliminate the characteristic oscillations, we might consider extending MRP. For the purposes of this study, the inaccuracies introduced by the simple MRP are not significant enough to warrant re-modelling with extra parameters, due to the inherent uncertainty in the HMF due to varying halo finders and physics models. However, we refer the reader to appendix A.3 for some details concerning possible future extensions.

We turn from analysing the residuals of the MRP fits to the structure of the parameter values across models and redshifts, which exemplifies the physical insights that may be gained from such a simple model. Figure 4.6 shows the MRP parameters determined from the various cosmo-OWLS simulations. We note that  $\mathcal{H}_s$  is roughly constant across baryon models, but shows a smooth decrease over redshift (cf. §4.4). Furthermore, the Planck cosmology exhibits both tighter constraints across models and slightly higher turnover masses.

The parameter which changes the most significantly over models is the power-law slope  $\alpha$ . Particularly, models which include AGN heating exhibit a much steeper slope, which corresponds to the suppression of halos already mentioned. The REF model, on the other hand, has a relatively shallow slope.

The value of  $\beta$  shows a significantly decreasing trend with redshift for all AGN-heated models, but only weak dependence, if any, for the other models. From the simple analysis in §4.4, we expect such a dependence. On the other hand, all values are surprisingly low, which may be due to poor resolution at the high-mass end, but is more likely due to the different truncation mass used here ( $m_{\min} = 10^{12} h^{-1} M_{\odot}$ ).

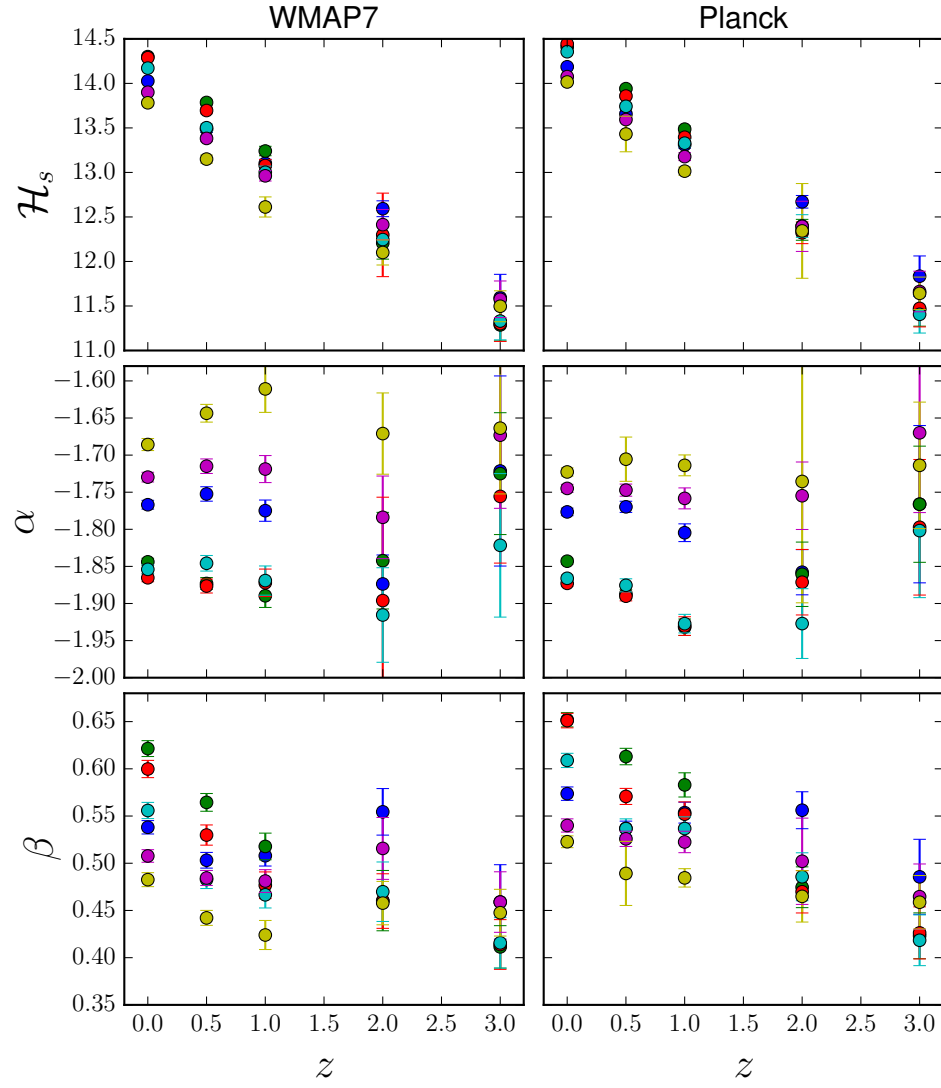


Figure 4.6: MRP parameters determined from the various cosmo-OWLS simulations. The rows of panes represent the 3 MRP parameters, while the columns represent cosmologies. The  $x$ -axis is redshift. Different coloured markers delineate baryon physics models, according to the colour-scheme in figure 4.5.

We remark that such a plot is useful diagnostic and exploratory tool in quickly assessing the properties of multiple HMF models, and the full MCMC fits for all 60 points here were produced in under 2 hours on a standard laptop.

#### 4.3.5 *Fitting halos with measurement uncertainty*

Cluster cosmology forms a crucial part of the cosmological analysis toolset. Specifically, analyses of cluster counts are sensitive both to the expansion rate and growth rate of structure, making them invaluable probes of the interplay between dark energy and general relativity (Mantz et al., 2010; Vikhlinin et al., 2009; Tinker et al., 2011). Though the likelihood models used in these analyses are complex and must take into account mass-observable relations, at some level they are reliant on fitting a theoretical HMF to measured halo masses.

This procedure is complicated by several forms of bias, not least of which is the so-called Eddington bias – that in data with symmetric uncertainties, there is a tendency for halos to be scattered into regions of lower density. In the case of the HMF, this can be restated as saying that at any measured mass, halos are more likely to be intrinsically lighter than measured than they are to be heavier than measured. This can be seen in figure 4.8, in which the red curve shows the density of uncertain masses, compared to the black line, which shows the density of the intrinsic masses. Two characteristics emerge: firstly, the sharp truncation scale is smeared out, and some halos are scattered down. Secondly, there is a push towards higher mass over the whole function, due to the aforementioned Eddington bias. Modelling this bias is intractable analytically (as the convolution of a Gaussian and GGD), but in this section we provide a method to naturally account for it statistically.

Properly modelling the uncertainties requires that each “true” mass become a parameter, which immediately increases the dimensionality of the problem to the tens of thousands – an increase which for most solvers will be intractable.

However, Hamiltonian Monte Carlo (HMC) and its derivatives, eg. the No-U-Turn Sampler (NUTS), are quite capable of dealing with such massively-multi-dimensional problems, if the likelihood can be parameterised in terms of known analytic functions. This makes MRP an ideal candidate for fitting large datasets with per-halo mass measurement uncertainties.

##### 4.3.5.1 *Development of the model*

As mentioned, the fundamental idea is to set each “true” mass as a parameter, with a prior distribution given by its uncertainty. Thus we simultaneously identify the best-fit “true” masses along with the MRP parameters. In the context of uncertain masses we no longer have a clear-cut truncation mass,  $m_{\min}$ , so this becomes a parameter in the estimation also.

Finally, we must decide whether to estimate the masses in log- or linear-space. We find that it is much simpler to transform all masses into log-space, and perform all the calculations there – not least because common uncertainty distributions for the masses

are lognormal. In this case, we transform to a parameter  $y = \log_{10} \frac{m}{\mathcal{H}_s}$ , and in the space of  $y$ , the pdf becomes

$$\ln \mathcal{L} = \sum_i^N \ln \beta + \ln \ln 10 + \ln 10 y_i (\alpha + 1) - 10^{y_i \beta} - \ln \Gamma(z, x), \quad (4.21)$$

where clearly the slope has increased by 1 due to the transformation.

We list STAN code describing this model in Appendix A.6, as we find that it more concretely defines our model choice. The code begins with a functions block, which defines custom functions for use in the model. In particular, here we have defined `gammainc`, which evaluates  $\Gamma(z, x)$  for arbitrary values of  $z$  (using the recursion relation), and also `truncated_logGGD_log`, which defines the log-likelihood according to Eq 4.21.

The data block defines values that are passed by the user, including the number of halos, their measured masses, and uncertainties. To assist with optimal choices of parameter bounds for a given dataset, we also define the boundaries as data here.

The parameters block defines the parameters that are to be fit, which include the 3 MRP parameters with their bounds, the truncation mass and a vector of “true” masses. We note that the lower-bound of `log_mtrue` is dynamically shifted according to the choice of `log_mmin` on each iteration.

Finally, the model block is the engine room of the model, and consists primarily of two lines: first an evaluation of the log-likelihood, Eq 4.21, for the current parameters and estimated masses, and secondly, an evaluation of the log-likelihood of the current iteration of estimated masses based on their uncertainty distribution.

#### 4.3.5.2 Testing on mock datasets

To test the performance of our model, we have chosen to create mock datasets loosely based on specifications from the Galaxy and Mass Assembly (GAMA) survey (Driver et al., 2011). We sample 25,000 halo masses from an MRP with  $\vec{\theta} = (14.2, -1.85, 0.72)$ , which is a good fit to the HMF in the P13 cosmology at  $z = 0.25$  (the mean redshift of GAMA). We use a truncation scale of  $\log m_{\min} = 12$ , roughly consistent with expected GAMA catalogues.

We augment these intrinsic masses with standard normal uncertainties, where  $\sigma \sim \mathcal{N}(S, 0.05)$  dex. In reality, we expect the mass-uncertainties from GAMA to have some mass-dependence, so that smaller haloes have a higher mean dex uncertainty. However, for simplicity here we choose it to be appreciable yet constant.

For each of  $S \in (0.2, 0.3, 0.5, 1.0)$ , we perform 3 fits. The first fit uses the intrinsic masses before scattering, and a model which does not include the uncertainty modelling. The second fit uses the scattered masses with the same model which cannot account for the uncertainty, but in which we place an imposed truncation of mass at a best-guess of  $m_{\min}$ . Thirdly, we use the full model that we have here developed.

The estimated masses in each case follow a characteristic pattern, which is illustrated in figure 4.7. Here the  $x$ -axis shows the measured masses, while the  $y$ -axis shows the difference between the true and measured mass. The grey-scale points are the intrinsic masses, while the coloured points are the estimated masses. Clearly, the breadth of the scatter is not reproduced in the estimation. To see why this is the case, consider the case

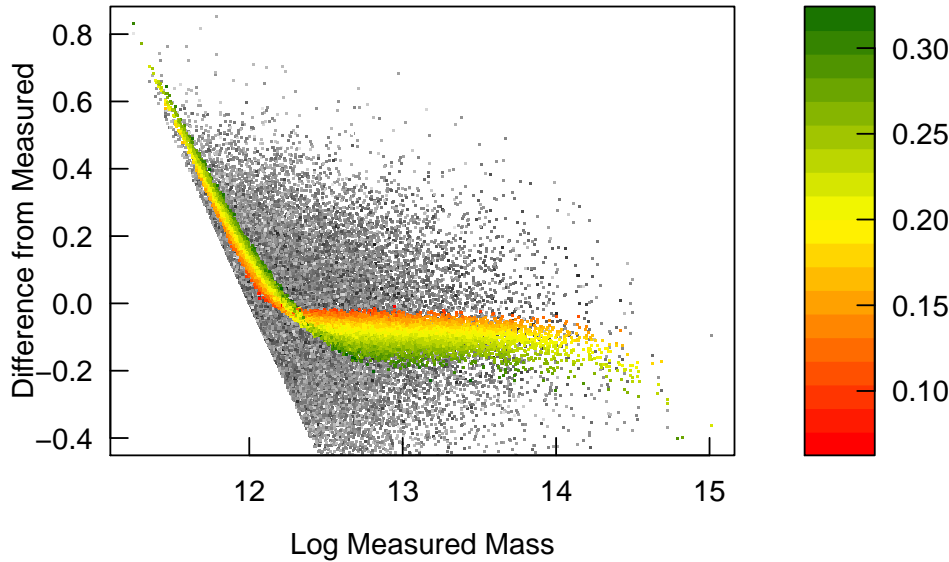


Figure 4.7: Grey-scale points are the difference between intrinsic log mass and measured log mass, with the shading indicating the uncertainty in each mass (black (white) is low (high) uncertainty, on the scale given by the colour-bar). Coloured points are the difference with respect to the estimated masses. Colouring again indicates the level of uncertainty. A clear pattern emerges in the estimated masses, corresponding roughly to a central estimator for the intrinsic masses. This was produced with  $S = 0.2$ .

when  $\sigma$  is a perfect constant. In this case, every measured mass has the same uncertainty, and so equal masses, though they may have arrived there from different intrinsic masses, must be treated equivalently. Thus the procedure at best will converge towards a central estimator for  $m_{\text{true}}(m_{\text{obs}}, \sigma)$ , rather than the full distribution (when  $\sigma$  is a constant this simplifies to a univariate function).

The form of the function  $m_{\text{true}}(m_{\text{obs}}, \sigma)$  is quite characteristic, with low masses following a linear relationship close to the line  $\Delta \log m - \log m_{\text{obs}} = 0$ , and intermediate-to-high masses flattening out with  $\Delta \log m < 0$ , before a dip at high masses. As may be expected, though the scatter in the curve is tight, it tends towards halos with small uncertainty remaining closer to  $\Delta \log m = 0$ .

One aspect to notice is that the estimated masses do not trace the line  $\Delta \log m - \log m_{\text{obs}} = 0$ , but sit marginally above it. This causes a characteristic peak feature in the density of the estimated masses just above  $m_{\text{min}}$  (see figure 4.8). We expect this behaviour since the estimated masses must trace a central estimator of the intrinsic masses, which have a sharp truncation at  $m_{\text{min}}$ , so that the central estimator must be greater than  $m_{\text{min}}$ . Though this causes the number density of halos to be mis-estimated, we find that it does so in such a manner as to generate the correct MRP parameters.

Figure 4.9 shows a pairs plot of the MRP parameters for  $S = 0.2$ , and we remark that the distributions are quite distinct from multivariate normal, exhibiting ‘banana-like’ structures and asymmetries. Thus we must use MCMC methods to adequately explore the posterior distributions, and cannot resort to approximations of the covariance matrix at the solution.

Finally, in figure 4.10 we show a comparison between the estimated parameters in each method. The blue markers show the values estimated using the sampled masses without



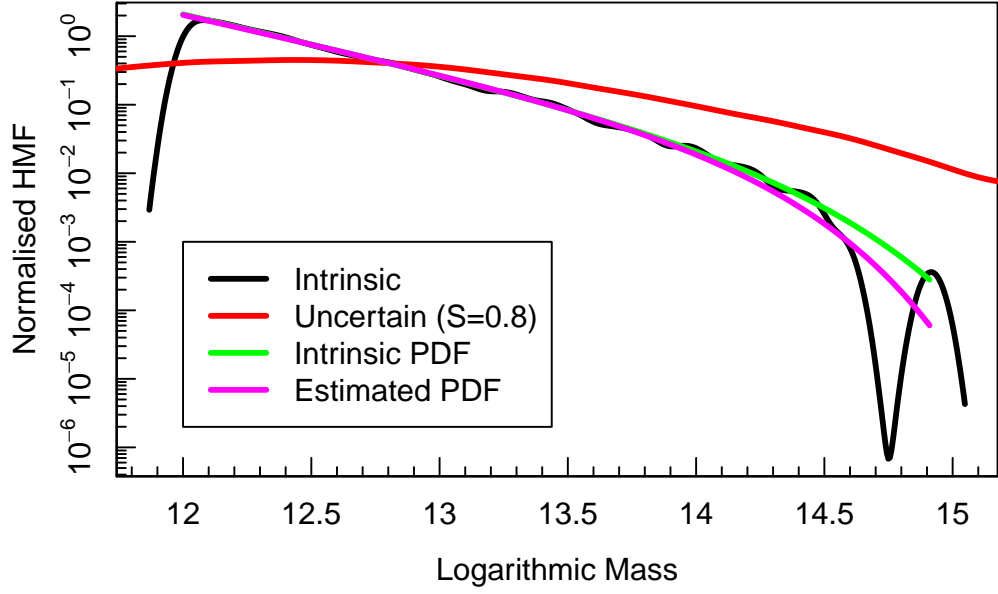


Figure 4.8: Density plots of various realisations of the halo catalogue for  $S = 0.8$ : black is the intrinsic sampled masses, red is the sampled masses with uncertainty, and blue is the estimated masses. Green and magenta are the analytic MRP pdf of the input and estimation respectively. In this sample, it is difficult to constrain  $\beta$  even with masses with no uncertainty, which causes the divergence of the analytic curves at high mass. Note the peak feature in the estimated masses just above  $m_{\min}$ .

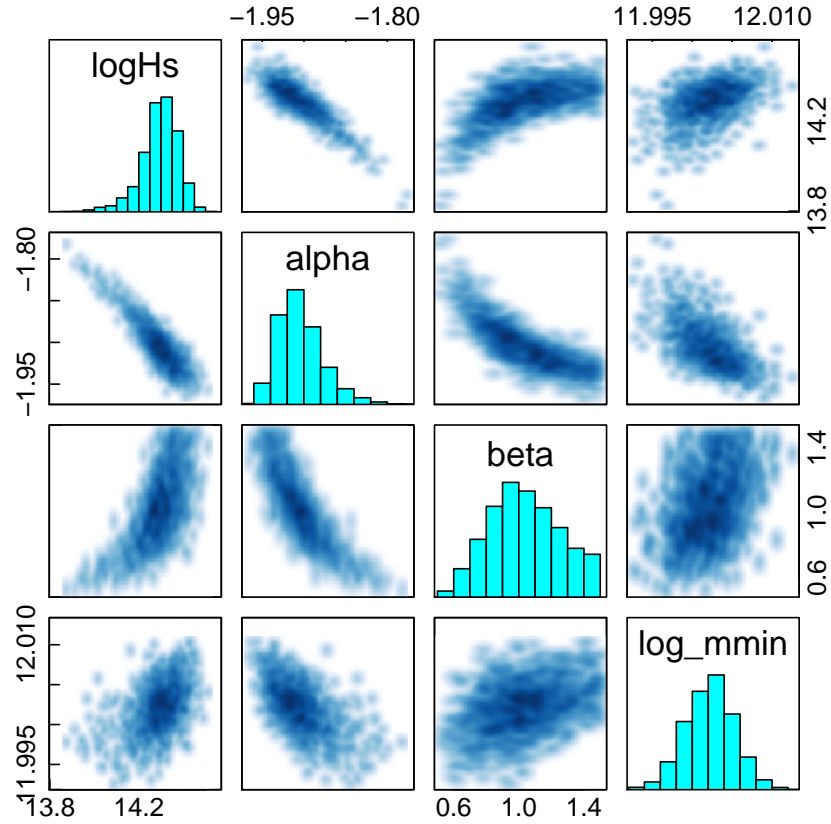


Figure 4.9: The marginalised joint-posteriors for all parameter pairs in the  $S = 0.2$  case (excluding mass estimates). Note both the high correlations and non-gaussianity in the distribution, which necessitate a full MCMC approach to exploring the posteriors.

uncertainties applied (as ratios of the input values), and thus give a baseline expectation for the achievable constraints on the parameters for a sample with the raw specifications of GAMA (ie. 25,000 halos with a minimum mass of  $10^{12}h^{-1}M_{\odot}$ ). The  $x$ -axis is the mean uncertainty applied to the masses (in each case the same intrinsic masses were used to facilitate comparison). Errorbars are derived from the MCMC chains and show the 16<sup>th</sup> and 84<sup>th</sup> quantiles.

Green markers show the results using the simple method without uncertainty modelling, while red markers show the fully modelled results. Several basic results are initially worth mentioning: firstly, in such a small sample,  $\beta$  is constrained far worse than its counterparts for all samples, though this effect is somewhat dampened for very uncertain masses. Secondly, as expected, the parameters are less constrained for more uncertain mass samples, though as we have pointed out, this effect is less pronounced for  $\beta$ .

We also note that neither of the methods completely fails even for very high uncertainty in the masses. However, it is clear that the full per-object error method results in up to double the precision, and sometimes considerably less bias than its simple alternative. The simpler method also contains an arbitrary choice of truncation mass, which will affect results (though it seems that a reasonable visual choice generates reasonable results).

We conclude that the method presented here, which fully takes into account arbitrary per-object mass uncertainties, provides exceptional constraints on the MRP parameters for samples which roughly simulate current surveys. We remark that the method has unprecedented precision even for samples with uncertainties that span 1/4 of the logarithmic mass range, and thus could form an invaluable component of next-generation HMF and cluster cosmology studies.

#### 4.4 DEPENDENCE ON PHYSICAL PARAMETERS

While it is envisaged that the MRP will generally be used as a standalone description of halo mass data, it is instructive to determine the dependence of the parameters  $\vec{\theta} \equiv (\mathcal{H}_s, \alpha, \beta, A)$  on the more important physical parameters.

The parameter with the most influence over the HMF is the redshift, which through the growth rate affects the normalisation of the power spectrum, and is also explicitly modelled into the HMF fit in several recent studies (eg. our fiducial model of To8). Along with this, Murray, Power, and Robotham (2013b) showed that the parameters  $\Omega_m$  and  $\sigma_8$  have the most significant effects across a broad range of masses. Thus, in this section we set out to model the dependence of the  $\vec{\theta}_i$  on  $\phi = (\Omega_m, z, \sigma_8)$ .

We note that the choice of base model to fit to is somewhat arbitrary. Fits in the literature vary at the level of tens of percent across interesting mass ranges. Furthermore, there is a dichotomy between fits based on haloes found using Friends-of-friends (FOF) methods, and Spherical-overdensity (SO) methods. Given the inherently approximate nature of this dependency modelling, we opt to fit to a chosen fiducial model, and use the uncertainty of this fit (both intrinsically and with respect to other fits) to guide our judgement of the appropriateness of the model.

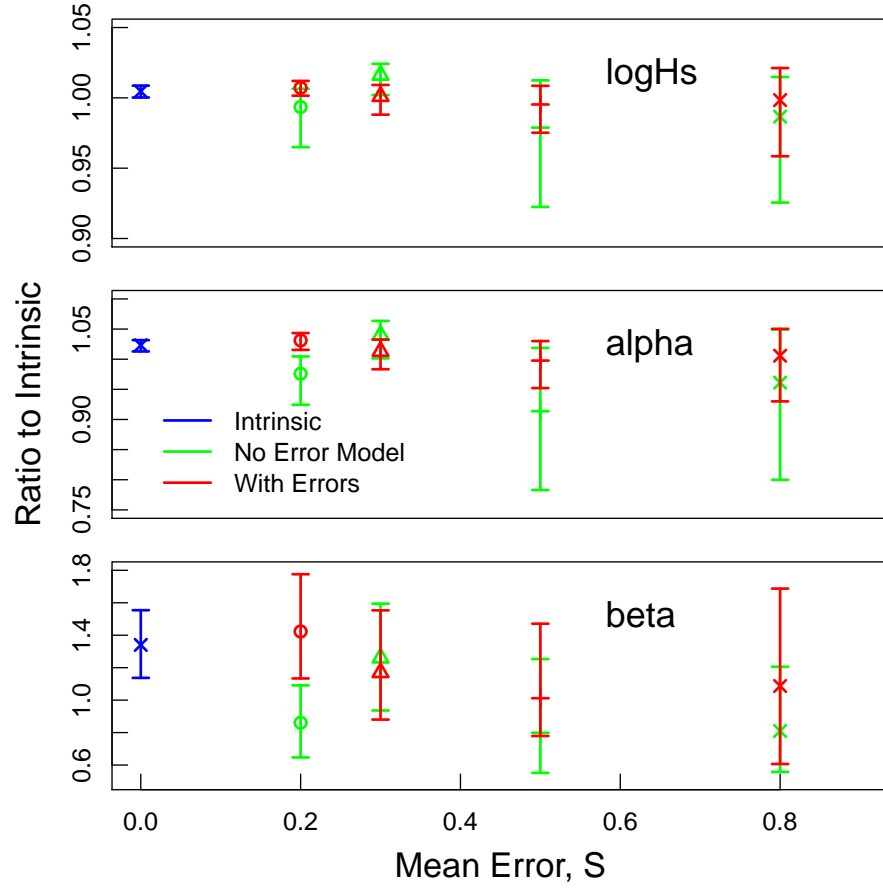


Figure 4.10: Summary of all methods to estimate MRP parameters from uncertain masses. Top to bottom the panels are  $\log \mathcal{H}_s$ ,  $\alpha$ , and  $\beta$ . The  $x$ -axis is the mean amount of uncertainty on the masses,  $S$ , and the  $y$ -axis is the ratio of the result to the true input value. Error-bars indicate 68% uncertainty regions. Blue markers indicate results using a method without error modelling on uncertainty-free masses, green markers use the same method but on masses with uncertainties, and red markers use the full error modelling.

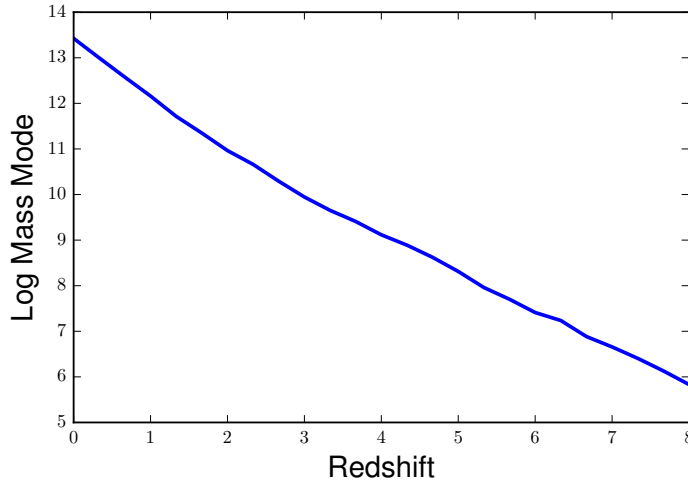


Figure 4.11: Logarithmic mass mode  $\mathcal{H}_T$  as a function of redshift for the B13 HMF.

Since we are interested in modelling both low and high redshifts, we take as our fiducial fit the form of Behroozi, Wechsler, and Conroy (2013) (hereafter B13), which is an empirical modification of To8 that increases accuracy at high redshift (up to the Epoch of Reionization).

When calculating the HMFs for modelling the MRP parameters, we ensure that in each case the wavenumber range is wide enough to encapsulate all relevant information (see §2 of Murray, Power, and Robotham (2013a) for a brief discussion of relevant limits on the product of the radii and wavenumber), and also that the resolution is high enough to avoid small oscillatory artefacts.

An important consideration is the mass range over which to fit the MRP, as this choice affects the derived parameters (cf. figure 4.1). In order to probe the domains of influence of all three shape parameters, and to provide a standardized system, we opt to define the mass range as constant with respect to the logarithmic mass mode  $\mathcal{H}_T$ . This value is calculated explicitly as the zero of the derivative of a quartic spline interpolation of the B13 HMF for each  $\phi_i$ . Figure 4.11 shows  $\mathcal{H}_T$  as a function of redshift for fiducial values of  $\Omega_m$  and  $\sigma_8$  (the dependence on these latter parameters is comparatively negligible). We have verified that  $\mathcal{H}_T$  is within the valid mass range of B13 for all samples.

It must be urged that the fitted parameters can change significantly given a different truncation mass with all other parameters fixed (cf. Appendix A.3). Specifically, if a smaller mass range is of interest, a more accurate fit can generally be attained (cf. Fig. 4.1). Furthermore, an optimization of the MRP parameters for any input model will achieve a greater accuracy than using the models presented in this section (and doing so is not computationally expensive). An extension to this analysis may explicitly include truncation mass as a parameter. However, the most important benefit of this section is the insight gained from determining the approximate relationship between physical and MRP parameters, and that remains relatively consistent with variable  $m_{\min}$ .

We produce 2000 samples,  $\vec{\phi}_i$ . The redshifts are drawn uniformly in log-space, so as to lightly increase the weight of lower redshifts. The cosmological parameters are drawn

Parameter	Formula, $\mu + \sigma f$	$\mu$	$\sigma$
$\log_{10} \mathcal{H}_s$	$\Omega_m + 1.6955\sigma_8 + 5.4038 \times 0.8845^z - 5.7651$	12.216	1.6413
$\alpha$	$3.0460\Omega_m + \sigma_8^2 + 0.175997\Omega_m z^{-1.8861z} - 1.5235 \ln(1+z)$	-1.91	0.02682
$\beta$	$6.2701\sigma_8\Omega_m + 2.0153 \times 0.53101^z - 1.8976 - 0.56778\Omega_m z$	0.50056	0.12893
$\ln(A/k(\vec{\theta}))$	$1.0587z + 0.0029051z^3 - 0.25536 - 1.3090s$ $- 0.049780z^2 - 0.32113sz - 0.42027mz$	3.9293	4.6086

Table 4.3: Parameterisations of  $\vec{\theta}_i$  as functions of  $\phi$ . All parameterisations are expressed as rescaled by the mean and standard deviation of the input data, which highlights the typical values and associated sensitivity of each MRP parameter.

from the normal distributions presented in P13, so as to highly weight values which are more likely to be physically appropriate. In summary,

$$\begin{aligned} \log(1+z) &\sim U(0, \log 9) \\ \sigma_8 &\sim \mathcal{N}(0.829, 0.012) \\ \Omega_m &\sim \mathcal{N}(0.315, 0.017). \end{aligned}$$

A value  $\vec{\theta}_i$  is fit to each model using  $\chi^2$ -minimisation, where in each case the shape parameters are fit to an appropriately normalised HMF, and the normalisation is set by the mass-weighted integral. We note that the normalisation can alternately be defined as that which reproduces the known mean density over all mass. We find that using this definition provides quite poor fits, especially at high redshift, and thus focus on the more accurate direct fits here. The reason for the discrepancies in the alternate method appears to be that the behaviour of the HMF below the halo resolution limit is undefined, and thus we may expect that the best-fit MRP at high redshifts will be a poor approximation to the as-yet-unknown true HMF below the fitted range.

To find an optimal parameterisation for  $\vec{\theta}$  as a function of  $\phi$ , we use the symbolic regression program EUREQA. This program iterates over models using a genetic algorithm, with constraints set by the user, to find underlying trends in input data. To assist the convergence of the program, we first fit a one-dimensional model in redshift only, which is the most effective parameter. Using this model as a baseline, we introduce  $\Omega_m$  and  $\sigma_8$  for a full 3D parameterisation. Each run of EUREQA produces a series of potential parameterisations of varying complexity and we choose models that offer a balance between accuracy and simplicity.

We present our chosen models in table 4.3. Each model is normalised and scaled to its mean and standard deviation, highlighting the primary dependencies amongst the parameters. A more intuitive presentation of the models can be found in figure 4.12, in which the independent dependencies of  $\vec{\theta}$  on  $\vec{\phi}$  are shown. Here,  $\sigma_8$  and  $\Omega_m$  are varied at  $z = 0$ , and otherwise take their fiducial values.

In particular we note that across the board, the most effective parameter is  $z$ , with which all three MRP parameters are anti-correlated. The MRP parameter most sensitive to the physical parameters is  $\beta$ , followed by  $\mathcal{H}_s$ , with  $\alpha$  changing very slowly (mainly with redshift).

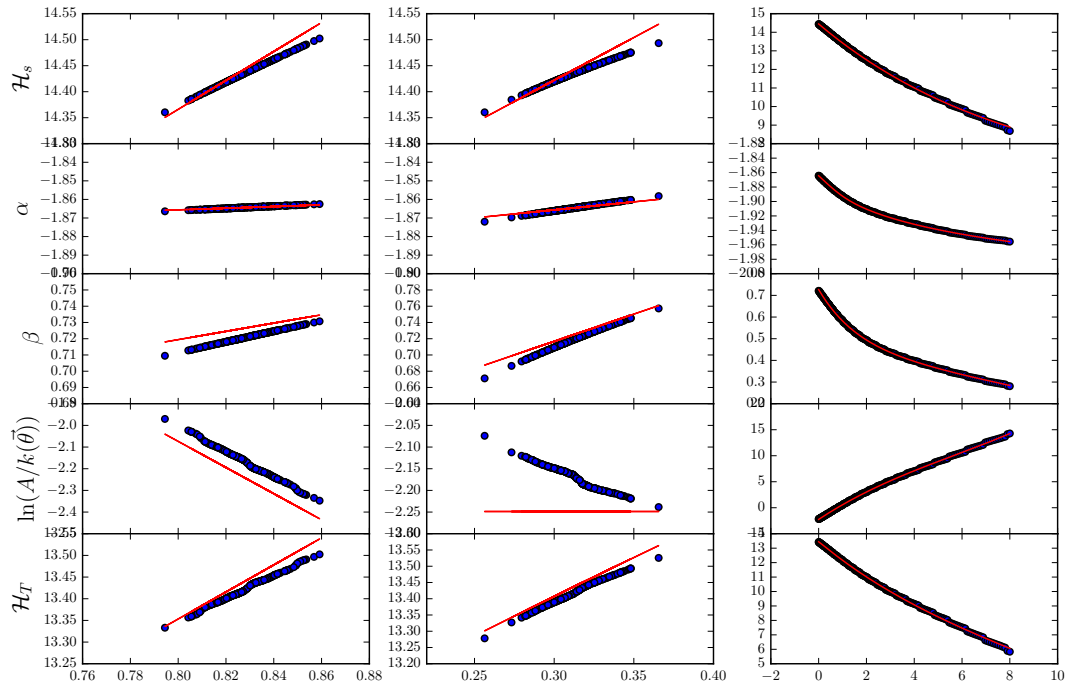


Figure 4.12: Dependence of  $\vec{\theta}_i$  (rows) on  $\vec{\phi}$  (columns). The bottom row shows the logarithmic mass mode,  $\mathcal{H}_T$ . Blue markers show best-fit values derived from  $\chi^2$ -minimization on the B13 HMF. Red lines show the EUREQA-derived models from table 4.3. The lower effectiveness of the cosmological parameters renders their independent models visually poorer, especially in the case of the normalisation.

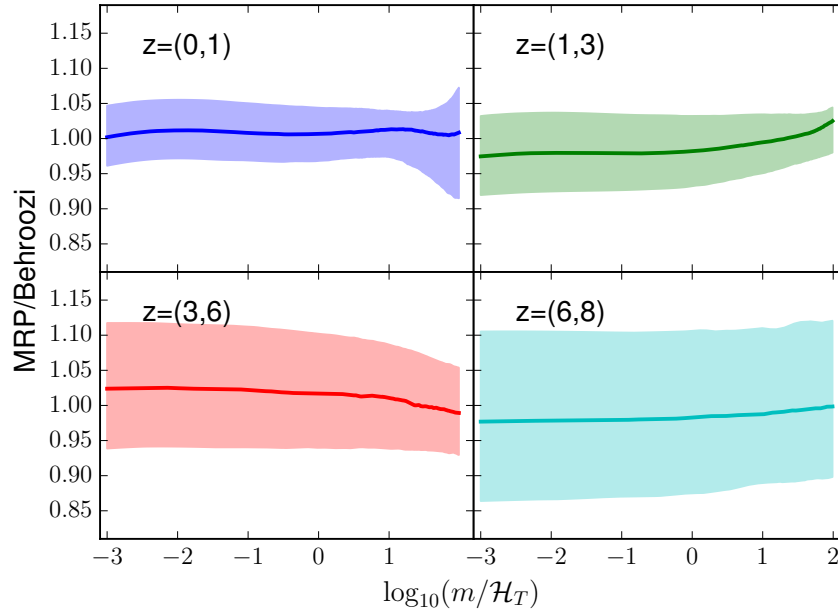


Figure 4.13: Central 68% regions of the relative error for the 3D parameterisations of the MRP parameters compared to the fits of B13. The  $x$ -axis is the mass normalised to  $\mathcal{H}_T$ , in which each sample has the same range. Each panel indicates a different redshift range, as marked. The solid line is the median value of each subsample. The uncertainty in the models sub-ten-percent over the entire redshift and mass range.

The visually poor performance of the models over independent cosmological parameters is due to the low effectiveness of these parameters. In particular, the normalisation has no dependence on  $\Omega_m$  at  $z = 0$ , which is clearly not the case in detail. This suggests that the “automatic” fits from EUREQA can be improved.

Summarily, the redshift dependence of the parameters indicates a crossover between two behaviours. At low redshifts there are two regimes – a low-mass power-law tail, and a high-mass cut-off region. At higher redshifts, the high-mass cut-off moves to become so far beyond  $\mathcal{H}_T$  that the function resembles a power-law over all relevant scales.

Figure 4.13 shows the 68% region of relative uncertainties in the models, over the distribution of physical parameters we have employed. The MRP model attained by simply using the formulas from Table 4.3 produces HMFs within 5% of B13 for almost the entire mass range between  $0 \leq z \leq 1$ , but at high redshifts can deviate by up to  $\sim 10\%$ . Given that uncertainties within the B13 HMF itself are of order 5-10%, in addition to uncertainties due to cosmology, choice of fitting function and halo finder, this is accurate enough for our purposes.

#### 4.5 STELLAR-MASS HALO-MASS RELATION

The stellar-mass halo-mass (SMHM) relation connects galaxies to halos by relating the typical stellar mass content of a halo to its total mass. Determining the form of this relation has been the subject of numerous studies (Moster et al., 2010; Behroozi, Conroy, and Wechsler, 2010; Mutch, Croton, and Poole, 2013; Durkalec et al., 2015), and is important for quantifying the effects of physical processes such as Supernova and AGN heating.

One simple method of estimating the SMHM relation is to use observed galaxy stellar mass functions (GSFMs), along with theoretical HMFs. If using a subhalo-calibrated HMF, there must be at most one galaxy per halo. It is generally assumed also that each (sub)halo contains a galaxy, so that each halo contains exactly one galaxy. If we also assume that the relationship between a halo mass and its corresponding stellar mass is both deterministic and monotonic (an unrealistic assumption, though studies suggest that the scatter in the SMHM relation is small and constant over mass), then we may derive a relationship between  $m_\star$  and  $m_h$  by

$$n_g(> m_\star) = n_h(> m_h). \quad (4.22)$$

Typically,  $n_h(> m_h)$  is derived by the EPS formalism, and thus can be treated somewhat as a black box in this equation. Though the equation can be solved numerically for  $m_\star(m_h)$  given any observed GSMF, it is convenient to be able to solve it, at least approximately, analytically in terms of the parameters of the underlying GSMF and HMF.

The GSMF is commonly accurately parameterized as a double-Schechter function (Baldry et al., 2012), which has the integral

$$n_g(> m_\star) = \Phi_1 \Gamma\left(\alpha_1 + 1, \frac{m_\star}{M_\star}\right) + \Phi_2 \Gamma\left(\alpha_2 + 1, \frac{m_\star}{M_\star}\right), \quad (4.23)$$

where without loss of generality we will assume that  $\alpha_2 < \alpha_1$ . If we let the HMF be in the form of the MRP, we then arrive at the expression

$$\Phi_1 \Gamma\left(\alpha_1 + 1, \frac{m_\star}{M_\star}\right) + \Phi_2 \Gamma\left(\alpha_2 + 1, \frac{m_\star}{M_\star}\right) = A \mathcal{H}_s \Gamma(z_h, x_h). \quad (4.24)$$

Unfortunately, this is analytically unsolvable. However, it is easy to solve by root-finding, as it has a single root and simple derivatives. Thus a numerical procedure is able to produce the empirical SMHM relation.

This procedure is illustrated in figure 4.14, where the cumulative halo mass function is shown in blue, and the cumulative stellar mass function in green. Given equation 4.24, the procedure to calculate the ratio  $m_\star/m_h$  is to identify the ratio of the curves in the horizontal direction. This can be simply performed by splines, or more robustly by root-finding, to find the red curve in the lower panel. This curve displays the efficiency of producing stellar mass as a function of halo mass, and exhibits a characteristic peak around  $10^{12} h^{-1} M_\odot$ .

It is convenient to have a closed form for the SMHM relation (or equivalently  $f = m_\star/m_h$ ). Here we compare three parameterisations from the literature (Moster et al., 2010; Behroozi, Conroy, and Wechsler, 2010; Mutch, Croton, and Poole, 2013, hereafter Mo9, B10, M13) to our empirically-derived function, and then make some comments as to their validity in light of analytic approximations. The parameterisations are listed in table 4.4.

Figure 4.15 shows the numerical solution (corresponding to figure 4.14) accompanied by these three parameterisations. In the figure, all three parameterisations are fit by minimizing  $\chi^2$  in log-log space (these fits shown with the solid curves). This leads to reasonable agreement over all mass scales, though better agreement can be achieved around



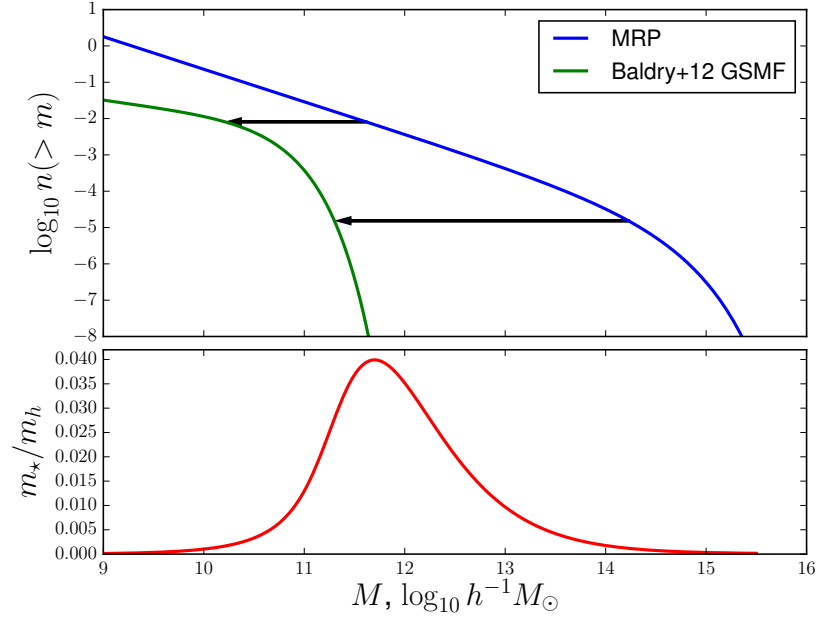


Figure 4.14: *Upper panel:* the cumulative MRP corresponding to To8 at  $z=0$  in blue, and the cumulative double-Schechter parameterisation of the  $z=0$  GSMF from Baldry et al. (2012) in green. Identifying the SMHM relation equates to computing the transform of halo mass to stellar mass, as shown by the arrows. *Lower panel:* The numerically derived fraction  $m_\star/m_h$  from the functions in the top panel. There is a strong peak efficiency at  $m_h \sim 10^{11.8} h^{-1} M_\odot$ .

REF	# par.	Formula
M13	3	$\epsilon \exp \left( - \left[ \frac{\log_{10} m_h - \log_{10} m_{peak}}{\sigma} \right]^2 \right)$
M09	4	$2f_0 \left[ \left( \frac{m_h}{m_1} \right)^{-\beta} + \left( \frac{m_h}{m_1} \right)^\gamma \right]^{-1}$
B10	5	$\frac{m_\star}{m_0 \left( \frac{m_\star}{m_1} \right)^\beta 10^{\left( \frac{m_\star}{m_1} \right)^\delta / \left( 1 + \left( \frac{m_\star}{m_1} \right)^{-\gamma} \right) - 1/2}}$

Table 4.4: Parameterisations of the SMHM relation from the literature used in this study. Note that B10 is in terms of  $m_\star$  rather than  $m_h$ , and so must be numerically inverted for comparison here (except in limiting cases where it can be analytically inverted). Note also that parameters with the same names do not have the same meaning across parameterisations.

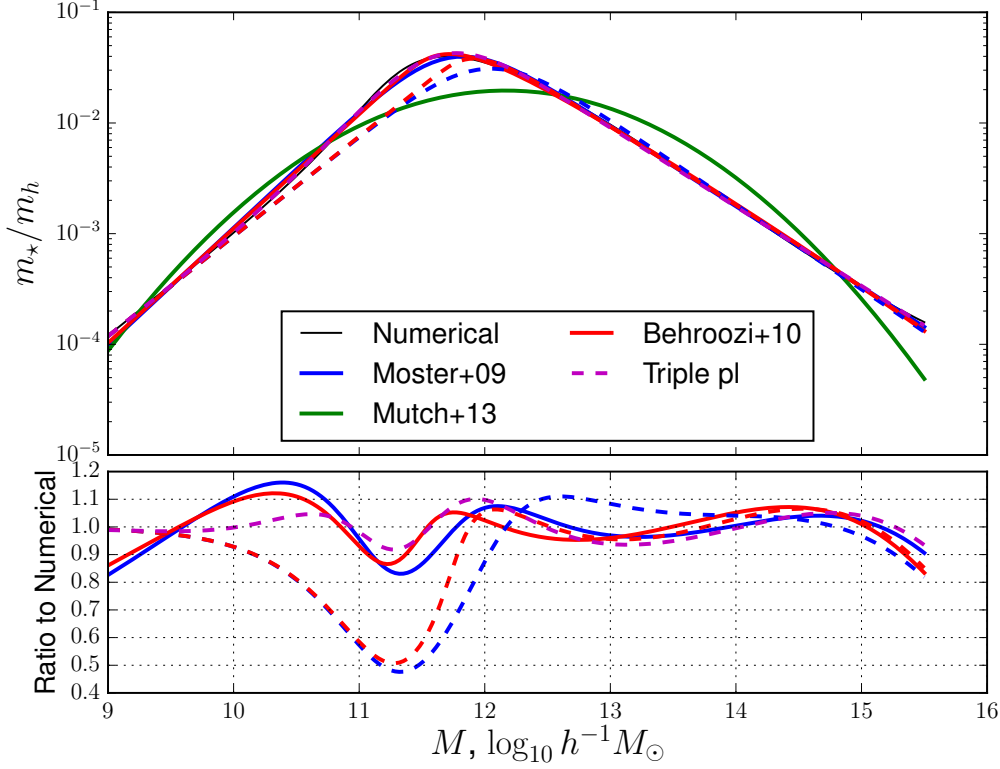


Figure 4.15: The numerically calculated SMHM relation (black), with three best-fit models from the literature. The lower panel shows their ratios. Solid lines are best fits in log-log space, with all parameters free, while dashed lines for Mo9 and B10 are fits with two parameters fixed based on the low-mass approximation.

the turning point by performing the fit in linear space, and even better by weighting the fit by  $f$ . For our purposes, we are more interested in the behaviour of the fit across a large range of scales, so the log-log fit is ideal.

It is clear that the simple form of M13 is not able to capture the behaviour of the SMHM relation adequately over a wide range of scales. However, both Mo9 and B10 provide reasonable approximations (within 20%) for the entire range. The poorest performance is at the turnover for both models, arguably the most important regime, with the extra parameter of B10 helping it get closer to the solution. Furthermore, both parameterisations clearly underestimate the ratio at very low masses in their best-fit form. Interestingly, both Mo9 and B10 do roughly equally as well at high masses, with Mo9 performing slightly better. This is surprising, as one of the primary reasons for the extra parameter in B10 was to ensure proper behaviour at high mass.

We now investigate the behaviour of the parameterisations based on what we can learn from analytic approximations to the solution. Though we cannot solve Eq. 4.24 for all  $m_h$ , we do well to identify its behaviour in the limits.

As  $x \rightarrow 0$ , we have the identity  $\Gamma(z, x) \rightarrow -x^z/z$ , for  $z < 0$  (this can be simply derived from the recursion relation, Eq. A.1). The speed of convergence of this limit depends heavily on the value of  $z$  (the more negative, the faster the convergence).

In this case, at small mass, the  $\alpha_2$  term dominates the  $\alpha_1$  term (which generally has positive shape parameter), so we can simply write

$$-\frac{\Phi_2 \left( \frac{m_\star}{M_\star} \right)^{(\alpha_2+1)}}{\alpha_2 + 1} = -\frac{A \mathcal{H}_s x_h^{z_h}}{z_h}, \quad (4.25)$$

so that we find that the ratio is a power law:

$$\frac{m_\star}{m_h} = K m_h^p, \quad (4.26)$$

$$p = \frac{\alpha_h - \alpha_2}{\alpha_2 + 1} \quad (4.27)$$

$$K = M_\star \mathcal{H}_s^{-\frac{\alpha_h}{\alpha_2+1}} \left( \frac{A(\alpha_2 + 1)}{\Phi_2 z_h} \right)^{\frac{1}{\alpha_2+1}}, \quad (4.28)$$

While M13 cannot replicate this behaviour in the low-mass limit, both Mo9 and B10 can. Specifically, Mo9 is equivalent in the low-mass limit if  $\beta = p$  and  $2f_0/m_1^p = K$ , while B10 requires  $\beta = 1/(p+1)$  and  $m_1(\sqrt{10}/m_0)^{1/\beta} = K$ .

Using these relations to fix two parameters in each of Mo9 and B10, we again find the best-fit models for these parameterisations, and plot the results as the dashed lines in figure 4.15. It is clear that the two solutions are identical and asymptotically approach the numerical solution at low mass. By setting these parameters we have decreased the flexibility of the parameterisations and therefore decreased the goodness-of-fit, especially around the turning point, however this solution has a more fundamental motivation.

Specifically, if we are to find an optimal parameterisation, it should be able to reproduce the low-mass and high-mass behaviour, ideally along with correctly identifying the turning point. The high-mass limit also affords an analytic asymptotic approximation, however unfortunately we are more interested in scales  $m_\star/M_\star \lesssim 10$  (higher masses are almost non-existent). In figure 4.15, one can just begin to see the eventual asymptotic behaviour as the slight upturn in the numerical solution at the most extreme masses. We argue that getting this behaviour correct is not important, but the behaviour between  $m = 10^{12} - 10^{13}$  is.

Unfortunately, we have not been able to derive an approximation in this regime. However, the power-law behaviour encoded in Mo9 seems sufficient for the present.

The most severe deficiency of both models, and this is exacerbated by correctly setting the low-mass behaviour, is the position of the turning point. A simple way to change the position of the turning point without affecting the behaviour in the limits, is to use an extension of Mo9:

$$\frac{m_\star}{m_h} = \frac{w \left( \frac{m_h}{m_1} \right)^{-\delta} + 2f_0}{\left( \frac{m_h}{m_1} \right)^{-\beta} + w \left( \frac{m_h}{m_1} \right)^\gamma + k}, \quad (4.29)$$

where  $w$  controls the peak position,  $k$  is able to correct the amplitude of the high-mass power-law, and  $\delta$  adds the flexibility needed to induce an upturn left of the turning point. This extension obeys the same relations as Mo9 in terms of the low-mass approximation, but has the difference that the high-mass power law has the slope  $\delta - \gamma$ .

CASE	$f_0/m_0$	$m_1$	$\beta$	$\gamma$	$\delta$	$k$	$w$
Mo9 Free	0.0379	11.735	1.0596	0.7269			
B10 Free	12.299	10.694	0.4915	1.963	0.5684		
Mo9 Fixed	0.0299	12.01	0.9031	0.7667			
B10 Fixed	12.52	10.8	0.5255	2.958	0.5977		
Triple pl Fixed	0.01871	11.7823	0.9031	1.805	0.0713	8.734	15.28

Table 4.5: Best-fit parameters for all SMHM parameterisation cases.

Using the two fixed parameters from these relations, the resultant best-fit curve is plotted in magenta in figure 4.15 (“Triple pl”). We find much better agreement in terms of the turning point position, though the precise shape of the maximum is clearly still not reproduced. With the restrictions enforced, the model has the same number of parameters as the original B10 form, but has more satisfying properties and a better residual (at  $< 10\%$ ).

For completeness, we list the values of the best-fit parameters (except for M13) in table 4.5.

#### 4.6 CONCLUSION

We have introduced a new analytic formula to describe the halo mass function over a broad range of scales, without needing to calculate relatively expensive transfer functions and numerical integrals. We have shown that this formula is an accurate model, within the uncertainties inherent in the mass function due to halo finders, halo definition and simulation resolution. We did however note that there seem to be some characteristic oscillations in the residuals when compared to fully modelled HMFs, dependent on the truncation mass.

We presented in some detail the machinery required to use the MRP form to fit to halo catalogues – both simulated and observed – and specifically focussed on sources of error and characteristic difficulties that arise. Furthermore, we showed the utility of using the simple MRP for fitting suites of simulations simultaneously, and to gain insight into hydrodynamic simulations suites. Importantly, we developed the Bayesian methods required to incorporate arbitrary per-object measurement uncertainties into the analysis, showing the power of this efficient method to characterise the posterior uncertainty in the HMF.

As examples of corollary applications of the MRP form, we determined simple fitting functions for the dependence of the MRP parameters on the physical parameters  $z$ ,  $\Omega_m$  and  $\sigma_8$ , showing that each of the parameters primarily has a negative correlation with redshift. Furthermore, we showed that it is possible to analytically motivate more precise parameterisations of the stellar-mass-halo-mass relation by using the known properties of the MRP in conjunction with the common double-Schechter form for the stellar mass function.

We have highlighted several areas in which the work presented here might be improved throughout the course of this chapter. In particular, proper investigations into

alternate parameterisations which have better covariance properties, extended models like a “double-MRP” which could improve accuracy significantly over wide mass ranges, physical dependence fits which include the truncation mass as a parameter, and improved models for the SMHM relation.



---

HALOMOD : A NEW AND FLEXIBLE HALO MODEL CALCULATOR

---

## ABSTRACT

The halo model is a successful framework for describing cosmological spatial statistics – from weak lensing observables to galaxy 2-point correlation functions. We review the basic formulation of the halo model and several of its components in the context of galaxy two-point statistics, developing a coherent framework for its application.

We use this framework to motivate the presentation of a new tool for simple and efficient calculation of halo model quantities, called HALOMOD. This tool is algorithmically efficient, simple to use, comprehensive and importantly provides a great deal of flexibility in terms of custom extensions. These characteristics will be especially important as the halo model is continually developed in the years to come.

## 5.1 INTRODUCTION

The continuing success of large galaxy surveys (eg. 2dFGRS, Colless et al. 2001; SDSS, SDSS Collaboration 2008; 6dFGS, Jones et al. 2009; WiggleZ, Parkinson et al. 2012; GAMA, Driver et al. 2011) has resulted in a period of unprecedented development in our understanding of physical cosmology and galaxy formation. Current and future instruments and surveys (eg. BOSS, Dawson et al. 2013; Euclid, Laureijs et al. 2011; LSST, Ivezić et al. 2014) will expand survey sizes by orders of magnitude, leading to a new era in precision cosmology based on large-scale spatial statistics. The challenge of such a promising future is that to maximally benefit from the data, the theoretical models underpinning interpretation must be not only well understood, but well-implemented, to enable efficient and insightful analysis.

There are a diverse network of such models, but we focus here on a particularly successful description of the galaxy population, called the halo model (hereafter HM; Neyman, Scott, and Shane, 1953; Peacock and Smith, 2000; Seljak, 2000; Ma and Fry, 2000; Cooray and Sheth, 2002). The halo model describes the statistics of the dark matter density field well into the nonlinear regime, beyond the reach of perturbation theory. It does so by combining linear theory predictions with empirical properties of dark matter haloes, via the assumption that the sum total of dark matter resides in these clumps, and that a handful of simple functions can universally describe them.

In combination with a *halo occupation distribution* (HOD) model (Kauffmann, Nusser, and Steinmetz, 1997; Scoccimarro et al., 2001; Berlind et al., 2003; Zheng et al., 2005),

the predictions of the halo model can be extended to galaxy populations, and therefore used to model clustering in large galaxy surveys. In principle, any clustering statistic on any scale is accessible through the halo model (or HOD) formalism (Zehavi et al., 2011), from real-space or projected two-point correlation functions (2PCFs), to galaxy-galaxy lensing, to higher-order correlations.

In practice, the HOD formalism has been widely used in the interpretation of galaxy populations in the past decade. Most of these studies have focused on determining the parameters of the HOD (i.e. the galaxy-halo connection) from the 2PCF of galaxies (Moustakas and Somerville, 2002; Bullock, Wechsler, and Somerville, 2002; Magliocchetti and Porciani, 2003; Yan, Madgwick, and White, 2003; Zheng, 2004; Zehavi et al., 2005; Hamana et al., 2006; Zheng, Coil, and Zehavi, 2007; Blake, Collister, and Lahav, 2008; Brown et al., 2008; Quadri et al., 2008; Wake et al., 2008a; Zheng et al., 2009; Ross, Percival, and Brunner, 2010; Zehavi et al., 2011; Wake et al., 2011; Beutler et al., 2013; Mostek et al., 2013; Palamara et al., 2013; Dolley et al., 2014; Guo et al., 2014; Skibba et al., 2014; Skibba et al., 2015; Kim et al., 2015; McCracken et al., 2015) using the analytical framework that we present in this chapter. However, other observables such as galaxy-galaxy lensing (Mandelbaum et al., 2006), the 2PCF of radio galaxies (Wake et al., 2008b; Kim et al., 2011), galaxy-quasar cross-correlations (Shen et al., 2013) and Near-UV cross-correlations (Krause et al., 2012) have also received application through the same framework. Furthermore, several recent studies have successfully combined observables to aid the breaking of degeneracies (Leauthaud et al., 2011; Leauthaud et al., 2012; More, 2013), and the combination of the galaxy 2PCF with weak-lensing (More et al., 2015), stellar mass functions (Coupon et al., 2015) and group mass-to-number ratios (Reddick et al., 2014) has resulted in the ability to simultaneously constrain cosmological parameters along with those of the HOD.

Several related approaches that employ different means for connecting galaxies to halos have been developed, including conditional luminosity functions (CLF; Yang, Mo, and Van den Bosch, 2003; Bosch, Yang, and Mo, 2003; Yang et al., 2005; Cooray and Milosavljević, 2005a; Cooray and Milosavljević, 2005b; Cooray, 2006; Bosch et al., 2013; Cacciato et al., 2013) and subhalo abundance matching (SHAM; eg. Conroy, Wechsler, and Kravtsov 2006).

Clearly the HM framework, complemented by an HOD or comparable mechanism, is of wide utility in the interpretation of large surveys. The frequency of application of the framework is also growing, suggesting that it will be an invaluable tool in the years to come.

This highlights the need for the HM framework (and its several constituents) to be both well-understood and well-implemented. Having a reliable, flexible and simple implementation of the HM framework is crucial to the next-generation of modelling for large galaxy surveys, as it enables research to be focussed on development of more sophisticated models, assessment of systematics, and combination of observables, rather than the basic but time-consuming minutia of ensuring reliability in the calculations. Indeed, it is often the case that with a new and reliable tool comes new discovery – and this is not the case only for instrumentation.



However, while the HM has become an increasingly popular tool, it is still difficult to find a publicly-available, stable and efficient code for its calculation. To be clear, it is not that implementations do not exist (eg. `CHOMP`<sup>1</sup>, `HMcode`<sup>2</sup> and `AUM`<sup>3</sup>), nor that they are poor implementations in general (indeed, all of these were useful in testing our own code). It is however our experience from conversations within the community that a remarkable number of researchers are using the HM in various ways, but almost ubiquitously they find need to write their own code – whether based heavily on existing (public or private) code or from the ground up. The reasons for this are potentially many and varied – difficulty of discovery of existing public code, difficulty of understanding existing codes (either the API or internal algorithms), a requirement for the implementation to be in a specific language, potential inadequacy of the code with respect to accuracy or performance, general inflexibility of the code leading to inability to perform an intended application, or even the overhead associated with learning the code usage when the intended application is not a high priority (we think here of observational astronomers for whom generating a fiducial 2PCF for their data may be a useful diagnostic).

What is clear is that whatever the reason for avoidance of available codes, the process of re-implementation is both time-consuming and fraught with the danger of subtle errors and inconsistencies. Indeed, while the HM is seemingly reasonably simple, several of the integrations involved are rather delicate and there are many instances where subtleties of interpretation lead to varying results. In short, it is typically at best a waste of the researcher’s time, and at worst a way to proliferate errors<sup>4</sup>.

Summarily, it seems advantageous to have a publicly-available, stable, flexible and simple code that implements the HM framework. Such a code would be useful as (i) a baseline standard for user-specific private codes (ii) a simple interface for those not actively researching in the field, but who may wish to calculate clustering statistics for their data (iii) a tool for fast exploratory analysis, and comparison between models and (iv) a stable framework for more rapid development of theoretical extensions to the HM, and modelling of its various components.

To this end, we present such a code, the `HALOMOD` package, which inherits (both philosophically and technically) from the `HMF` halo mass function package presented in Chapter 2. This package offers a comprehensive suite of component models, all integrated seamlessly into the broader HM framework. Its top-level architecture supports ease-of-use and algorithmic efficiency, and its component-level structure provides plug-and-play capabilities based on Python’s class inheritance system. We note that this package focuses on the *analytic* HM, whereas other tools (eg. `HALOTOOLS`<sup>5</sup>) are dedicated to halo modelling which samples galaxies in halo catalogues.

<sup>1</sup> <https://code.google.com/p/chomp/>

<sup>2</sup> <https://github.com/alexander-mead/HMcode/>

<sup>3</sup> <https://github.com/surhudm/aum>

<sup>4</sup> We acknowledge the counter-argument that writing such code is the best way to learn the field. However as a general counter, this is like saying that one should rewrite CAMB if they use a transfer function. For most, that is clearly over-kill, and testament to the reliability that CAMB engenders. For those to whom this is true, CAMB becomes a very useful standard against which to test.

<sup>5</sup> <http://halotools.readthedocs.org/>

This chapter is structured as follows: §5.2 details the theory of the HM particularly in the context of the galaxy 2PCF, collating the various components involved in a manner consistent with our implementation. Following this, we describe our code and its usage in §5.3, and in §5.4 we present an illustrative example. In §5.5 we define a prospectus for the future, before summarising and concluding in §5.6.

## 5.2 THE HALO MODEL

The broad assumption underpinning the HM is that in hierarchical structure formation scenarios, all mass is expected to be bound to halos at some scale<sup>6</sup>. If this is the case, then the entire nonlinear density field may be reconstructed by summing contributions from the individual halos. If, in addition, we may describe the average radial density profile of halos as spherically-symmetric, with a shape that depends solely on the mass of the halo itself, we can write:

$$\rho(\mathbf{x}) = \sum \rho_{h,i}(|\mathbf{x} - \mathbf{y}_i|, m_i) \quad (5.1)$$

where  $\rho_h(r|m)$  is the density of a halo with mass  $m$  at radius  $r$ .

The application of the HM rests in converting the premise of Eq. 5.1 into a semi-analytic integration. This requires knowledge of three key components:

1. The average radial density profile of halos,  $\rho(r|m)$
2. The abundance of halos of a given mass (termed the halo mass function, or HMF),  $n(m)$
3. The expected overdensity of halos of mass  $m$ ,  $\delta_h(\mathbf{x}, m)$  in a region with a given overdensity of mass  $\delta(\mathbf{x})$ , called the *halo bias*.

In practice, considering two-point clustering also requires a knowledge of the spatial clustering of halos, most commonly specified through the spatial clustering of matter. Furthermore, detailed calculations require the following extra components:

4. A concentration-mass relation,  $c(m)$ , linking the shape of a halo profile to its mass.
5. A model for “halo exclusion”, which accounts for double-counting intra-halo correlations.
6. To extend the calculations to galaxies, a distribution function for the occupation of halos by galaxies, as a function of halo mass,  $N(m)$  (the HOD).

Though the halo model can be used to describe the density field at any level of the  $n$ -point hierarchy, its most common application to date has been two-point statistics. Thus in this outline we shall focus on the framework as it pertains to the 2PCF.

Our aim in this section is to introduce the theory in a manner conducive to our implementation, so we shall cover each of the six components in turn following the presentation of the core framework.

<sup>6</sup> This assumption is clearly an approximation, since even cold dark matter (CDM) has a free-streaming scale in the early Universe below which we expect non-virialized mass (Frenk and White, 2012; Schneider, 2014)

Throughout the following we discuss isotropic two-point statistics, which measure the over-density of pairs of points at a certain scalar separation. This can be formulated either in Euclidean space, as the correlation function  $\xi(r)$ , or in Fourier-space, as the power spectrum  $P(k)$ . We will occasionally alternate between the two formulations for simplicity of calculation. Transformation from an (isotropic) power spectrum  $P(k)$  to an (isotropic) correlation function  $\xi(r)$  can be performed using a *Hankel* (or *Fourier-Bessel*) transform, which is the 3D Fourier transform of a spherically symmetric distribution:

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty P(k) k^2 j_0(kr) dk, \quad (5.2)$$

where  $j_0$  is the zeroth-order spherical Bessel function,

$$j_0(x) = \frac{\sin x}{x}. \quad (5.3)$$

### 5.2.1 Clustering Framework

It is convenient to formulate the framework of clustering in two regimes (Seljak, 2000), intra- and inter-halo pairs, called the 1-halo and 2-halo terms respectively. These approximately correspond to small- and large-scale structure (i.e. the contribution of each in the opposite regime is negligible), where ‘small’ is sub-megaparsec, and ‘large’ is  $> \sim 5h^{-1}\text{Mpc}$ .

Throughout the following, the superscripts  $1h$  and  $2h$  will be used to denote this segregation. Furthermore, a subscript of  $DM$  will denote a dark matter statistic, while  $g$  will denote galaxy statistics.

#### 5.2.1.1 1-halo term

The pair-counts, proportional to  $1 + \xi$ , within a halo of mass  $m$ , are given by the self-convolution of the halo profile. Thus the total pair counts, normalised by a random field (specified by the constant  $\bar{\rho}^2$ ) is given by

$$1 + \xi_{DM}^{1h}(r) = \int n(m) \left( \frac{m}{\bar{\rho}} \right)^2 \lambda(r|m) dm, \quad (5.4)$$

where  $\lambda$  is the mass-normalised self-convolution (i.e. the 3D integral of the profile multiplied by itself after a shift of length  $r$ , cf. §5.2.8), and  $n(m)$  is the halo mass function (HMF).

It is common to express this in Fourier-space, since the self-convolution becomes a simple multiplication,

$$P_{DM}^{1h}(k) = \int n(m) \left( \frac{m}{\bar{\rho}} \right)^2 |u(k|m)|^2 dm, \quad (5.5)$$

where  $u(k|m)$  is the normalised Fourier transform of the halo profile. This form has the distinct advantage that higher-order correlations merely involve an increase in the index, rather than an intractable multi-dimensional convolution. Conversely, for numer-

ical purposes, if an analytic formulation of the self-convolution exists, and the intended quantity is the correlation function, then it is more efficient to use Eq. 5.4 directly.

### 5.2.1.2 2-halo term

The two-halo term is most elegantly expressed in Fourier-space,

$$P_{\text{DM}}^{2h}(k, r) = \int \int I_{\text{DM}}(k, m_1) I_{\text{DM}}(k, m_2) P_{hh}(k, r | m_1, m_2) dm_2 dm_1. \quad (5.6)$$

where

$$I_{\text{DM}}(k, m) = n(m) u(k|m) \frac{m}{\bar{\rho}}. \quad (5.7)$$

Here,  $P_{hh}(k, r | m_1, m_2)$  is the power spectrum of halo centres for halos of mass  $m_1$  and  $m_2$ . This is in general a complicated function of mass and scale, but the most common approach is to approximate it by a first-order linear bias of the matter power,

$$P_{hh}^c(k, r | m_1, m_2) \approx b(m_1, r) b(m_2, r) P_m(k), \quad (5.8)$$

where  $b(m, r)$  is the first-order bias at  $m$  (cf. §5.2.7), possibly with a scale-dependent correction, and  $P_m(k)$  is the matter power spectrum (cf. §5.2.5)<sup>7</sup>. Typically, empirical nonlinear corrections on the linear matter power spectrum from HALOFIT (Smith et al., 2003) are used for this quantity. We leave its definition free in our implementation, since we may expect that it will be heavily developed in the coming years (cf. Jose, Lacey, and Baugh, 2015).

The general form for the 2-halo term then becomes

$$P_{\text{DM}}^{2h}(k, r) = \int \int I_{\text{DM}}(k, m_1) I_{\text{DM}}(k, m_2) \times b(m_1, r) b(m_2, r) P_m(k) dm_2 dm_1 \quad (5.9)$$

This is separable, so long as the integration limits are not inter-dependent (this is not the case in general, cf. §5.2.10), and in this simplest case we have

$$P_{\text{DM}}^{2h}(k, r) = P_m(k) \left[ \int I_{\text{DM}}(k, m) b(m, r) dm \right]^2. \quad (5.10)$$

### 5.2.1.3 Total

The total power spectrum and 2-point correlation function can be written simply as

$$P(k) = P^{1h}(k) + P^{2h}(k), \quad (5.11)$$

and

$$\xi(r) = [1 + \xi^{1h}(r)] + \xi^{2h}(r). \quad (5.12)$$

<sup>7</sup> Note that throughout, we use a subscript  $m$  to denote either the linear matter power or the nonlinear matter power derived using analytic corrections, as opposed to the full halo model.

### 5.2.2 HOD models

Before we can extend the previous theory to a treatment of galaxies, we must have a model for the expected number of galaxies within a given halo,  $P(N|m)$ , termed the halo occupation distribution. Note that this model is assumed to depend solely on the halo mass, which perhaps surprisingly captures the majority of the behaviour of interest. There is a known second-order dependence on environment, the so-called ‘assembly bias’ (Sunayama et al., 2015), but we do not pursue it in this work.

Any HOD model is principally composed of two parts – a parametrisation of the mean occupation of halos of a given mass,  $\langle N_t \rangle_m$ , and a distribution about this mean  $P(N_t|\langle N_t \rangle_m)$  (note that throughout this work, we imply that any values  $\langle N \rangle$  are functions of mass, and drop the notation from here on). The mean occupation plays the major role in providing a form for the mass dependence of typical galaxy abundances, while the distribution is crucial in calculating pair (and triplet etc.) count statistics.

Historically, the parametrisation of the mean occupation began as a single function, but early work (Kauffmann et al., 2004; Zheng et al., 2005; Zehavi et al., 2005) showed an advantage in treating ‘central’ and ‘satellite’ galaxies independently. In hydrodynamical simulations, these are empirically identifiable classes of galaxies, and semi-analytic methods also treat them separately. Central galaxies tend to be large and bright, and occupy the region near the halo’s central potential well. Satellite galaxies are less bright, and tend to follow the halo’s density profile.

Finally then, we have a mean occupation function for each class, and the total occupation is their sum:

$$\langle N_t \rangle = \langle N_c \rangle + \langle N_s \rangle. \quad (5.13)$$

Identifying these classes as independent is helpful in determining the distributions about the means. Halos may only contain a single central galaxy (or none at all if the sample selection does not allow it). Thus their distribution is simply Bernoulli (or Binomial with a single trial):

$$N_c \sim \text{Bern}(\langle N_c \rangle). \quad (5.14)$$

Satellite galaxies have been found to approximately follow Poisson distributions (Kauffmann et al., 2004; Zheng et al., 2005), a condition which we enforce throughout:

$$N_s \sim \text{Pois}(\langle N_s \rangle). \quad (5.15)$$

Most studies also impose what we will term the ‘central condition’. Explicitly, the condition states that  $\langle N_c \rangle = 0 \Rightarrow \langle N_s \rangle = 0$ . It arises from the fact that central galaxies are more luminous than their satellite counterparts, and therefore for any sample selection based on luminosity (or similar) thresholds, the central galaxy will be the first to be included. Our implementation does not enforce this condition (as there are conceivably sample selections which circumvent it), but does enable it by default, which simplifies some calculations.

For two-point statistics, an important quantity arising from the HOD is the mean pair counts,  $\langle N_t(N_t - 1) \rangle$ . We wish to derive this quantity in a form explicitly dependent only on the mean occupation functions  $\langle N_c \rangle$  and  $\langle N_s \rangle$ , for which we have specific parameterisations. We begin by noting, after Zheng et al. (2005), that

$$\langle N_t(N_t - 1) \rangle = \langle N_c(N_c - 1) \rangle + \langle N_s(N_s - 1) \rangle + \langle N_c N_s \rangle, \quad (5.16)$$

in which the first term on the right vanishes since  $N_c$  is either 0 or 1, and under the assumption that  $N_s$  has a Poisson distribution, we can simplify to

$$\langle N_t(N_t - 1) \rangle = \langle N_s \rangle^2 + \langle N_c N_s \rangle. \quad (5.17)$$

The last term is problematic, however if we impose the central condition, then  $\langle N_c N_s \rangle = \langle N_s \rangle$ , since we need only count halos in which  $N_c = 1$ . Thus, if the central condition holds, we achieve the final form

$$\langle N_t(N_t - 1) \rangle = \langle N_s \rangle(1 + \langle N_s \rangle). \quad (5.18)$$

If the parameterisations for  $\langle N_c \rangle$  and  $\langle N_s \rangle$  do not impose the central condition then we can approximate, following Zehavi et al. (2005), with  $\langle N_c N_s \rangle = \langle N_c \rangle \langle N_s \rangle$ , given that  $\langle N_s \rangle \ll 1$  when  $\langle N_c \rangle < 1$ . For  $\langle N_c \rangle = 0$  or  $1$ , it is no longer an approximation, and thus for step-function parameterisations of  $\langle N_c \rangle$ , it is completely accurate (for arbitrary parameterisations of  $\langle N_s \rangle$ ).

Finally, if the central condition is required, but the parameterisations of choice do not impose it, one can manipulate the equations to enforce it using a common technique (eg. Beutler et al., 2013), in which we set  $\langle N_s \rangle = \langle N_c \rangle \langle N_s \rangle'$ , where  $\langle N_s \rangle'$  is the original parameterisation. In this case, the truncation of  $\langle N_c \rangle$  forces the appropriate truncation of  $\langle N_s \rangle$ . One must be careful in this method, however, to ensure that the HOD parameters retain the same meaning as the original. For a step-function  $\langle N_c \rangle$  this is not a problem, but there may be subtle differences induced for smoothed truncations (these are likely to be negligible for reasonable models).

There are several parameterisations for the HOD in the literature. In this work, and correspondingly in the halomod framework, we focus on those that utilise the separation of central and satellite galaxies. We present a summary of those included in |halomod| in Table 5.16.

Most formulations contain the same essential features. so the most important features of the HOD are present in the simplest case. For the central term: a step function characterised by a step at the parameter  $M_{\min}$ , which is heavily dependent on the sample selection. The deeper the survey, the lower this parameter will be. The satellite term is primarily characterised by a power law which quantifies the growth of galaxies with halo mass by the index  $\alpha$ , and a scaling mass  $M_1$  which is the mass at which an average halo in the sample contains a single satellite galaxy.

The various refinements to this general structure include smoothing the truncation of central galaxies assuming a log-normal distribution of halo masses at this threshold, and truncating the satellite power law at an arbitrary mass scale  $M_0$ . Further refinements

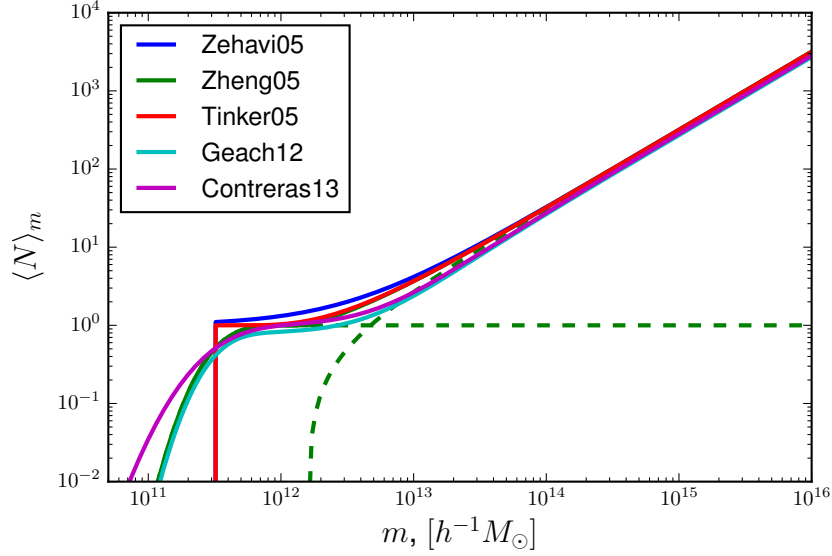


Figure 5.1: Total halo occupation for several parameterisations, highlighting the core similarities and fine differences between the models.

can be specific to a certain kind of sampling, for example samples gained at different wavelengths. An illustrative plot of the core similarities and detailed differences can be found in figure 5.1.

### 5.2.3 Extension of framework to galaxies

The extension of equations 5.4, 5.5 and 5.10 to include galaxies is relatively simple under the assumption that all galaxies reside in halos, and their intra-halo abundance follows that of the density profile.

The primary modification of the DM equations is to replace factors of  $m/\bar{\rho}$  with  $N_t(m)/\bar{n}_g$ , where  $N_t(m)$  is the mean occupation function from the HOD (where for simplicity we omit angle brackets for the rest of this work), and  $\bar{n}_g$  is the mean number density of galaxies,

$$\bar{n}_g = \int n(m) N_t(m) dm. \quad (5.19)$$

With this modification the 2-halo term is simply described by

$$I_g(k, m) = n(m) u(k|m) N_t(m) / \bar{n}_g, \quad (5.20)$$

which serves as a direct replacement for  $I_{\text{DM}}$  in equations 5.9 and 5.10.

The 1-halo term needs a little more care. The primary modification gives

$$1 + \xi_g^{1h}(r) = \frac{1}{\bar{n}_g^2} \int n(m) \frac{\langle N_t(N_t - 1) \rangle}{2} \lambda(r|m) dm, \quad (5.21)$$

however using the formalism of §5.2.2, we can separate the pair counts into contributions from central-satellite ( $c-s$ ) pairs, and satellite-satellite ( $s-s$ ) pairs (cf. Eq. 5.17). The

first of these is a one-point quantity dependent only on the profile, rather than the self-convolution. Thus we have, for the correlations

$$1 + \xi_g^{1h}(r) = \frac{2}{\bar{n}_g^2} \int n(m) [\langle N_c N_s \rangle \rho(r|m) + N_s^2 \lambda(r|m)] dm, \quad (5.22)$$

where the quantity  $\langle N_c N_s \rangle$  is treated differently depending on whether the central condition is imposed. The Fourier-space counterpart is defined likewise.

This framework provides the structure with which to combine the several components of the HM, which we describe in further detail in the following subsections.

#### 5.2.4 Derived Quantities

There are several quantities of interest which can be derived from the HM/HOD framework.

Firstly, on large scales, where  $u \rightarrow 1$ , and the 1-halo term is negligible, the 2-halo term is approximated by

$$P_{gg}(k) \approx b_{\text{eff}}^2 P_m(k), \quad (5.23)$$

where

$$b_{\text{eff}} = \frac{1}{\bar{n}_g} \int n(m) b(m) N_t(m) dm \quad (5.24)$$

is called the “effective bias”.

We may likewise calculate an “effective concentration” which is the mean halo concentration within the sample:

$$c_{\text{eff}} = \frac{1}{\bar{n}_g} \int n(m) c(m) N_t(m) dm, \quad (5.25)$$

and an “effective mass”:

$$M_{\text{eff}} = \frac{1}{\bar{n}_g} \int n(m) m N_t(m) dm. \quad (5.26)$$

An important quantity in studies of galaxy formation and evolution is the fraction of galaxies that are satellites. The evolution of this quantity through time can trace the effects of various physical processes. It is simply defined as

$$f_{\text{sat}} = \frac{\int n(m) N_s(m) dm}{\int n(m) N_t(m) dm}. \quad (5.27)$$

#### 5.2.5 Matter Power Spectrum

The matter power spectrum  $P_m(k)$  characterises the distribution of matter density perturbations as a function of wavenumber  $k$ ; formally it quantifies the average of  $|\hat{\delta}(\vec{k})|^2$ , the



square of the Fourier-transformed real-space overdensity fluctuations  $\delta(\vec{x}) = \rho(\vec{x})/\bar{\rho} - 1$ , where  $\vec{x}$  and  $\vec{k}$  are the spatial coordinate and wave vector respectively.

When density fluctuations are small, so that  $\delta \ll 1$ , the statistical properties of the perturbations can be derived analytically via perturbation theory. Generally, it is assumed that the fluctuation spectrum was initially scale-free (i.e. a power-law), and the present day (linear) power spectrum is modified by a transfer function which captures the effects of the transition from a radiation-dominated to matter-dominated Universe (Bond and Efstathiou, 1984):

$$P_{\text{lin}}(k) = Ak^n T^2(k), \quad (5.28)$$

where  $A$  is the normalisation constant and  $n$  is the spectral index. We follow convention and use the parameter  $\sigma_8$ , which measures the mass variance on a scale of  $8h^{-1}\text{Mpc}$  at  $z = 0$ , to calculate  $A$ . The transfer function is particularly sensitive to the nature of the dark matter and the baryon density parameter  $\Omega_b$ .

A commonly used dimensionless equivalent can be defined as

$$\Delta^2(k) = \frac{k^3}{2\pi^2} P(k), \quad (5.29)$$

which is the contribution to the variance in logarithmic wavenumber bins.

The CDM transfer function has been modelled extensively in the literature. At least three functional forms have been proposed (Bond and Efstathiou, 1984; Bardeen et al., 1986; Eisenstein and Hu, 1999, hereafter EH), along with a series of *Boltzmann* codes, which calculate the transfer function from first principles with perturbation theory techniques (Zaldarriaga and Seljak, 2000; Lewis, Challinor, and Lasenby, 2000; Blas, Lesgourgues, and Tram, 2011).

Figure 5.2 shows the typical shape of the CDM linear power spectrum, while highlighting differences between forms for  $T(k)$  found in the literature. Both small- and large-scales have small fluctuations, with a characteristic bend at scales of  $\sim 0.02h/\text{Mpc}$ . Only the Boltzmann code is able to fully capture the wiggles introduced by the baryon acoustic oscillations (BAOs), though the EH fit closely follows in all other respects. In fact, all curves (except that of Bond and Efstathiou 1984) are within  $\sim 10\%$  over the plotted range.

The calculation of the halo-centre power spectrum is commonly approximated by Eq. 5.8, i.e.

$$P_{hh}^c(k, r|m_1, m_2) \approx b(m_1, r)b(m_2, r)P_m(k),$$

in which an estimate of the matter power spectrum is linearly biased. Typically, a *non-linear* estimate of the matter power is used, for which the corrections of Smith et al. (2003, hereafter HALOFIT) are employed. We do not present the details here, except to mention that the modifications are applied as ratios of  $P_{\text{lin}}(k)$ , which are obtained via arguments motivated by the halo model. These modifications are expressed as a series of equations with coefficients fitted to numerical simulations. Updated coefficients were introduced in Takahashi et al. (2012). Note also that the normalisation of  $P_{\text{nl}}(k)$  is iden-

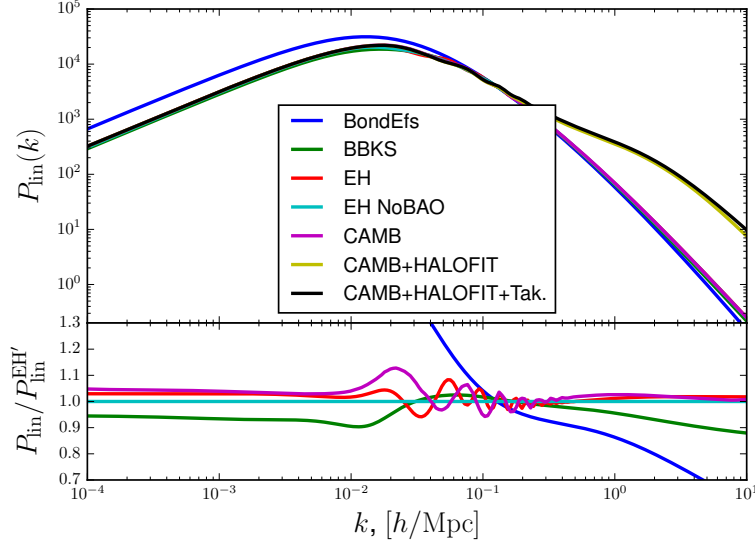


Figure 5.2: A selection of transfer function models, including analytic approximations (Bond and Efstathiou, 1984; Bardeen et al., 1986; Eisenstein and Hu, 1999) and Boltzmann solutions. The general shape is reproduced in each case, with details of baryon acoustic oscillations differing. Also plotted are two versions of the empirical nonlinear corrections.

tical to  $P_{\text{lin}}(k)$ , and should be calculated using the linear form. The effect of introducing non-linearities is to increase small-scale power (cf. figure 5.2), with the recent coefficient updates slightly enhancing this effect.

### 5.2.6 The Mass Function

The mass function quantifies the number of dark matter haloes per unit mass per unit comoving volume of the Universe.

In brief, the most common theoretical formulation of the mass function is via the extended Press-Schechter (EPS) formalism (Press and Schechter, 1974; Bond et al., 1991), in which density peaks form through gravitational instability and evolve into halos when they cross a density threshold. In this formalism, the mass function is a universal function of the mass variance,  $\sigma$ , which captures the variation from cosmological parameters and the redshift (Jenkins et al., 2001, but note that several recent studies find significant non-universality with redshift, and have provided fits that explicitly account for this departure).

The mass variance is expressed as

$$\sigma^2(R(m)) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) W^2(kR) dk, \quad (5.30)$$

where  $W(x)$  is a window function (common choices are a top-hat in real-space, Gaussian, and top-hat in Fourier-space, though most mass function studies use the first).

The mass function is then defined as

$$n(m) = \frac{\rho_0}{M^2} \sqrt{\nu} f(\nu) \left| \frac{d \ln \sigma}{d \ln m} \right|; \quad (5.31)$$

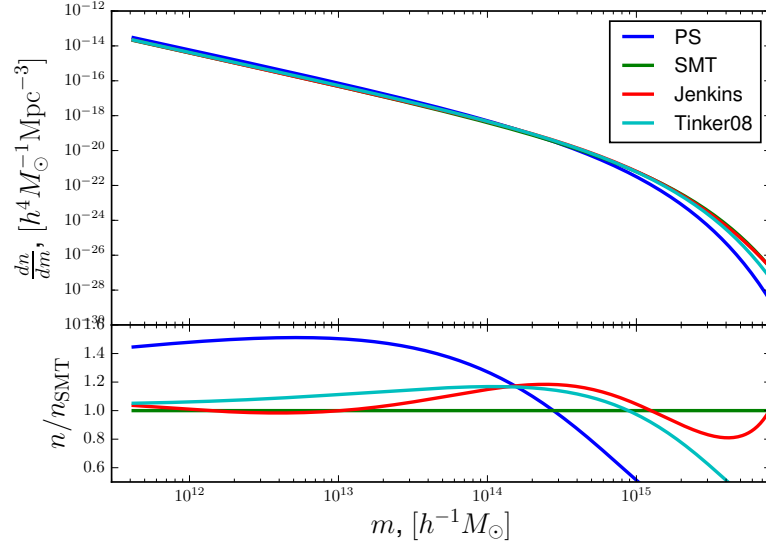


Figure 5.3: Four HMFs from the literature, including two physically motivated fits (PS motivated by spherical collapse from Press and Schechter (1974) and SMT motivated by ellipsoidal collapse from Sheth, Mo, and Tormen (2001)) and two empirically motivated fits (Jenkins et al., 2001; Tinker et al., 2008)

where  $\nu = \delta_c^2 / \sigma^2(m)$  is called the peak-height<sup>8</sup>. The function  $\sqrt{\nu}f(\nu)$  has been modelled extensively in the literature, and is generally a fit to numerical simulations.

It is advantageous (but not strictly necessary) to choose a fitting function which is normalisable, ie.

$$\int f(\nu) d\nu = 1. \quad (5.32)$$

Several fits in the literature obey this relation (eg. Sheth, Mo, and Tormen, 2001; Peacock, 2007; Tinker et al., 2010).

Figure 5.3 shows the HMF for four of the fits available in HALOMOD (via HMF). The HMF behaves as a power-law at low masses, with an exponential cut-off above a typical mass  $m_* \approx 5 \times 10^{14}$  (see eg. Chapter 4). There is clearly significant variations between the fits, especially at high masses, for a range of reasons, including halo finder differences, alternate mass definitions, and simulation resolution (cf. Murray, Power, and Robotham, 2013b; Knebe et al., 2011).

However, as seen in Figure 5.4, these variations are not as pronounced in the resulting  $\xi_{gg}(r)$ , due to integration with the variation occurring at small values. The two-halo term is modified by a near-constant factor dependent primarily on the amplitude of  $n(m_*)$ . The one-halo term has a more complicated dependency.

### 5.2.7 Bias

It is well-known that dark matter halos tend to trace the underlying density field in a biased manner. In principal, the form of this bias can be of arbitrary order, non-local

<sup>8</sup> Note that some authors use  $\nu = \delta_c / \sigma(m)$ , i.e. the square root of our definition.

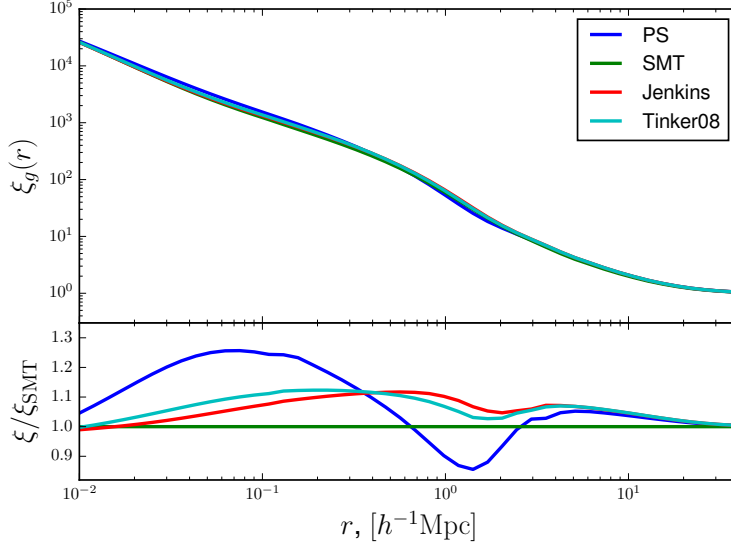


Figure 5.4: Effect of the varying HMF fits on the two-point correlation function, where the mean galaxy density is matched in each case by shifting the HOD parameter  $M_{\min}$ , where  $M_{\min} = 10^{11} h^{-1} M_{\odot}$  for the SMT function. The variation is reduced to  $< 10\%$  for fits other than PS.

(i.e. dependent on the surrounding environment) and stochastic. However, it is most common within studies involving the halo model to consider only the first-order, local, deterministic bias as a function of the halo mass (Mo and White, 1996):

$$\delta_h(\vec{x}, m) = b(m)\delta(\vec{x}). \quad (5.33)$$

The EPS formalism is able to make predictions of the form of  $b(m)$ , by considering the probability of collapse of a region of given smoothing scale  $m$  in the context of an altered density threshold  $\delta_c - \delta(\vec{x})$ , often called the peak-background split method (for more details, see eg. Bond et al., 1991; Zentner, 2007; Tinker et al., 2010; Manera, Sheth, and Scoccimarro, 2010). These predictions are specified in terms of the peak-height parameter  $\nu$ , and are therefore expected to exhibit a degree of universality with respect to cosmology and redshift.

Within this formalism, the bias is intimately related to the mass function,

$$b(m) = 1 - \frac{\partial}{\partial \delta_c} \ln n(m), \quad (5.34)$$

which when combined with Eq. 5.31, gives

$$b(m) = 1 - \frac{1}{\delta_c} - \frac{\partial}{\partial \delta_c} \ln f(\nu). \quad (5.35)$$

In this formalism, if all mass is contained within halos, according to Eq. 5.32, then by construction, the following consistency relation holds,

$$\int b(\nu) f(\nu) d\nu = 1, \quad (5.36)$$

which merely says that the average bias is unity, so that halos are not biased against themselves.

In the approximation of spherical collapse,  $f(\nu) \propto \exp(-\nu^2/2)$ , which yields a bias of (Cole and Kaiser, 1989; Mo and White, 1996)

$$b(\nu) = 1 + \frac{\nu - 1}{\delta_c}. \quad (5.37)$$

The general qualities of the spherical collapse form are (recall that  $\nu = 1$  at  $M_\star \sim 5 \times 10^{12}$  at  $z = 0$ ):

- a transition from low-mass anti-bias to high-mass bias at  $\nu = 1$
- a small-scale asymptotic limit of  $b = 1 - 1/\delta_c \approx 0.4$
- a large scale asymptotic behaviour of  $b \propto \nu$ .

Given the simplistic nature of the spherical collapse model, it is remarkable that this bias function roughly reproduces the results of  $N$ -body simulations. However, in detail, this form over-predicts the bias at high masses ( $\nu > \sim 1.5$ ), and consequently slightly under-predicts the bias for low-mass halos.

The approximations induced by the spherical collapse assumption have been relaxed in several studies, most notably by Sheth, Mo, and Tormen (2001), who obtain a form motivated by ellipsoidal collapse, but calibrated with simulations. Forms motivated by stochastic barriers (Corasaniti and Achitouv, 2011), non-Markovian barriers (Ma et al., 2011) and local density-maxima (Paranjape, Sheth, and Desjacques, 2013) are more recent developments.

Unfortunately, it appears that the formalism itself leads to an artificial enhancement of the bias for halos with  $\nu < \sim 1$  (Manera, Sheth, and Scoccimarro, 2010; Tinker et al., 2010), which has resulted in a proliferation of direct empirical fits to high-resolution  $N$ -body simulations (Jing, 1998; Jing, 1999; Seljak and Warren, 2004; Tinker and Weinberg, 2005; Mandelbaum et al., 2005; Pillepich, Porciani, and Hahn, 2010). Motivated by the broad success of the EPS formalism, these fits are usually (but see Jing, 1998; Seljak and Warren, 2004) still formulated in terms of the peak-height (indeed, often they utilise the forms yielded by the peak-background split, with updated parameters based on direct fits to simulations).

In figure 5.5 we show a sample of linear scale-independent bias relations found in the literature and included in HALOMOD. We note the similarity of high-mass behaviour, though the high-mass amplitude is substantially modified between several of the forms.

### 5.2.7.1 Scale-dependence

In reality, the assumption that the bias is a function only of halo mass, to the exclusion of environment, is overly simplistic (Sunayama et al., 2015). The scale-dependence of halo bias is a complex phenomenon, which has received a great deal of attention in recent studies (eg. Paranjape et al., 2013; Lapi and Danese, 2014; Poole et al., 2015). Foregoing

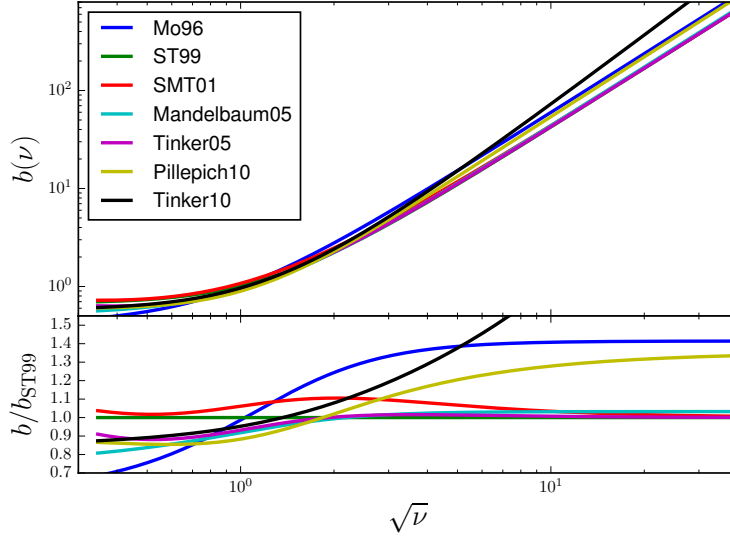


Figure 5.5: Several linear bias functions from the literature. Most functions, including Mo and White (1996), Sheth and Tormen (1999), Sheth, Mo, and Tormen (2001) and those with the same form (Mandelbaum et al., 2005; Tinker and Weinberg, 2005) have high-mass asymptotic behaviour of  $b \propto \nu$ . The form of Tinker et al. (2010) is an empirical fit with an arbitrarily chosen form, and has a high-mass asymptotic limit of  $b \propto \nu^{1.2}$ .

these complexities, simple empirical models specify the scale-dependence as a function  $S(\xi_m(r))$ ,

$$b^2(m, r) = b^2(m) S(\xi_m(r)). \quad (5.38)$$

A commonly-used form for  $S$  is given by Tinker and Weinberg (2005) as

$$S(r) = \frac{[1 + 1.17\xi_m(r)]^{1.49}}{[1 + 0.69\xi_m(r)]^{2.09}}. \quad (5.39)$$

At large scales, where  $\xi \ll 1$ , the scale-dependence is negligible, but scales around the transition can receive significant ( $\sim 50\%$ ) suppressions.

### 5.2.8 Halo Density Profiles

It is widely accepted that DM halos exhibit a universal spherically-averaged density profile, which is necessary for our formulation of the HM. We here define our nomenclature concerning this profile, which is a systematic formulation enabling the simple insertion of specific profiles of one's choice. This system is designed specifically for elegant implementation, and at this stage only considers profiles with two parameters (one of which is set by the halo overdensity  $\Delta_h$ , and the other we label as “concentration”). Multi-parameter profiles (eg. the ‘Einasto’ profile, Navarro et al., 2004) will be left for future versions.

The universal density profile has the form

$$\rho(r|m, z) = \rho_s(c) f(r/r_s), \quad (5.40)$$

where  $\rho(r|m)$  has units of density, and  $c = c(m, z)$  is a *concentration parameter* which we shall treat as an explicit function of the halo mass and the redshift (though it exhibits appreciable scatter around any such relationship). The factor  $\rho_s(c)$  controls the amplitude of the profile, and is set by the known mass inside the halo radius (here we use halo radius generically as  $r_{\Delta_h}$ , where  $\Delta_h$  is the halo overdensity, often set to 200), dependent on the concentration.  $f(r/r_s)$  controls the shape of the profile, and is a function of the scaled radius  $x = r/r_s$ , where  $r_s = r_{\Delta_h}/c$ .

Throughout the following discussion we shall often refer to a ‘reduced’ form of the profile (or related quantities) in which simple normalisation factors are removed. An example of such a reduced form is  $f$  in equation 5.40.

As mentioned, for a particular profile the amplitude is set by the total mass inside the halo radius. We may integrate Eq. 5.40 to recover the total mass of the halo:

$$m = 4\pi\rho_s(c)r_s^3 \int_0^c x^2 f(x) dx, \quad (5.41)$$

where we have made the substitution  $r = xr_s$  in which  $x(r_{\Delta_h}) = c$ . If we set

$$h(c) = \int_0^c x^2 f(x) dx \quad (5.42)$$

then we obtain

$$\rho_s(c) = \frac{m}{4\pi r_s^3 h(c)} \equiv \frac{c^3 \Delta_h}{3h(c)} \bar{\rho} \quad (5.43)$$

Thus for all profiles,  $h(c)$  uniquely defines the amplitude.

The calculation of two-point statistics involves two more functions derived from the halo density profile: the normalised Fourier transform  $u(k|m, z)$  and the self-convolution  $\lambda(r|m, z)$ . We now proceed to set up the general case of these.

The normalised Fourier transform of the halo profile is defined as (Cooray and Sheth, 2002)

$$u(\mathbf{k}|m, z) = \frac{1}{m} \int \rho(\mathbf{r}|m, z) e^{-i\mathbf{r}\cdot\mathbf{k}} d^3\mathbf{r} \quad (5.44)$$

in which  $\mathbf{k}, \mathbf{r}$  are the Cartesian wave-vector and position vector respectively. In the case of a spherically symmetric distribution truncated at the halo radius (as we assume) this may be treated by a Fourier-Bessel (or Hankel) transform, in which we use  $\kappa = k r_s$ , and  $x$  as previously defined:

$$u(\kappa|m, z) = \frac{4\pi r_s^3}{m} \rho_s(c) \int_0^c x^2 \frac{\sin(\kappa x)}{\kappa x} f(x) dx. \quad (5.45)$$

Using Eq. 5.43 we have

$$u(\kappa|m, z) = \frac{p(\kappa, c)}{h(c)} \quad (5.46)$$

where

$$p(\kappa, c) = \int_0^c x^2 \frac{\sin(\kappa x)}{\kappa x} f(x) dx \quad (5.47)$$

is the reduced form.

The self-convolution of the density profile is

$$\lambda(x|m, z) = \rho_s^2(c) r_s^3 l(x, c) \quad (5.48)$$

and has units  $M^2/V$ . The reduced form is given by

$$l(x, c) = \int f(\mathbf{y}, c) f(|\mathbf{y} + \mathbf{x}|, c) d^3y, \quad (5.49)$$

in which  $x$  and  $y$  are in units of  $r_s$  as usual. It is generally difficult, if not impossible, to solve this integral analytically. However, there are several profiles for which do afford a solution, including the popular profile of Navarro, Frenk, and White (1997, hereafter NFW).

Thus, any profile is fully defined by its reduced profile shape,  $f(x)$ , which can be used to produce the dimensionless quantities  $h(c)$ ,  $p(\kappa, c)$ , and  $l(x, c)$ . We utilise this system in the `halomod` implementation.

### 5.2.9 Concentration-Mass Relation

The concentration of halos, defined as  $c = r_{\Delta_h}/r_s$ , where  $r_s$  is the radius at which the logarithmic slope is -2, is known to be correlated with halo mass (Navarro, Frenk, and White, 1997). Early studies developed toy models to explain this correlation, based on associating the concentration in the central regions with the density of the Universe at the epoch at which the halo collapsed,  $a_c = 1/(1+z_c)$ .

Initially, NFW defined the epoch of collapse via the spherical-collapse model of Press and Schechter (1974) after Lacey and Cole (1993), as the epoch most likely to be the first time that a certain fraction  $f$  of a halo of mass  $m$  at epoch  $a$  had crossed the density threshold for collapse:

$$\text{erfc} \left( \frac{\delta_c(a_c) - \delta_c(a)}{\sqrt{2 [\sigma_0^2(fm) - \sigma_0^2(m)]}} \right) = 1/2. \quad (5.50)$$

They then related the inner density of the halo,  $\rho_s$ , to the density of the Universe by a constant of proportionality to yield

$$\rho_s = k \bar{\rho}(a) \left( \frac{a_c}{a} \right)^3, \quad (5.51)$$

which determines the concentration via Eq. 5.43.



This model tends to over-estimate the concentration at high redshift, which motivated Bullock (2001, hereafter Bo1) to refine the toy model. They specify the collapse epoch as the epoch at which a typical collapsing mass was some fraction  $F$  of the halo mass:

$$M_{\star}(a_c) = Fm, \quad (5.52)$$

where

$$\nu(M_{\star}, a) = 1. \quad (5.53)$$

Defining a generalised inner density

$$\tilde{\rho}_s \equiv \frac{3m_{\text{vir}}}{4\pi r_s^3}, \quad (5.54)$$

they again set the inner density of the halo proportional to the density of the Universe:

$$\tilde{\rho}_s = K^3 \Delta_{\text{vir}}(a) \bar{\rho}(a) \left( \frac{a}{a_c} \right)^3, \quad (5.55)$$

to yield the simple form

$$c(m, a) = K \frac{a}{a_c}. \quad (5.56)$$

In these equations, for a given cosmology,  $F$  and  $K$  are assumed to be constant for all halos. Thus, for a given halo mass, the concentration is proportional to  $a$ , and so is increasing with time. Furthermore, at a given epoch, lower masses have younger collapse epochs, and are therefore more highly concentrated.

In fact, for the simplest possible cosmological model, an Einstein-de Sitter universe with scale-free power spectrum  $P(k) = k^n$ , the resulting mass dependence is a simple power law. Indeed, for the more complex  $\Lambda$ CDM model, the local behaviour of the model is also power-law-like, which motivated Bo1 to give the following simple form from simulation results:

$$c(m_{\text{vir}}, z) \simeq \frac{9}{1+z} \left( \frac{m}{M_{\star}} \right)^{-0.13}. \quad (5.57)$$

Several significant refinements have been made to the simple analytic models, especially by including results from mass-accretion-histories (MAHs, Wechsler et al., 2002; Zhao et al., 2003; Zhao et al., 2009; Correa et al., 2015a; Correa et al., 2015b; Correa et al., 2015c), using both simulated results and EPS predictions. However, the most common prescription used in observational and halo model studies is a generalisation of the approximation of Bo1, taking the form of a simple power-law in mass and scale-factor (Dolag et al., 2004; Kuhlen et al., 2005; Buote et al., 2007; Comerford and Natarajan, 2007; Macciò et al., 2007; Neto et al., 2007; Duffy et al., 2008; Macciò, Dutton, and Bosch, 2008)

$$c(m_{\Delta_h}, a) = A \left( \frac{m}{M'} \right)^B a^C, \quad (5.58)$$

where from the seminal work by Bullock (2001),  $A \sim 9$ ,  $B \sim -0.13$  and  $C \sim 1$ . The precise values of these parameters change with cosmology, the halo definition, and the choice of pivot mass  $M'$ . It is common to choose  $M' = M_{\star,0}$ , but this is arbitrary in the power-law model.

This property remains an active area of research, requiring high-resolution and large-scale simulations to properly model universality over a broad range of epochs and masses.

### 5.2.10 Halo exclusion

Modelling halo exclusion in the 2-halo term attempts to account for the fact that pairing two galaxies from separate halos is extremely unlikely if the radius of one of the halos is larger than the distance between galaxies. In this case, it is far more probable that the galaxies will reside in the same halo. In practice, this affects the correlation function at the crossover between the 1-halo and 2-halo terms at the 30% level (Schneider et al., 2012), by suppressing the 2-halo term.

#### 5.2.10.1 Empirical approaches

There have been a number of proposed models to account for this effect. At a purely empirical level, one may consider modifying  $P_{hh}^c(k)$ . While it is customary to use a (biased) nonlinear estimate of the matter power, Cooray and Sheth (2002) suggest that using the linear matter power will decrease the power at small scales, therefore increasing the fidelity of the model. In a similar way, Schneider, Smith, and Reed (2013) use a smoothing of the power on roughly transition scales to eliminate the excess correlations:

$$P_{hh}^c(k, m_1, m_2) \approx b(m_1)b(m_2)W(kR_T)P_m(k), \quad (5.59)$$

where  $P_m(k)$  is the matter power from HALOFIT,  $W$  is the Fourier transform of a top-hat in real space,

$$W(x) = 3 \frac{\sin x - x \cos x}{x^3}, \quad (5.60)$$

and  $R_T \approx 2\text{Mpc}h^{-1}$ . This latter argument can be used to reduce errors to below 10%, though its physical motivation is unclear.

#### 5.2.10.2 Physical approaches

By far the more common approach is to set upper limits (say  $m'_1$  and  $m'_2$ ) on the 2-halo integral, Eq. 5.9, using physical arguments. This approach has been described in some detail in Tinker and Weinberg 2005 (hereafter T05), and we briefly reproduce that discussion here. While we follow T05 and discuss the galaxy quantities, precisely the same arguments hold for DM quantities with the replacement  $I_g \rightarrow I_{\text{DM}}$  (and  $\bar{n}_g \rightarrow \bar{\rho}$ ).

The underlying idea is to set  $m'_1$  and  $m'_2$  such that the scale in question is not smaller than the sum of the halo radii, i.e.  $r \geq R_{\text{vir}}(m'_1) + R_{\text{vir}}(m'_2)$ .

Note that in this approach, the density is also modified, so that

$$\bar{n}_g'^2 = \int^{m_1'} \int^{m_2'} \prod_{i=1}^2 n(m_i) N_t(m_i) dm_i. \quad (5.61)$$

In this case, the final 2-halo correlation function needs to be re-scaled by

$$\xi_g^{2h}(r) = \left( \frac{\bar{n}_g'}{\bar{n}_g} \right)^2 \left[ 1 + \xi_g'^{2h}(r) \right] - 1. \quad (5.62)$$

The simplest model is to assume spherical halos, and that the two halos in question are the same size, so that (Zheng et al., 2005):

$$m_1' = m_2' = \frac{4}{3} \pi \left( \frac{r}{2} \right)^3 \Delta_h \bar{\rho}. \quad (5.63)$$

This is efficient since the limits are independent, and we can use the separated form of Eq. 5.10. However, it severely under-counts pairs due to edge effects.

We can do better by relaxing the assumption that the two halos are the same size. In this case,  $m_1'$  and  $m_2'$  are related by

$$R_\Delta(m_2') = r - R_\Delta(m_1). \quad (5.64)$$

In this case, we cannot separate the integrals, and so double-integration must be performed, which is relatively inefficient.

It is well-known that halos are typically triaxial (Bullock et al., 2001; Taylor, 2011; Zemp et al., 2011), and therefore one can improve the results by considering the probability that each pair is in different halos, given an empirical distribution of axis ratios. In this case, Eq. 5.9 is augmented by the distribution:

$$\frac{P_g^{2h}(k, r)}{P_m(k, r)} = \int \int \prod_{i=1}^2 I_g(k, m_i) b(m_i, r) \mathcal{P}(x) dm_i. \quad (5.65)$$

Here,  $\mathcal{P}(x)$  is the probability that pairs at a given separation are in separate halos, where  $x = r / (R_\Delta(m_1) + R_\delta(m_2))$ . To5 suggest using

$$\mathcal{P}(y) = 3y^2 - 2y^3 \quad (5.66)$$

$$y = \frac{x - 0.8}{0.29} \quad (5.67)$$

where  $\mathcal{P}(y < 0) = 0$  and  $\mathcal{P}(y > 1) = 1$ . This is a significant improvement over the spherical case, however it still requires double-integration (at every value of  $k$ !) due to the appearance of  $m_1$  and  $m_2$  in  $\mathcal{P}(x)$ .

To5 propose a solution to the efficiency problem by performing one double-integral to calculate  $\bar{n}_g'$  for the full ellipsoidal case (where  $\mathcal{P}(x)$  also augments Eq. 5.61). From

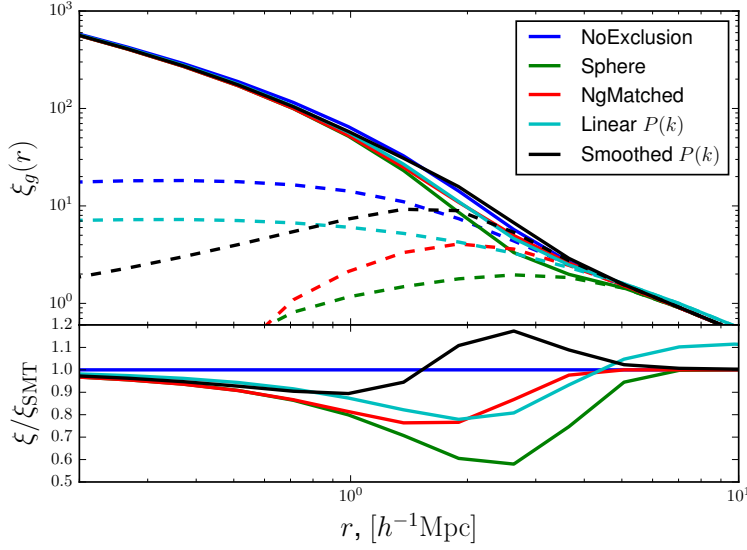


Figure 5.6: Effects of various halo exclusion models around the transition region. Solid lines are the full 2PCF, while the dashed lines show the 2-halo term. In both cases where the halo-centre power spectrum is modified (cyan and black curves), the halo exclusion is not performed.

here, one can “match” the number density numerically with the simpler single-integral:

$$\bar{n}'_g = \int_0^{m'} n(m) N_t(m) dm, \quad (5.68)$$

cumulatively integrating until the integral matches  $\bar{n}'_g$ , thus acquiring  $m'$ . Setting  $m'_1 = m'_2 = m'$ , one can use the separated form, Eq. 5.10, to generate the final galaxy power spectrum.

We show the effect of several of these models in figure 5.6. The spherical exclusion is the most effective at reducing pairs, with the  $n_g$ -matched method exhibiting a more gentle suppression. Interestingly, merely using the linear power without any exclusion is the closest to the  $n_g$ -matched case at the transition scales.

### 5.3 THE HALOMOD LIBRARY

In this section we present our new implementation of the HM, HALOMOD. Our intention is first to give an overview of the philosophy behind, and general characteristics and features of the code. From there, we describe in detail various parts of the code, beginning in §5.3.2 with the “frameworks” which tie together the various components, and then the individual components themselves in §5.3.3. These sections have a pragmatic focus, introducing any pertinent numerical techniques, and summarising the relevant APIs and various models included. Finally, we comment on some of the extra functionality present in the package in §5.3.4.

### 5.3.1 Overview

HALOMOD is a Python package, built on the HMF package<sup>9</sup> (Murray, Power, and Robotham, 2013a), which contains the necessary components and algorithms to calculate HM quantities in the context of the HOD<sup>10</sup>.

As discussed in the introduction, HM studies are a burgeoning field, especially in the interpretation of large galaxy surveys via HOD fits to galaxy two-point statistics. Furthermore, the compartmentalisation of the technique into various components (eg. halo profile, mass function etc.) provides opportunity to enhance analyses by improving individual components. These basic factors motivate the development of a code which has the following properties:

- **Simple.** Though many aspects of the calculations are unavoidably non-trivial, a simple layout of the code within a highly structured framework, with a well-specified and intuitive API for the user is important. This is always desirable, of course, but in this case it promotes future development, and usage by a broad cross-section of researchers.
- **Flexible/Extendible.** The HM is a rapidly evolving framework, with individual components constantly improving, and the framework itself being extended. Building a static implementation is therefore non-conducive to the development of the field. Components need to be as plug-and-play as possible, with new models easily created and inserted on the fly.
- **Comprehensive.** The code should act like a repository for all the modelling that has been done. It should collate and compile the various models and extensions in a cohesive way so that new models can be quickly compared, and insights gained.
- **Efficient.** Though not as immediately important as flexibility, it is important that the code be efficient in the coming era of enormous volumes of data. This includes both algorithmic and numerical efficiency, but also efficiency of the writing of user-side code.
- **Open.** It must be open-source (and it is). In addition, it must be well-documented both within the code and externally.

These have been the guiding principles in the development of HALOMOD (and also HMF). Our code provides a well-defined architecture for defining components and frameworks for halo modelling, including many basic and derived quantities of interest, and a large fraction of the available models in the literature. Each component is completely flexible, either by supplying custom parameters or entirely new parameterisations *without touching the source code*. We now expand upon some of these aspects.

Note that many of the features that we discuss in this section reflect updates to the HMF package (compared to its presentation in Murray, Power, and Robotham (2013a)), which are inherited in HALOMOD. Furthermore, clearly this chapter represents a snapshot of the package in time, with updates likely to occur.

<sup>9</sup> <https://github.com/steven-murray/hmf>

<sup>10</sup> For now, it only implements the two-point statistics, but this will be expanded in future versions.

### 5.3.1.1 General usage

The backbone of HALOMOD is split into two categories: ‘frameworks’ and ‘components’, which are well-defined by a base class, Framework and Component respectively. Individual elements of the calculation, such as a halo profile, HMF fitting function, or any other unit which may be calculated in more than one way (eg. growth factors) are represented by a Component. Frameworks are the structures which tie each Component together to calculate a set of quantities (eg. correlation functions, power spectra or mass functions).

This structure makes it simple to identify a component of interest for extended modelling, and provides a common interface and set of methods to each type. For example, if one is developing a new improved model of halo bias, it is clearly identified as a Component, with a particular structure. Its place in the overall framework is well-specified, it by default receives any defining parameters it requires, and specific examples of implementations of other models of the component are easily found in the bias module.

It is expected that the most common usage will be to create a Framework instance, and access available quantities through it. Our general approach has been to treat frameworks as a self-contained unit. This means that all parameters are passed to the initial constructor, after which every relevant quantity is ready for immediate access, without method-specific parameters. Indeed, to the user, once a Framework instance has been created (with all its options), any derived quantity (such as mass functions, correlation functions, mass variance etc.) is accessed just like a stored variable<sup>11</sup>.

This approach is not typical of much software – usually there is benefit in segregating object-wide variables from method-specific ones. However, we find in our case that most variables are used across many methods, and are more naturally interpreted as being a part of the framework itself.

In fact, our adherence to this approach is bolstered by our generic distinction (within Frameworks) between a parameter and a property. Every input variable is treated as a class-wide parameter, which receives its own characteristics (this is very useful for our caching mechanism, see §5.3.1.2). Every evaluated quantity is treated as a property, which likewise has its own properties. The set of all parameters of a given Framework is then a unique description of the entire object.

Though the state of a Framework is defined by its parameters, we provide a simple mechanism of modifying the parameters for dynamic calculations, through an update method, which can receive any parameter<sup>12</sup>.

Our approach to the passing of parameters has been to provide defaults for *every* variable in a Framework<sup>13</sup>. These defaults are chosen to be broadly in line with the most common usage, which greatly simplifies common tasks. However, it must be noted that this has the drawback that if user code relies on defaults, and code updates modify them, then results can vary without being noticed.

In summary, a brief example may be most illustrative. Generating a two-point correlation function with reasonable default parameters is simply three lines of code:

<sup>11</sup> In reality, the quantities are ‘lazy’, and are evaluated on access

<sup>12</sup> This method is defined in the base Framework class, and is therefore simply inherited by any specific framework.

<sup>13</sup> These defaults are currently majoritarilly static, but we plan on dynamifying them in future versions.

```

from halomod import HaloModel
h = HaloModel()
xir = h.corr_gg

```

### 5.3.1.2 Efficiency

Efficiency, while important, is not the primary aim of `halomod`. As a broad approach, we have aimed for flexibility and ease-of-use, along with a robust design, over efficiency. This is made clear from the outset by the choice of programming language, as Python is known for being relatively slow in several regimes. We believe that flexibility and robust design will be of much greater benefit for the community and the future of modelling with the HM than raw efficiency. Despite this, with the era of ‘big data’ on the horizon it is important to optimize wherever possible. We have done this in two ways: at a functional level, and an algorithmic level.

We make heavy use of the excellent `NUMPY`<sup>14</sup> package to enable optimized compiled code performance in the background for many of the array calculations. With careful use of broadcasting and vectorization, the bulk of the calculations are close to as efficient as one would expect from compiled code, within the simplicity and readability of the interpreted language.

In cases of computational bottlenecks where vectorization is not an option due to memory consumption, we have implemented optional “just-in-time” (JIT) compiled functions using the new `NUMBA` package<sup>15</sup>. Though these functions reduce code readability and extendibility to some extent, they can improve speeds in small bottleneck routines to almost-compiled-code standards. Furthermore, we always implement a pure-python basic function for testing, and the JIT-compiled function is only called in its place if the user has `NUMBA` installed.

We make various other functional-level performance improvements which we will refer to in their own context. However, possibly the most useful performance enhancement is `HALOMOD`’s caching system.

The `HMF` package has always included some form of caching, whereby derived quantities are stored for later retrieval, and only updated when absolutely necessary. At the time of publication, `HMF` relied on each parameter being hard-coded into an update method in a specific way so as to trigger re-calculation of all child properties. This required great care when modifying the code, and especially when adding new parameters, as breaking the chain of dependencies could cause subtle errors.

We have since updated the caching system of `HMF` to be much more robust. We use Python’s function decorators, a kind of functional programming technology, to abstract the caching logic from the main Framework. There is now a specific module containing two methods and a single base class which contain all the machinery for the cache system. This increases the readability of the main code, and maximises the usage of the DRY (Don’t Repeat Yourself) principle.

The caching system makes heavy use of the general distinction between a parameter and property. All parameters receive a `@parameter` decorator, which uses function clo-

<sup>14</sup> <http://www.numpy.org/>

<sup>15</sup> <http://numba.pydata.org/>

sure to modify its behaviour. Likewise, all properties receive a `@cached_property` decorator, with optional arguments specifying the names of all direct parents of the property (whether other properties or parameters). This is simple to deduce from the code, since the definition of the property contains these parents.

The `@cached_property` decorator performs several operations each time its function is called. On the first call, the appropriate value is calculated, and then *before returning*, a private dictionary is populated with a cross-reference to each parent, and their relationship to the children. Because this happens after the calculation, but before the return, the first method that actually finishes this cross-referencing by definition can only have parameters as parents. Thus all child properties can access that information and link back to their ultimate parent parameters. Finally, a table is built which contains, for each parameter, a list of the properties it affects. The calculated value is stored in a private variable, and each subsequent call merely returns this value, unless the property is flagged as needing recalculation.

The `@parameter` decorator merely ensures that each time a parameter is updated, the list is traversed, and each child property is flagged as needing recalculation.

This method is as efficient as possible for such a generalised process. No re-calculations are performed unless necessary. However, higher efficiency could be achieved by simply using low-level functions chained together in an appropriate way, storing only the variables that needed storing for a particular application. This is due to the extra memory usage required to store all variables in the caching system, and the minimal overhead of re-checking the cache.

Importantly, this process is completely transparent to the user, who merely accesses properties as if they were pre-stored variables.

### 5.3.1.3 Flexibility

We have referred to our imposed dichotomy between Framework and Component objects, and in the previous subsection we outlined a performance benefit derived from the internal structure of Frameworks. Likewise, we have implemented a uniform convention for Component specification which enhances flexibility.

Any component that is implemented as a Component is a variable within a Framework. As such, it is determined by passing a parameter, in the form of either a string representing the name of the model, or the class itself. In order to avoid updating endless lists of available models to be selected by the string identifier, we implement a basic `get_model` function which locates and returns appropriate models based on our conventions.

Every Component has a top-level dictionary called `_defaults`, which contains the model-specific parameters and their default values (these are not fundamental quantities used in the calculation, eg.  $\nu$  in the  $f(\nu)$  for the HMF, but rather free parameters of the model, if any). These are processed by the base class so that any newly defined model automatically has them available. In the Framework, the model parameters are passed, by convention, as `<component>_params`. Furthermore, the update method in Frameworks in-



telligently updates parameter dictionaries so that previous updates are persistent unless explicitly overwritten<sup>16</sup>.

This general system offers the highest level of flexibility and extendibility. In a single session, it is extremely simple to switch between several models for a given component, thereby making useful comparisons (and this without needing to recalculate basic quantities that do not depend on the variable component). On the other hand, using entirely new custom models can be achieved without modifying the source code, by inheriting from the base Component and passing the new model directly. We expect this to be particularly advantageous in the current climate of rapid development.

Furthermore, to a greater or lesser degree, each Component has built-in methods which either remove the burden on the user, or supply extra functionality for free when the user develops a new model. This is perhaps most clear in the Profiles component (cf. §5.3.3.6), though some components have very little to add to the basic model (eg. the Bias component, cf. §5.3.3.5).

Components are primarily designed to integrate with the Frameworks, however, they are not limited to this usage. They are also intended to be useful in their own right, for more specific applications.

Summarily, the Component architecture of HALOMOD is geared to provide flexibility and extendibility for crucial developments in the field in the years to come.

### 5.3.2 Frameworks

In this section we present two of the frameworks that HALOMOD provides. We note that several other frameworks are found in the HMF package, from which those presented here inherit.

#### 5.3.2.1 Base Halo Model

The basic framework in HALOMOD is `HaloModel`, which provides general access to most derived quantities relating to two-point statistics. Table 5.1 contains a useful summary of parameters for `HaloModel`, giving an indication of the flexibility involved. Likewise, table 5.2 lists the available properties. We note that these are expected to be updated and expanded in future versions, and thus are not intended to provide a strict API, but rather an illustrative summary.

Some of the features of the `HaloModel` framework, ignoring contributions of individual components, are as follows.

**MEAN DENSITY MATCHING.** Most HOD parameterisations have a parameter  $M_{\min}$ , which either explicitly or loosely defines the minimum halo mass expected to host galaxies in the sample. This parameter is tightly correlated with the mean galaxy density, as it determines the lower limit of the integral in Eq. 5.19. It is common (though not ubiquitous) in survey analyses to let the known value of  $\bar{n}_g$  determine  $M_{\min}$  (given other parameters of the HOD). `HaloModel` supports this technique for all HOD models, either

<sup>16</sup> Note that this behaviour, while useful in general, can cause some headaches if one needs to *delete* model parameters

PARAMETER	DEFAULT	DESCRIPTION
<b>COMPONENTS/PARAMETERS*</b>		
cosmo_<>	Planck13	Underlying cosmology. Must be an FLRW object from <code>ASTROPY</code> . This model differs from other Components, being passed as an <i>instance</i> , not a <i>class</i> .
transfer_<>	CAMB	A model to use for the transfer function.
growth_<>	GrowthFactor	A model for the growth function. The default is to perform the explicit integration.
hmf_<>	Tinker08	A model specifying the HMF fitting function, $\sqrt{v}f(v)$ .
filter_<>	TopHat	A filter function model.
hod_<>	Zehavi05	A HOD model.
profile_<>	NFW	A halo profile.
concentration_<>	Duffy08	A concentration-mass-redshift relation
bias_<>	Tinker10	A scale-independent linear bias model.
sd_bias_<>	Tinker_SD05	A model for the scale-dependence of the linear bias.
exclusion_<>	NgMatched	A halo exclusion model.
<b>RESOLUTION/LOCATION</b>		
lnk_min, lnk_max, dlnk	-8,8, 0.05	Min/Max/ $\Delta$ logarithmic wavenumber
dlog10m	0.01	Logarithmic mass interval
rmin, rmax, rnum	0.1, 50.0, 20	Min/Max/# for pair-separation vector.
<b>PHYSICAL</b>		
sigma_8	0.8344	Normalisation of power spectrum, $\sigma_8$
n	0.9624	Spectral index, $n_s$
z	0.0	Redshift
delta_h, delta_wrt	200, "Mean"	The halo overdensity, $\Delta_h$ , and what it is with respect to (mean or critical).
delta_c	1.686	Critical overdensity for collapse, $\delta_c$ .
ng	None	Optional specification of mean galaxy density. If present, an HOD parameter (generally $M_{\min}$ ) is fixed.
<b>OPTIONS</b>		
takahashi	True	Apply updated HALOFIT coefficients from Takahashi et al. (2012)
hc_spectrum	Nonlinear	Model for $P_{hh}^c(k)$ . Either "linear", "nonlinear" (HALOFIT), or "filtered-nl" (smoothed on a scale of $2h^{-1}\text{Mpc}$ ).

Table 5.1: All parameters of `Halomod`. Parameters in the "Components" section are each specified by two parameters, ending in `model` and `params`. The first is the actual model specification as mentioned in the table, while the second is a dictionary of parameters for the model.

TYPE	QUANTITIES*
Scalar	$b_{\text{eff}}, f_{\text{cen}}, f_{\text{sat}}, D^+(z), M_{\text{eff}}, M_*, \bar{\rho}(z), \bar{n}_g, n_{\text{eff}}$
Length- $m$	$R, b, \bar{c}, \sqrt{v}f(v), \frac{dn}{d \ln m}, \frac{dn}{d \log_{10} m}, \frac{dn}{dm}, n(> m), N_c, N_s, N_t, L(n=1), \ln \sigma^{-1}, v, \rho(> m), \rho(< m)$
Length- $k$	$T, P_m, P_{m,\text{halofit}}, \Delta^2, \Delta_{\text{halofit}}^2, P_{gg}^{ss}, P_{gg}^{2h}, P_{mm}, P_{mm}^{1h}, P_{mm}^{2h}$
Length- $r$	$\xi_{gg}^{1h}, \xi_{gg}^{cs}, \xi_{gg}^{ss}, \xi_{gg}^{2h}, \xi_{gg}, \xi_{DM}^{1h}, \xi_{DM}^{2h}, \xi_{DM}, \xi_m, \xi_{m,\text{halofit}}, S(\xi_m(r))$

Table 5.2: All properties of `HaloModel`. Listed are those quantities that are directly accessible as explicit properties. Other quantities, such as the halo profile, its Fourier transform and self-convolution, and the window function, are accessible indirectly through a model instance variable.

by direct cumulative integration, or by numerical minimization. This is caught at every update of parameters.

**MASS RANGE SETTING.** The mass range in `HaloModel` is not user-definable, but rather the class sets intelligent limits. A hard upper limit of  $M_{\text{max}} = 10^{18} h^{-1} M_{\odot}$  is set to ensure convergence, and the lower limit is based on the value of  $M_{\text{min}}$  in the HOD. For step-function models, the lower limit is exactly  $M_{\text{min}}$ , which increases accuracy in an important regime. For smooth-cutoff models, the lower limit is determined by the `mmin` property of the HOD (see §5.3.3.8).

**HALO-CENTRE POWER SPECTRUM.** A key quantity in the large-scale clustering is the power spectrum of halo centres,  $P_{hh}^c(k)$  (cf. §5.2.1.2 and §5.2.10). This can be modelled via the linear matter power, nonlinear matter power, or even some smoothed version of either. We leave this choice free for the user, through the parameter `_hc_spectrum`.

**CONVOLUTIONS.** The satellite-satellite term can be calculated directly in real-space if and only if there is an analytic solution for the self-convolution of the halo profile. `HaloModel` uses this form directly if possible, saving integration time, but resorts to a calculation in Fourier-space (and subsequent Hankel transform) if necessary.

**HANKEL TRANSFORMS.** The Hankel transform is a delicate operation, involving the integration of a highly oscillatory function. We use a novel method for performing this transform, which we detail in appendix B.1.

### 5.3.2.2 Projected correlation function

In analyses of galaxy surveys, it is more common to measure *projected*, rather than real-space, correlation functions (another alternative is the *angular* CF, which we outline in the next section). We provide an extended framework, inheriting from `HaloModel`, for this calculation. The primary reason it is separated from the basic class is that it requires extra parameters specific to its calculation. A summary of the extra parameters and properties in the `ProjectedCF` class is given in table 5.3.

---

<b>PARAMETERS</b>	
rp_min, rp_max, rp_num	Min/Max/Length of projected scale vector
rp_log	Whether to use log-scales for the projected scale vector
proj_limit	Optional upper limit on integral (Eq. 5.69)

---

<b>REMOVED PARAMETERS</b>	
r_min	Automatically set to rp_min
r_max	Automatically set to rlim

---

<b>PROPERTIES</b>	
Scalar	$r_{\text{lim}} \equiv \max(80.0, 5r_{p,\text{max}})$ , or alternatively set with proj_limit
Length- $r_p$	$w_p$

---

Table 5.3: Extra parameters and properties introduced in ProjectedCF, as compared to HaloModel. Some parameters are also automatically overwritten.

The primary addition to the base HaloModel class is the conversion of the real-space 2PCF to projected space. The transformation from real-space to projected is defined as

$$w_p(r_p) = 2 \int_{r_p}^{\infty} \frac{r \xi(r)}{\sqrt{r^2 - r_p^2}} dr. \quad (5.69)$$

The implementation of this integral is rather delicate due to the singularity at the lower bound. This renders both limits of the integration non-physical, and convergence requirements must be met by both. We discuss our prescription for these limits in appendix B.2.

### 5.3.2.3 Warm dark matter models

For each of the frameworks, we also implement a version suited for warm dark matter (WDM) models. We give a fuller description of the various models that have been modified for WDM in Chapter 6, but here we mention that the WDM frameworks are designed to set relevant default models and perform any framework-level WDM-specific modifications.

### 5.3.3 Components

This section provides a brief summary of each of the components within HALOMOD. For each, we provide a summary table, which specifies the variables and quantities available within the model, any necessary definitions, and any extra quantities provided for free. Furthermore, we provide a table or otherwise concise summary of the models already implemented within HALOMOD, and a comparison plot if relevant.

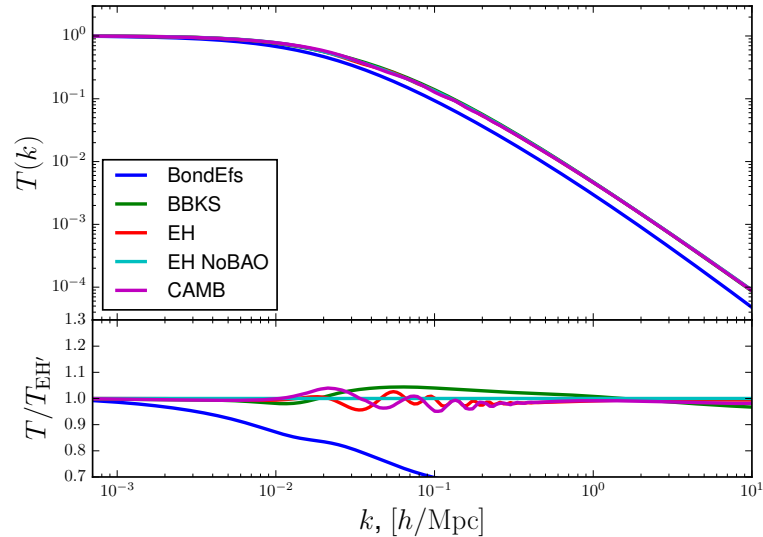


Figure 5.7: All transfer function models in HALOMOD. In each case, the transfer function is normalised to unity at very large scales. The general shape is reproduced in each case, with details of the BAO differing. The oldest fit, of Bond and Efstathiou (1984), fails to reproduce the break scale, though this is likely due to being fit to a differing cosmology.

#### 5.3.3.1 Cosmology

Instead of defining our own basic cosmology models, we use the mature `ASTROPY` package<sup>17</sup> (Robitaille et al., 2013). In brief, this package implements a connected series of cosmographic models based on the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, with several standard instances based on results from CMB missions.

We do not present the summary details here, but note that basic cosmographic quantities such as comoving distance, and density parameters such as the mean density as a function of redshift, are implemented in these models.

#### 5.3.3.2 Transfer functions

We provide a common interface for calculating transfer functions using a variety of methods: from analytic fits, to Boltzmann code, to importing data from file. In each case, the required function returns the logarithmic transfer as a function of logarithmic wavenumber.

We use the `pycamb`<sup>18</sup> wrapper to provide on-the-fly access to `CAMB` routines. We also note that we include a standalone version of `HALOFIT` that may be used to apply nonlinear corrections to any of the included power spectra.

We summarise the API of the Transfer Component in table 5.4, and the various models included in HALOMOD in table 5.5. A comparison of the models with default parameters is shown in figure 5.7.

<sup>17</sup> <http://www.astropy.org>

<sup>18</sup> <https://github.com/joezuntz/pycamb>. Note that `pycamb` is currently unmaintained and does not support the most recent versions of `CAMB`. We are working on providing more robust wrappers in the future.

Available	cosmo	A full instance of the <code>ASTROPY FLRW</code> class.
Required	<code>lnt(lnk)</code>	A function returning the natural log of the unnormalised transfer function, given the natural log of the wavenumber
Extra	None	

Table 5.4: API summary of the Transfer component.

REF	NAME	FORMULA	PARAMS
Bond and Efstathiou (1984)	BondEfs	$\left[1 + \left(\sum_{i=2}^4 (ck)^i\right)^\nu\right]^{-1/\nu}$	$\frac{\Omega_m h^2}{0.3 \times 0.7^2} c = (37.1, 21.1, 10.8), \nu = 1.12$
Bardeen et al. (1986)	BBKS	$\frac{\ln(1 + aq)}{aq} \left[\sum_{i=0}^4 (c_i q)^i\right]^{-1/4}$	$q = \frac{\Gamma}{k \exp(\Omega_b(1 + \sqrt{2h}/\Omega_m))},$ $\Gamma = \Omega_m h, a = 2.34, c = (1, 3.89, 16.1, 5.47, 6.71)$
Eisenstein and Hu (1999) (No BAO)	EH_NoBAO	Original paper, Eq. [26] and [28]-[31]	
Eisenstein and Hu (1999) (w/ BAO)	EH_BAO	Original paper, §2 and Eq. [16]-[24]	
Lewis, Challinor, and Lasenby (2000)	CAMB	Boltzmann Code	Many options
	FromFile	Read from file	Filename

Table 5.5: Summary of included Transfer models.

### 5.3.3.3 Window functions/Filters

The window/filter function is not generally considered a ‘component’ of the HM, however it is an integral part of the calculation, as it contributes to the basic unit of the EPS formalism, the peak-height. The filtering function is also responsible for the conversion of distance scales to mass scales.

Our general approach in this component is to specify all the quantities in terms of the Fourier co-ordinate,  $k$ , and the smoothing scale  $R$ . The class itself defines the relationship between mass  $m$  and the smoothing scale. The primary quantity of interest is the window function itself, specified in Fourier space,  $W(kR)$  (one can optionally specify the real-space version as well, though it is not involved in typical HM calculations). With this function, the mass variance can be calculated using Eq. 5.30.

To calculate the HMF requires also the logarithmic derivative of the variance with mass. To remain as general as possible, we use the following identity

$$\frac{d \ln \sigma}{d \ln m} \equiv \frac{1}{2} \frac{d \ln \sigma^2}{d \ln R} \frac{d \ln R}{d \ln m}, \quad (5.70)$$

where

$$\frac{d \ln \sigma^2}{d \ln R} = \frac{1}{\pi^2 \sigma^2} \int W(kR) \frac{dW(kR)}{d \ln(kR)} P(k) k^2 dk. \quad (5.71)$$

The factor  $d \ln R / d \ln m$  is typically  $1/3$ , though this is not necessarily the case for window functions of arbitrary shape. Thus, the defining quantity is the derivative of the window function,  $dW(x)/dx$ .

Table 5.6 summarises our approach to this important quantity. We also note that we generalise the calculation of the mass variance to all moments of the smoothed density field:

$$\sigma_n^2(r) = \frac{1}{2\pi^2} \int k^{2n} P(k) W^2(kR) dk, \quad (5.72)$$

for which  $\sigma_0^2$  is the usual mass variance.

Table 5.7 summarises the three most common window functions which we include in HALOMOD, while figure 5.8 displays the value of  $\sigma(m)$  for each. Note that there is a complex interplay between the role of  $W(kr)$  and the mass assignment which drives changes between the models.

### 5.3.3.4 HMF

The diversity among fitting functions for the HMF, not only in functional forms, but in general approach and explicit parameter dependence, requires that quite a number of variables be available in the general case. However, the typical fit uses just  $\nu$  (or sometimes  $\sigma$ ) and its own model parameters, reflecting the expected universality of the fit.

We summarise the API of the FittingFunction component in table 5.8, and figures 5.3 and 5.4 display a selection of the available fitting functions and their effect on the galaxy 2PCF respectively.

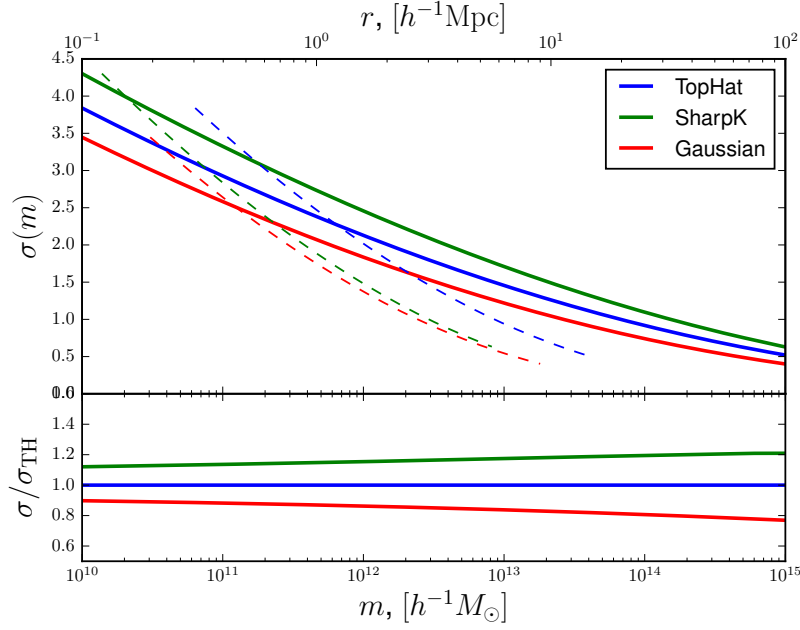


Figure 5.8: The mass variance determined by the three most common window functions. The dashed lines correspond to  $\sigma(r)$ , which is marked on the top  $x$ -axis, whereas the solid lines are  $\sigma(m)$ . The normalisation of the power spectrum in each case uses the TopHat filter, for which  $\sigma_8$  is defined.

<b>Available</b>	$\bar{\rho}, \delta_c, k, P(k)$	
<b>Required</b>	<code>k_space(kr)</code>	Fourier-transform of the filter
	<code>dw_dlnkr(kr)</code>	The derivative of the Fourier-transform of the filter w.r.t $kr$ .
	<code>mass_to_radius(m)</code>	Inverse of mass assignment
<b>Optional</b>	<code>dlnr_dlnm</code>	Dependence of scale on mass. Required if relationship is not default: $1/3$ .
	<code>real_space(R, r)</code>	Real-space window function at $r$ , with scale $R$ . Unnecessary for HM calculations.
	<code>radius_to_mass(r)</code>	Assignment of scale to mass. Unnecessary for HM calculations. Can be supplied for completeness.
<b>Extra</b>	<code>dlnss_dlnr(r)</code>	$\frac{d \ln \sigma^2(r)}{d \ln r}$ , Eq. 5.71
	<code>dlnss_dlnm(r)</code>	$\frac{d \ln \sigma^2(r)}{d \ln m}$ , Eq. 5.70
	<code>sigma(r, order=0)</code>	Moments of the power spectrum, given by order. Default is $\sigma(r)$ .
	<code>nu(r)</code>	Peak-height

Table 5.6: API summary of the Filter component.



NAME	FORMULA
TopHat	$W(r, R) = H(R - r)$
	$W(x = kR) = \frac{3}{x^3} (\sin x - x \cos x)$
	$m(R) = \frac{4\pi}{3} R^3 \bar{\rho}$
	$\frac{dW}{d \ln x} = \frac{9x \cos x + 3(x^2 - 3) \sin x}{x^3}$
SharpK	$W(r, R) = \frac{\sin(r/R) - (r/R) \cos(r/R)}{2\pi^2 r^3}$
	$W(x = kR) = H(kR - 1)$
	$m(R) = \frac{4\pi}{3} [cR]^3 \bar{\rho}$
	$\frac{dW}{d \ln x} = \delta_D(x - 1)$
	$\frac{d \ln \sigma^2(r)}{d \ln r} = -\frac{P(1/r)}{2\pi^2 \sigma^2(r) r^3}$
Gaussian	$W(r, R) = \frac{\exp(-r^2/2R^2)}{(2\pi)^{3/2} R^3}$
	$W(x = kR) = \exp(-x^2/2)$
	$m(R) = (2\pi)^{3/2} R^3 \bar{\rho}$
	$\frac{dW}{d \ln x} = -xW(x)$

Table 5.7: Summary of included Filter models. Note:  $H$  is the Heaviside step-function,  $\delta_D$  is the Dirac-delta function and  $P$  is the power spectrum. Also, in SharpK, the mass-assignment is not well-defined, but we use the given formula with  $c \approx 2.5$  a fit to simulations (Benson et al., 2012; Schneider, Smith, and Reed, 2013).

<b>Available</b>	$m, \nu, \delta_c, \sigma, n_{\text{eff}}, \ln \sigma^{-1}, z, \Delta_h, \text{cosmo}, \Omega_m(z)$	
<b>Required</b>	<code>fsigma(cut_fit)</code>	Returns $f(\sigma) \equiv \sqrt{\nu} f(\nu)$ , and <code>cut_fit</code> specifies whether to truncate in the mass range of the fit.
<b>Extra</b>	None	

Table 5.8: API summary of the FittingFunction component.

<b>Available</b>	$m, \nu, \delta_c, \Delta_h, \Omega_m, n_s, M_*, h, \sigma_8$	Note that most of these parameters are unnecessary for most fits.
<b>Required</b>	<code>bias()</code>	Return the bias corresponding to $\nu$ (or equivalently $m$ ).
<b>Extra</b>	None	

Table 5.9: API summary of the Bias component.

We refer to table 3 in Chapter 2 for a table of fitting functions, noting that four more fits have been added to the collection, specifically those of Tinker et al. (2010), Manera, Sheth, and Scoccimarro (2010), and Pillepich, Porciani, and Hahn (2010) and Ishiyama et al. (2015).

#### 5.3.3.5 Bias

Though the majority of bias functions are specified in ‘universal’ form with respect to the peak-height, there are some older functions that explicitly require cosmology and are specified in terms of a scaled mass,  $m/M_*$ . We thus provide access to these quantities, which may also serve to enable more fine-tuned models than possible with pure peak-background split arguments in the future.

We summarise the API of the Bias component in table 5.9, and compile the range of included models in table 5.10. Figure 5.5 displays a selection of the available models.

Table 5.10: Summary of included Bias models. Note that the fit of Manera10 is dependent on Friends-of-Friends linking length and redshift, and we give the result for  $l = 0.2$  at  $z = 0$ .

REF	NAME	FORMULA	PARAMS	HMF
Mo and White (1996)	Mo96	$1 + \frac{\nu - 1}{\delta_c}$		PS
Jing (1998)	Jing98	$(a/x^2 + 1)^{b-cn_s} \left( 1 + \frac{x-1}{\delta_c} \right)$	$x = (m/M_\star)^2, a = 0.5, b = 0.06, c = 0.02$	
Sheth and Tormen (1999)	ST99	$1 + \frac{q\nu - 1}{\delta_c} + \frac{2p/\delta_c}{1 + (q\nu)^p}$	$q = 0.707, p = 0.3$	SMT
Mandelbaum et al. (2005)	Mandelbaum05		$q = 0.73, p = 0.15$	
Manera, Sheth, and Scoccimarro (2010)	Manera10		$q = 0.709, p = 0.248$	
Sheth, Mo, and Tormen (2001)	SMT01	$1 + \frac{1}{\sqrt{a}\delta_c} \left[ s\sqrt{a}(av) + \sqrt{ab}(av)^{1-c} \right]$	$a = 0.707, b = 0.5, c = 0.6$	SMT
Tinker and Weinberg (2005)	Tinker05	$-\frac{(av)^c}{(av)^c + b(1-c)(1-c/2)}$	$a = 0.707, b = 0.35, c = 0.8$	
Seljak and Warren (2004)	Seljak04	$a + bx^c + \frac{d}{ex + 1} + fx^g$	$x = m/M_\star, a = 0.53, b = 0.39, c = 0.45, d = 0.13, e = 40, f = 5 \times 10^{-4}, g = 1.5$	
Seljak and Warren (2004)	Seljak04_Cosmo	$b_{\text{sel}} + \log_{10} x [a_1(\Omega'_m + n'_s) + a_2(\sigma'_8 + h')]$	$a_1 = 0.4, a_2 = 0.3, \Omega'_m = \Omega_m - 0.3, n'_s = n_s - 1, \sigma'_8 = \sigma_8 - 0.9, h' = h - 0.7$	
Pillepich, Porciani, and Hahn (2010)	Pillepich10	$B_0 + B_1\sqrt{\nu} + B_2\nu$	$B_0 = 0.647, B_1 = -0.320, B_2 = 0.568$	

Table 5.10: continued...

REF	NAME	FORMULA	PARAMS	HMF
Tinker et al. (2010)	Tinker10	$1 - A \frac{v^{a/2}}{v^{a/2} + \delta_c^a} + Bv^{b/2} + Cv^{c/2}$	$A = 1 + 0.24y \exp[-(4/y)^4], a = 0.44y - 0.88,$ $B = 0.183, b = 1.5,$ $C = 0.019 + 0.107y + 0.19 \exp[-(4/y)^4], c = 2.4,$ $y = \log_{10} \Delta_{li}.$	
Tinker et al. (2010)	Tinker10_PBsplit	$1 + \frac{\gamma v - (1 + 2\eta)}{\delta_c} + \frac{2\phi/\delta_c}{1 + (\beta^2 v)\bar{\phi}}$	Same as $f_{\text{Tinker10}}$	Tinker10

### 5.3.3.6 Halo Profiles

We developed a systematic representation of two-parameter universal halo profiles in §5.2.8. Our implementation follows this development closely. There is a richness of predictions available given just the basic profile shape, which makes the class structure of the `Profile` component extremely valuable. We have not yet implemented several generic predictions that may be added in future versions, such as the gravitational potential profile and circular velocity profile.

Table 5.11 outlines the basic API of the `Profile` component, highlighting the many derived quantities accessible with a simple definition of the profile shape. To make this more explicit, suppose that one wished to implement a “cored” NFW, such that the inner density was a constant  $\rho_s$ . The minimum code that the user must write (remember, this is user-side code, not touching the HALOMOD source) would be something like this:

```
from halomod.profiles import Profile
class CoredNFW(Profile):
    def f(self,x):
        return 1./(1 + x*(1+x)**2)
```

The class itself can then be used directly or passed to any relevant `Framework`, and all of the necessary quantities will be available. For increased efficiency, it may also be beneficial to analytically define  $p(\kappa, c)$ ,  $h(c)$  and even  $l(x, c)$  if available. If not specified, these quantities are evaluated via numerical integration (except for  $l(x, c)$  – it is usually more efficient to use an analytic Fourier-transform  $u^2(k|m)$  and then use a Hankel transform on the result).

We define our implemented profiles in table 5.12, and we note that for each, analytic forms have been given wherever possible to improve efficiency. In general it is non-trivial to produce analytic forms for the self-convolution. Fortunately, this can be achieved with the popular NFW profile (Sheth et al., 2001), as we show in table 5.12. The associated forms for the  $T_i$  are

$$T_1 = \frac{-4(1+a) + 2ax(1+2a) + a^2x^2}{2x^2(1+a)^2(2+x)}, \quad (5.73a)$$

$$T_2 = \frac{1}{x^2} \ln \left[ \frac{(1+a-ax)(1+x)}{1+a} \right], \quad (5.73b)$$

$$T_3 = \frac{\ln(1+x)}{x(2+x)^2}, \quad (5.73c)$$

$$T_4 = \frac{\ln[(1+a)/(ax+a-1)]}{x(2+x)^2}, \quad (5.73d)$$

$$T_5 = \frac{a^2x - 2a}{2x(1+a)^2(2+x)}. \quad (5.73e)$$

For the profile of Moore et al. (1998) and the generalised NFW and Moore profiles, the non-truncated self-convolution can be reduced to a single integral (Ma and Fry, 2000),

$$l(x) = \frac{2\pi}{x} \int_0^\infty y f(y) F_\alpha^X(x, y) dy, \quad (5.74)$$

<b>Available</b>	$\bar{\rho}, z, \Delta_h, \text{\_cm\_relation}$	The <code>\_cm\_relation</code> is a <code>CMRelation</code> model, with relevant methods.
<b>Required</b>	<code>\_f(x)</code>	$f(x)$ , cf. Eq. 5.40
<b>Optional</b>	<code>\_l(x, c)</code>	$l(x, c)$ , cf. Eq. 5.49
<b>Extra</b>	<code>\_h(c)*</code>	$h(c)$ , Eq. 5.42
	<code>\_p(K, c)*</code>	$p(\kappa, c)$ , Eq. 5.47
	<code>\_mvir_to_rvir(m)</code>	$r_\Delta(m)$
	<code>\_rvir_to_mvir(r)</code>	$m_\Delta(r)$
	<code>\_rs_from_m(m)†</code>	$r_s(m)$
	<code>rho(r, m)†‡∇</code>	$\rho(r m)$ , Eq. 5.40.
	<code>lam(r, m)†‡∇</code>	$\lambda(r m)$ , if <code>\_l(x)</code> is specified.
	<code>cdf(r, m)†∇</code>	Profile CDF, $m(< r)/m_\Delta$
	<code>u(k, m)†‡</code>	$u(k m)$ , Eq. 5.46
	<code>\_rho_s(c)‡</code>	$\rho_s$ , Eq. 5.43

Table 5.11: API summary of the Profile component. Functions marked by \* are pre-defined for free, but for optimization may be specified in subclasses with an analytic form. Functions marked with † have an option of passing  $c$  directly instead of  $m$ . Functions marked with ‡ are optionally normalised by mass or density. In functions marked with ∇, the radial coordinate can alternatively be specified as  $r/r_s$  or  $x$ .

where  $F$  describes the angular part of the 3D integral. In the case of the generalised NFW, this is fully analytic:

$$F_\alpha^N(x, y) = \frac{1}{2 - \alpha} \left[ \frac{(x + y)^{2-\alpha}}{(1 + x + y)^{2-\alpha}} - \frac{|x - y|^{2-\alpha}}{(1 + |x - y|)^{2-\alpha}} \right], \quad (5.75)$$

while in the generalised Moore profile, we have

$$F_\alpha^M(x, y) = \int_{|x-y|}^{x+y} \frac{z dz}{z^p (1 + z^{3-\alpha})}, \quad (5.76)$$

which for the Moore profile with  $\alpha = 3/2$  yields

$$F_{3/2}^M = \frac{1}{3} \left[ 2\sqrt{3} \tan^{-1} \left( \frac{-1 + 2\sqrt{z}}{\sqrt{3}} \right) + \ln \left( \frac{1 + 2\sqrt{z} + z}{1 - \sqrt{z} + z} \right) \right] \bigg|_{|x-y|}^{x+y}. \quad (5.77)$$

Table 5.12: Summary of included Profile models. Note that for  $p$  and  $l$ , functions explicitly including  $c$  are the usual profiles truncated at  $r_\Delta$ , whereas those that do not explicitly include  $c$  are not truncated (these are rather unhelpful in terms of the HM).  $G$  is the Meijer- $G$  function. The functions  $T_i$  in the NFW definition are found in the text, Eq. 5.73.

NAME & REFS.	FORMULAE
NFW	$f(x) = x^{-1}(1+x)^{-2}$
Navarro, Frenk, and White (1997)	$h(c) = \frac{c}{1+c} + \ln(1+c)$
Sheth et al. (2001)	$p(\kappa, c) = \cos(\kappa) (\text{Ci}(c\kappa + \kappa) - \text{Ci}(\kappa)) + \sin(\kappa) (\text{Si}(c\kappa + \kappa) - \text{Si}(\kappa)) - \frac{\sin(c\kappa)}{c\kappa + \kappa}$
Ma and Fry (2000)	$p(\kappa) = \frac{1}{2} [(\pi - 2\text{Si} \kappa ) \sin \kappa  - 2\cos(\kappa)\text{Ci} \kappa ]$ $l(x, c) = 4\pi \begin{cases} (T_1 + T_2 + T_3) & 0 \leq x \leq c \\ (T_4 + T_5) & c \leq x \leq 2c \end{cases}$ $l(x) = \frac{8\pi}{x^2(x+2)} \left[ \frac{(x^2 + 2x + 2) \ln(1+x)}{x(x+2)} - 1 \right]$
Moore	$f(x) = \frac{1}{x^{3/2}(1+x^{3/2})}$
Moore et al. (1998)	$h(c) = \frac{2}{3} \ln(c^{3/2} + 1)$
Ma and Fry (2000)	$p(\kappa) = \frac{G_{3,9}^{7,3} \left( \frac{\kappa^6}{46656} \middle  \frac{1}{12} \begin{pmatrix} 2, 5, 11 \\ 2, 2, 5, 6, 8, 10, 11, 0, 4 \end{pmatrix} \right)}{4\sqrt{3}\pi^{5/2} \kappa }$ $l(x) = \frac{2\pi}{x} \int_0^\infty y f(y) F_{3/2}^M(x, y) dy$
Hernquist	$f(x) = \frac{1}{x(1+x)^3}$
Hernquist (1990)	$h(c) = \frac{c^2}{2(1+c)^2}$
Sheth et al. (2001)	$p(\kappa) = \frac{1}{4} [- \kappa (2\text{Ci} \kappa  \sin \kappa  + \pi \cos(\kappa)) + 2\kappa \text{Si}(\kappa) \cos(\kappa) + 2]$ $p(\kappa, c) = 1/2 - \frac{\sin(c\kappa) + (c+1)\kappa \cos(c\kappa)}{2\kappa(c+1)^2} + \frac{k}{2} [\sin(\kappa)(\text{Ci}(c\kappa + \kappa) - \text{Ci}(\kappa)) + \cos(\kappa)(\text{Si}(\kappa) - \text{Si}(c\kappa + \kappa))]$
GeneralizedNFW	$f(x) = \frac{1}{x^\alpha(1+x)^{3-\alpha}}$
Ma and Fry (2000)	$h(c) = -(-c)^\alpha c^{-\alpha} B_{-c}(3-\alpha, \alpha-2)$ $p(\kappa) = \frac{2^{-\alpha}}{\sqrt{\pi}\Gamma(3-\alpha)} G_{2,4}^{3,2} \left( \frac{\kappa^2}{4} \middle  \begin{pmatrix} (\alpha-2)/2, (\alpha-1)/2 \\ 0, 0, 1/2, -1/2 \end{pmatrix} \right)$

Table 5.12: continued...

NAME & REFS.	FORMULAE
	$l(x) = \frac{2\pi}{(2-\alpha)x} \int_0^\infty y f(y) F_\alpha^N(x, y) dy$
GeneralizedMoore	$f(x) = \frac{1}{x^\alpha (1 + x^{3-\alpha})}$
Ma and Fry (2000)	$l(x) = \frac{2\pi}{x} \int_0^\infty y f(y) F_\alpha^M(x, y) dy$

### 5.3.3.7 Concentration-mass relations

The concentration-mass-redshift relation has a diversity of physical models, which is somewhat counter-productive for our purposes of collation. We begin the process with the most commonly used physical model of Bullock (2001), and the simple power-law approximations commonly used in observational studies.

The common thread between physical models is the determination of some notion of “formation time” of a halo – whether it is the first time a certain fraction of the final mass is assembled, or the typical epoch at which a fraction of the mass collapsed, or some characteristic of the mass accretion history. For analytic models, this is usually obtained via the peak height or some close variant, so the filter function becomes an important quantity in these calculations.

Conversely, the power-law approximations generally require nothing more than the model parameters, though we provide the option to calculate the nonlinear mass if it is not passed explicitly.

Our implementation of the Bullock01 model is quite efficient, using interpolated splines to calculate the inverse of the growth factor, in order to acquire the collapse epoch:

$$z_c = D^{-1}(\sqrt{\nu(Fm)}), \quad (5.78)$$

where  $D$  is the growth factor.

We summarise the API of the CMRelation component in table 5.13, and give a concise summary of the models included in table 5.14. We also illustrate the variation among the models in figure 5.9, which shows both dependence on mass (solid lines, lower axis) and redshift (dashed lines, upper axis). We note that the increased resolution of modern simulations results in a significant deviation of best-fit models from the original physical model.

### 5.3.3.8 HOD

We keep the implementation of the HOD simple, requiring only the two mean occupation functions for centrals and satellites. Furthermore, there is an optional definition of  $m_{\min}$ , which can be used to define a precise lower bound on halo mass. Generally this



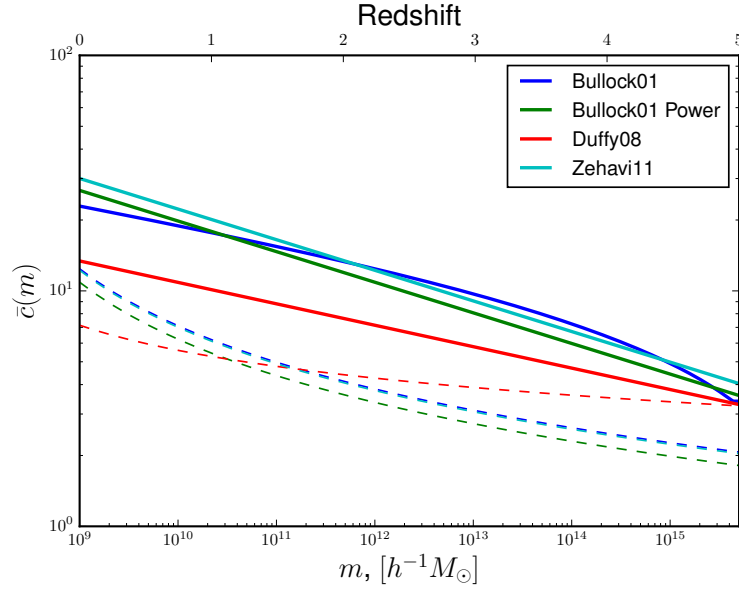


Figure 5.9: Three power-law models and one physical model (Bullock01) for the concentration-mass-redshift relation. Solid lines (lower axis) show the dependence on mass, while the dashed lines (upper axis) show dependence on redshift. We note that Zehavi11 is not a new fit to simulations, but rather a recasting of Bullock01\_Power to a different nonlinear mass.

<b>Available</b>	filter, growth, $\bar{\rho}_0$ , $\delta_c$ , $M_\star(z)$	The bulk of these quantities will only be required for physical models. $M_\star$ may be passed explicitly as a model parameter, an instance variable, or left to be calculated in the class.
<b>Required</b>	cm(m, z)	Return the concentration corresponding to $m$ at $z$
<b>Extra</b>	None	

Table 5.13: API summary of the CMRelation component.

Ref.	Name	Formula	Params.
Bullock (2001)	Bullock01	$K \frac{1+z_c}{1+z}$	$F = 0.001$ , $K = 3.4$ , $z_c$ given by Eq. 5.78
Bullock (2001)	Bullock01_Power	$\frac{a}{(1+z)^c} \left( \frac{m}{M_\star} \right)^b$	$a = 9$ , $b = -0.13$ , $c = 1$
Duffy et al. (2008)	Duffy08		$a = 6.71$ , $b = -0.091$ , $c = 0.44$ , $M_\star = 2 \times 10^{12}$
Zehavi et al. (2011)	Zehavi11		$a = 11$ , $b = -0.13$ , $c = 1$ , $M_\star = 2.26 \times 10^{12}$

Table 5.14: Summary of concentration-mass-redshift relations implemented in HALOMOD.

<b>Available</b>	None	
<b>Required</b>	<code>nc(m)</code> , <code>ns(m)</code>	Central and satellite mean occupation.
<b>Optional</b>	<code>mmin()</code>	Specification of the minimum possible halo mass. Defaults to $M_{\min}$ .
<b>Extra</b>	<code>ntot(m)</code>	Total mean halo occupation

Table 5.15: API summary of the HOD component.

Ref.	$\langle N_c \rangle_m$	$\langle N_s \rangle_m$
Zehavi et al. (2005)	$H(m - M_{\min})$	$\left(\frac{m}{M_1}\right)^\alpha$
Zheng et al. (2005)	$\frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{\log m - \log M_{\min}}{\sigma_{\log M}} \right) \right]$	$\left(\frac{m - M_0}{M'_1}\right)^\alpha$
Tinker and Weinberg (2005)	$H(m - M_{\min})$	$\exp \left( -\frac{M_{\text{cut}}}{m - M_{\min}} \right) \left(\frac{m}{M'_1}\right) H(m - M_{\min})$
Geach et al. (2012)	$F_c^B (1 - F_c^A) \exp \left[ \frac{\log_{10}(m/M_c)^2}{2\sigma_{\log M}^2} \right] \times F_c^A \left[ 1 + \operatorname{erf} \left( \frac{\log_{10}(m/M_c)}{\sigma_{\log M}} \right) \right]$	$F_s \left[ 1 + \operatorname{erf} \left( \frac{\log_{10}(m/M_{\min})}{\delta_{\log M}} \right) \right] \left(\frac{m}{M_{\min}}\right)^\alpha$
Contreras et al. (2013)	$N_c^{\text{Geach}}, \sigma_{\log M} \rightarrow x\sigma_{\log M}$	$N_s^{\text{Geach}}$

Table 5.16: Summary of included HOD parameterisations. Here  $H$  is the Heaviside step-function.

will be predominantly determined by the parameter  $M_{\min}$ , and it defaults to this value if not specified. In general it may be some function of the parameters.

We summarise the API of the HOD component in table 5.15, and present the implemented models in table 5.16. An illustration of the various effects introduced by each HOD is presented in figure 5.1.

### 5.3.3.9 Halo Exclusion

The primary task of the Exclusion component is to evaluate Eq. 5.9 (or Eq. 5.10 if appropriate) and the associated (potentially modified) mean density, Eq. 5.19. Each model, as discussed in §5.2.10, is quite unique, and therefore there is not a great deal of redundancy that can be assuaged by the class system. However, each implemented model is segregated into key methods that may provide a basis for extended models. Furthermore, the same classes can be used for either DM or galaxies, since they accept the quantity  $I_X$ .

Performance becomes a definite issue with these calculations, especially for those requiring many double-integrations. Thus we introduce alternative methods, where necessary, accelerated by just-in-time compilation.

We show the effects of the halo exclusion for four of the implemented models (two of them being modifications to  $P_{hh}^c(k)$  rather than physical halo exclusion models) in figure 5.6.

#### 5.3.4 *Extra functionality*

In addition to the core Framework and Component elements, we have implemented several pieces of functionality aimed at performing the most commonly required tasks.

##### 5.3.4.1 *Command-line interface*

Though we expect that the vast majority of usage will be interactive or via custom scripts, for additional flexibility we also provide a command-line interface (CLI), which provides a convenient way to compute specified quantities for a range of parameters. This may be useful for batch scripts which only require computation of basic quantities over a large parameter space, or for creating simple interfaces with other programming languages.

The underlying machinery for the CLI is contained in the `functional` module, which provides a top-level interface for calculating given quantities within a given Framework, for all combinations of a set of given parameters. Due to the homogeneity of the Framework definitions, this interface needs only be defined once for all frameworks.

The most challenging aspect of this function is the ordering of the implicit loops. As an example, consider a case where the user wishes to range over both redshift and  $\Omega_m$  to calculate  $n(m)$ . It is much more efficient to use redshift as the inner loop, since fewer quantities depend on it, and therefore fewer re-calculations will be performed. This implicit ordering of parameters is accounted for in the routine (if necessary) by performing a very fast low-resolution calculation and determining the number of child quantities for each parameter. This is made possible through the caching system which we have already described.

Along with the calculated quantities, the routine optionally returns unique labels for each of the iterations, since the return order is not pre-specified.

The CLI wraps this routine, allowing any parameter to be specified by name, either as a scalar or list of values to be iterated over. The output is segregated by the length of the quantity – functions of mass in one file, functions of wavenumber in another etc. This allows for simple post-calculation manipulation.

As an added benefit, the `functional` module is valuable for calculations in the web-app, in which the user may enter a series of parameters to be processed, and each must be returned with an appropriate label.

##### 5.3.4.2 *HOD population*

Quite apart from the analytic formalism outlined in this chapter, the HOD allows for creating galaxy catalogues by directly populating halos from  $N$ -body simulations. Though less efficient than its analytic counterpart, this method is more robust, since it does not depend on approximations for the several halo-based components. Indeed, several au-

thors have used this method, either alone or in conjunction with the analytic calculation (Skibba et al., 2015; Zheng and Guo, 2015). It is also useful as a sanity check.

We provide a very basic set of tools for populating halo catalogues with galaxies using the HOD models included in HALOMOD, synthesised into a CLI. Our implementation assumes a Bernoulli (Poisson) distribution for centrals (satellites), and a spherically symmetric profile for each halo, corresponding precisely to the analytic calculations. It is possible that this functionality will be extended in future versions to account for triaxial profiles (Jing and Suto, 2002) and a range of galaxy classes (such as colour, cf. Skibba and Sheth 2009).

#### 5.3.4.3 *Fitting routines*

Probably the most common application of the HM in the past decade has been to use the analytic model to fit HOD parameters using observed clustering data. We provide a very flexible interface for this purpose, both as a method and a CLI.

Again exploiting the homogeneity of the core Framework structure, we are able to present a unified interface to any Framework and an arbitrary selection of its properties. The features of the fitting capabilities include:

- CLI or functional interface
- Fit to any quantity from any any Framework
- Completely specify the base model parameters (with defaults given by normal usage).
- Arbitrary specification of variable parameters to be fit (including elements of any of the Component parameter dictionaries).
- Several parameter priors , including uniform, log-uniform, normal and multivariate normal distributions.
- Support for independent or covariant Gaussian uncertainties in the measured quantities.
- Extendible system for specifying fitting methods, with Markov Chain Monte Carlo (MCMC) fitting implemented with an Affine-Invariant Ensemble Sampler (provided by the EMCEE<sup>19</sup> package), and downhill-gradient fits using the minimization package in SCIPY<sup>20</sup>.
- Concurrent output of arbitrary properties, useful for identifying distributions of derived quantities (eg. effective bias or mass).
- CLI interface largely defined by a readable config file.

---

<sup>19</sup> [dan.iel.fm/emcee/](http://dan.iel.fm/emcee/)

<sup>20</sup> <http://www.scipy.org/>

## 5.4 EXAMPLE APPLICATION

To illustrate the utility of `HALOMOD` for common halo model applications, we here present an example motivated by typical analyses. In particular, we fit HOD models to the observed projected correlation function from the 6-degree Field Galaxy Survey (6dFGS; Jones et al., 2009), in much the same fashion as the analysis in Beutler et al. (2013, hereafter B13)<sup>21</sup>.

### 5.4.1 Setup and methodology

Our aim is to begin by performing the same basic analysis as B13 (in a limited sense), but then extend it somewhat to show the flexibility of `HALOMOD`. Thus, we use the closest sample from their analysis, called “S1”, which has a mean redshift of  $z = 0.0369$  and mean density of  $4.536 \times 10^{-3} h^3 \text{Mpc}^{-3}$ , and use the inbuilt MCMC capabilities of `HALOMOD` to fit increasingly complex sets of variables with it.

In all cases, we use the measured  $w_p(r_p)$  from S1, with errors given by a covariance matrix measured with jack-knife resampling. We also use the mean galaxy density as a measurement to be fit, with a 1% uncertainty.

All of our chains contain  $10^5$  samples before the removal of burn-in, generated with 100 ensemble walkers. Burn-in length is determined by spectral methods, as 5 times the maximum autocorrelation length. The time-per-sample required for the runs on a standard quad-core machine running two cores per chain ranges from 1/40 to 1/10 seconds depending on the variables involved.

In contrast to B13, we use the more recent cosmological parameter set of Planck Collaboration, 2015, hereafter P15, which is built-in to `HALOMOD`<sup>22</sup>. Furthermore, we use the HMF and bias models of Tinker et al. (2010), and for simplicity, we use the transfer function of Eisenstein and Hu (1999). A concise summary of our base model is found in table 5.17, which also highlights settings that differ from B13.

### 5.4.2 Results

Our first task is to re-perform the analysis of B13, which varies only the HOD parameters. Figure 5.10 shows the resulting marginalised joint-posterior for all combinations of parameters. In blue are the original results of B13 which differ appreciably from our results, especially in the mass scales. For these, the estimated uncertainty is very similar, indicating that each analysis is being performed correctly. However, the disparity in the estimated means illustrates the sensitivity of the analysis to systematic effects, such as fiducial cosmology, transfer model and numerical routines.

It is well-known that degeneracies between HOD and cosmology parameters prohibit constraints on the latter using only galaxy clustering. We explore this effect in figures 5.12 and 5.13, which show the marginalised joint-posterior for a similar run, but includ-

<sup>21</sup> It must be noted that the observed correlation functions and their covariance matrices were kindly provided by Florian Beutler.

<sup>22</sup> Note that the observable will have insignificant variation with cosmology due to the very low mean redshift of the sample.

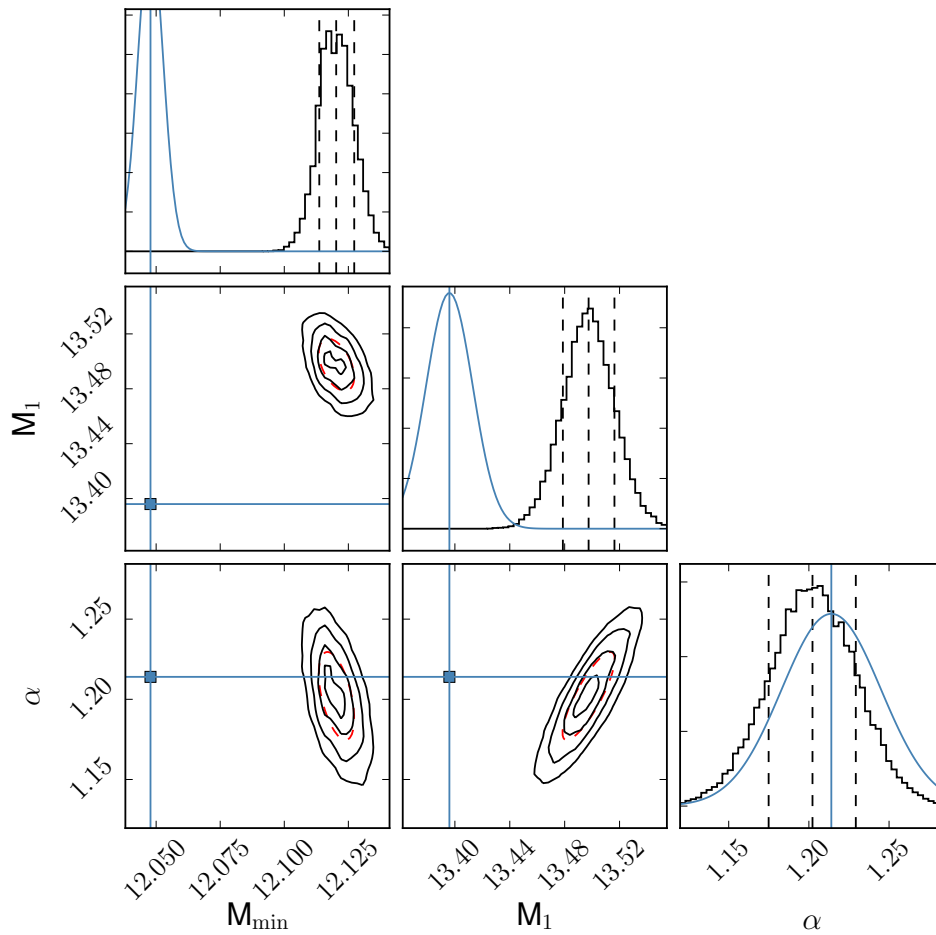


Figure 5.10: Marginalised joint-posteriors for the basic run, emulating the analysis of B13. Black histograms show the marginalised 1D posterior on each parameter, while black contours show the 2D joint-posteriors, smoothed with a Gaussian KDE. Red dashed contours indicate the  $1\text{-}\sigma$  region, when approximated by a bivariate Gaussian distribution, and should correspond to the second black contour perfectly if the true posterior is Gaussian. Blue lines and curves correspond to the result of B13 (the curves show the posterior approximated as a Gaussian).

PARAMETER	VALUE
Cosmology*, $\sqrt{v}f(v)$ *, $b(v)$ , $c(m, z)$	See table 5.18
$T(k)$ *	EH
$S(\xi_m(r))$	Tinker_SD05
$\rho(r m)$	NFW
exclusion_model	NgMatched_
lnk_min, lnk_max, dlnk	-4, 2, 0.02
dlog10m	0.04
rnum	50
proj_limit	50.0
z	0.0369
hc_spectrum	Nonlinear
takahashi	True

Table 5.17: Base model for our MCMC runs. The top section lists the various models used, the middle section lists the accuracy parameters, and the bottom section some extra options (some sample-specific). Note that the accuracy parameters were chosen to be as coarse as possible while only inducing a  $< 1\%$  error at any scale. Parameters marked with an asterisk (\*) are known to differ from the analysis of B13.

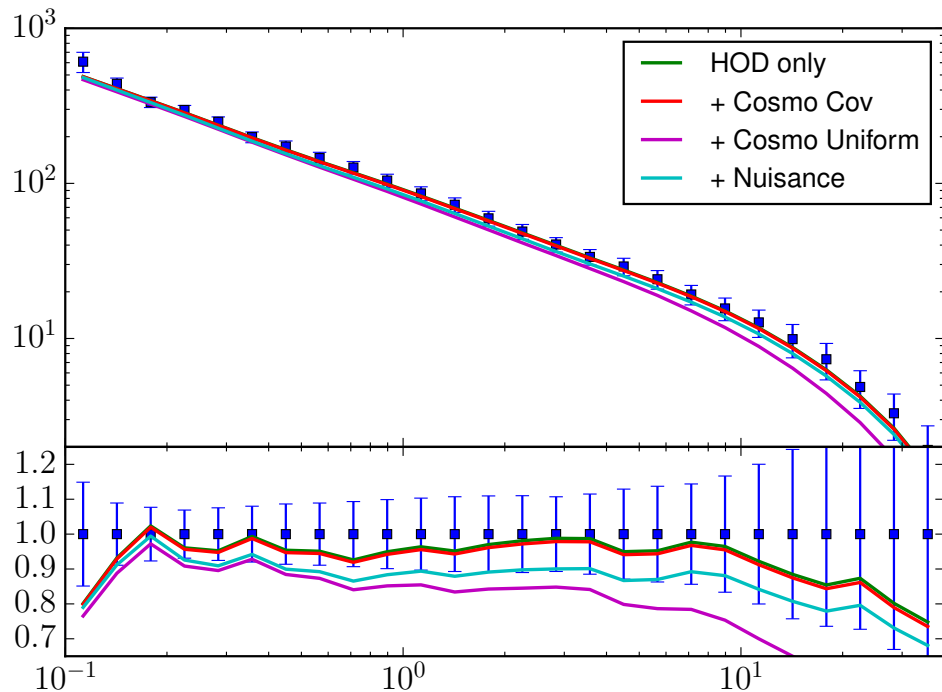


Figure 5.11: Residuals of all our fits. Data from B13 is shown as blue square markers with errorbars. Each of the fits is shown in a different colour. The only fit that is poor in general is that which had a uniformly varied cosmology. This is probably due to the way in which the ‘best’ fit was chosen: as the mean.

PARAM	PRIOR	BASELINE	DESCRIPTION
HOD			
$M_{\min}$	[10, 14]		Minimum halo mass containing galaxies.
$M_1$	[10.85, 15.2]		Halo mass on average containing one satellite.
$\alpha$	[0.5, 1.7]		Growth index of satellite occupation with halo mass
COSMOLOGY*			
$\sigma_8$	[0.4, 1.1]	0.846	RMS density within spheres of $8h^{-1}\text{Mpc}$
$n_s$	[0.7, 1.3]	0.9669	Spectral index.
$\Omega_m$	[0.2, 0.4]	0.3114	Matter density today divided by critical density
$\Omega_b$	[0.025, 0.075]	0.04888	Baryon density today divided by critical density
$H_0$	[55.0, 85.0]	67.58	Hubble constant today.
HMF			
$\alpha_0$		0.368	$z = 0$ parameters of the Tinker et al. (2010) fit.
$\beta_0$		0.589	
$\gamma_0$	$\mathcal{N}(\mu, 10\%)$	0.864	
$\phi_0$		-0.729	
$\eta_0$		-0.243	
BIAS			
$B$		0.183	$z = 0$ parameters of the Tinker et al. (2010) bias.
$b$	$\mathcal{N}(\mu, 10\%)$	1.5	
$c$		2.4	
CONCENTRATION			
$a$		6.71	Parameters of the Duffy et al. (2008) $c(m, z)$ relation
$b$	$\mathcal{N}(\mu, 10\%)$	-0.091	
$c$		0.44	

Table 5.18: Summary of the variable parameters used in MCMC runs. Listed are the parameter symbols (delineated by component), prior range, base value (in runs where it is not varied), and a description. In some runs, we impose a multivariate Gaussian prior on the cosmological variables, derived from the MCMC chains from P15. Furthermore, when we vary only  $\sigma_8$ , it takes a normal prior based on these same chains. The mean of these chains is the value listed in the baseline column. The HMF, bias and concentration parameters receive a normal prior, with 10% width and mean given in the baseline column.



ing five cosmological parameters:  $\sigma_8$ ,  $\Omega_m$ ,  $\Omega_b$ ,  $H_0$  and  $n_s$ . In figure 5.12, we leave these parameters to vary with a uniform prior, whereas in figure 5.13 they assume their prior from the covariance matrix of P15. Here, the blue lines in the HOD parameter panels are the results from the pure-HOD analysis, and the blue lines in the cosmology panels represent the best-fit P15 values.

We first note that uniform priors on the cosmology lead to virtually completely unconstrained values (within their ranges), which is in agreement with our expectation. Furthermore, while the HOD values are consistent with their estimated values with a fiducial cosmology, the uncertainties grow considerably (especially in the mass scale parameters), and the distributions are highly non-Gaussian. The latter is most likely due to the effects of the truncated posteriors of the cosmology.

When using covariant priors given by the P15 constraints, we find again that the 2PCF is not effective at constraining the cosmology, with the posteriors dominated by the priors. We do however find that the uncertainties on the HOD parameters are increased. In particular, the uncertainty on  $M_{\min}$  is approximately doubled. This highlights the need to include all known uncertainties in the analysis, even if they are to be marginalised over.

Of course, including known uncertainties in the cosmological parameters is not the final word. The various components involved in the calculation require fitted parameters, which include some uncertainty. The HALOMOD system makes it possible to easily include these uncertainties, by providing the model parameters as variables in the fit.

We take a simplistic approach for illustration and set each model parameter for the HMF, the bias and the concentration to have 10% uncertainty. We again perform the same analysis, and in figure 5.14 we plot the final marginalised distributions (without the nuisance parameters from the components and cosmology). In this case, the blue curves represent the distributions derived from the fit with covariant cosmology. We find that again, the best-fit parameters are consistent between analyses, but the estimated uncertainty is increased. Interestingly, again it is primarily in the mass scale parameters that the increase occurs, and predominantly  $M_{\min}$ . With the increased uncertainty, the tension between our results and those of B13 is relaxed considerably, with the best-fit  $M_{\min}$  from B13 just outside the  $1-\sigma$  region.

Figure 5.11 shows the model calculated from the mean of all the fitted posteriors against the data. Large scales do not contribute significantly to the fit due to the large errors, and so the fits diverge on these scales. Otherwise, the fits are all quite good except for that with a uniformly varying cosmology.

## 5.5 FUTURE DEVELOPMENT

The current implementation of HALOMOD (likewise HMF) forms a solid, but basic, framework for evaluating halo model quantities. Its flexible and extendible architecture enables user-side development with updated components and paradigms (eg. warm dark matter, or alternate dark energy scenarios). However, several features considered for future versions are worth mentioning:

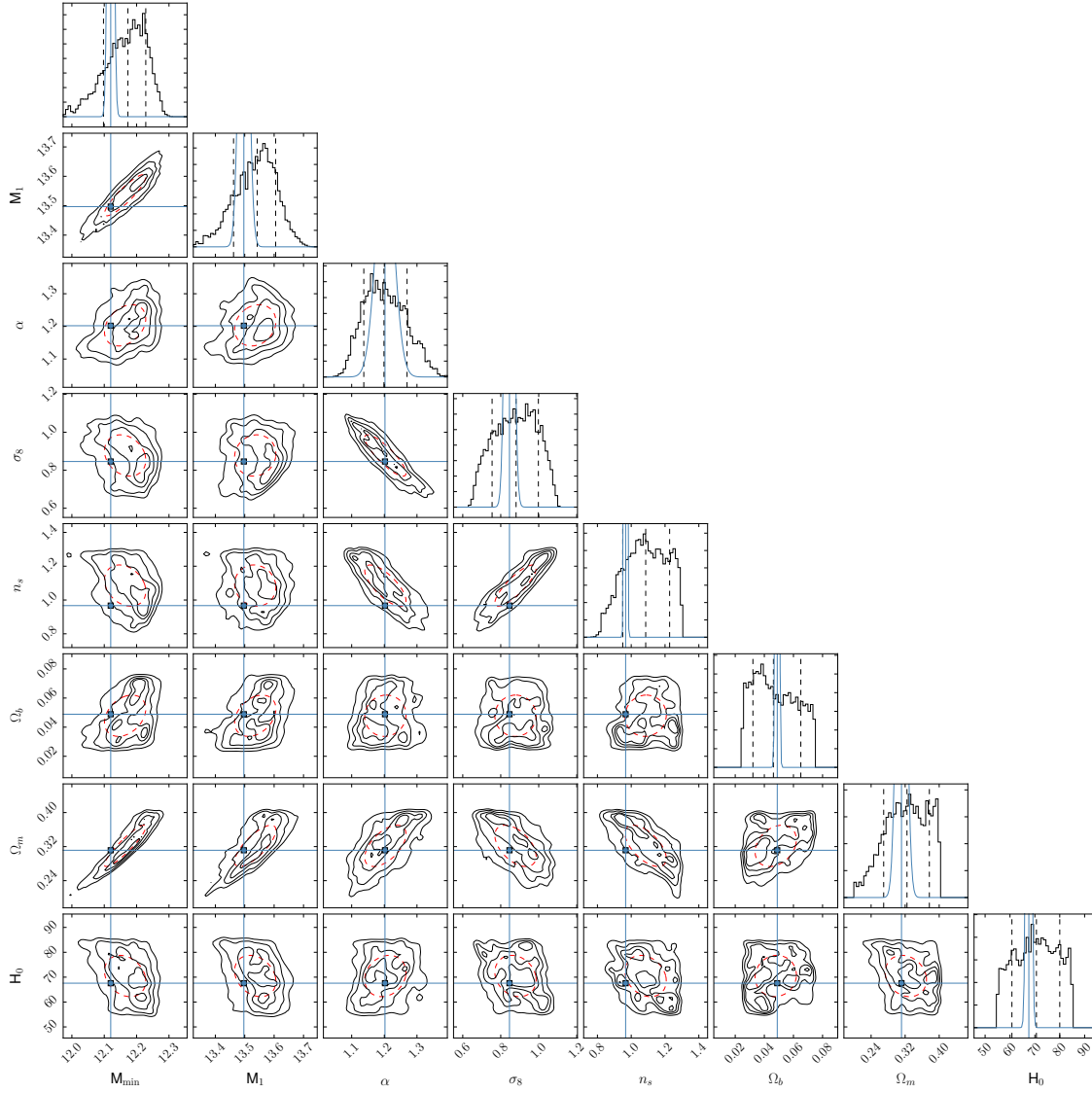


Figure 5.12: Marginalised joint-posteriors for a run with cosmology free to vary with uniform priors. The plot is similar to figure 5.10, but in this case, the blue lines/curves correspond to the best-fit HOD-only case (for the HOD parameters) and to the P15 uncertainties (for cosmological parameters). Noticeably, the cosmological parameters are virtually unconstrained.

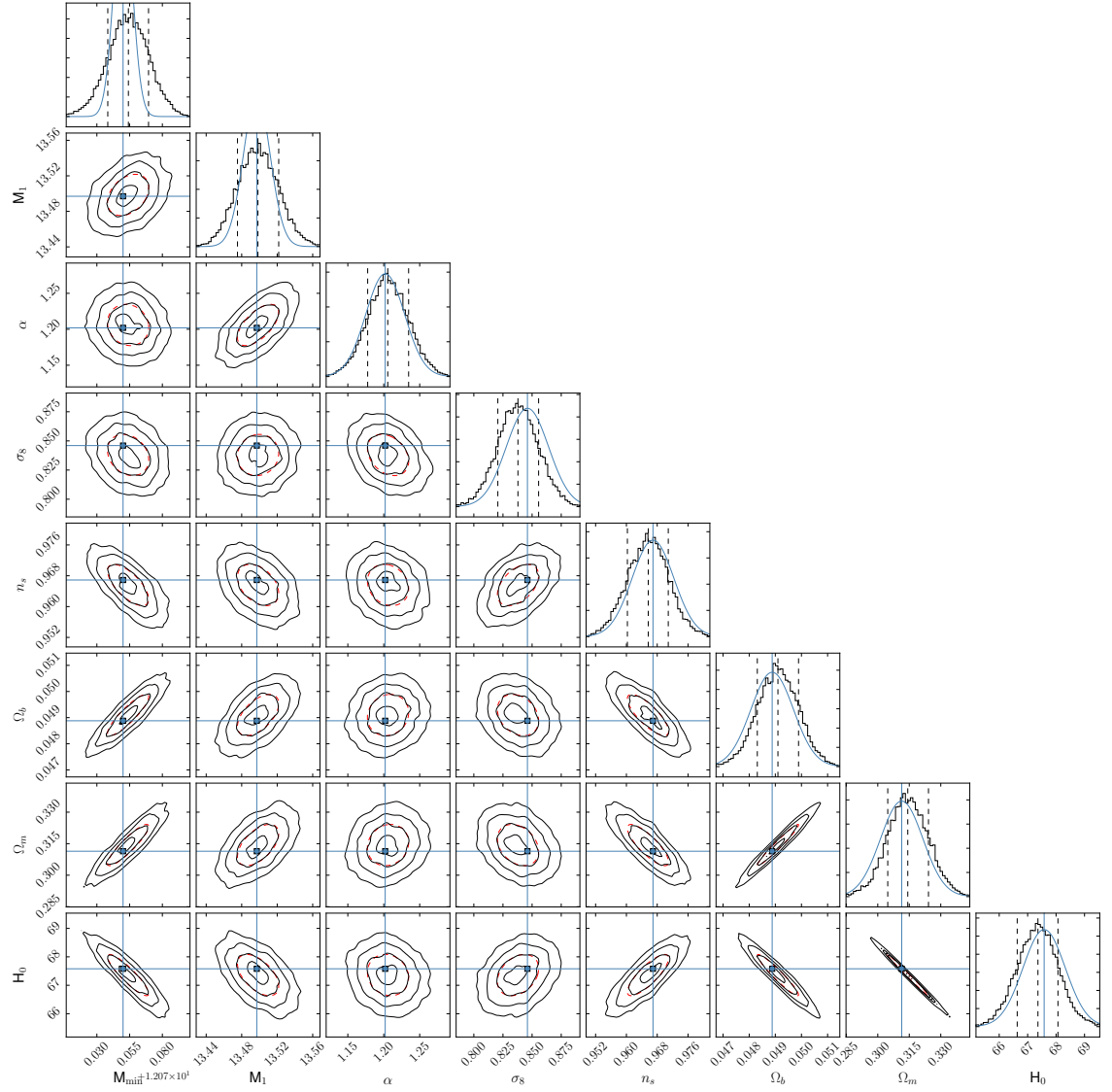


Figure 5.13: Marginalised joint-posteriors for a run with cosmology free to vary with covariant normal priors. The plot is otherwise precisely the same as figure 5.12. Cosmological parameters are relatively unchanged, but HOD parameters have looser constraints than when cosmology is fixed.

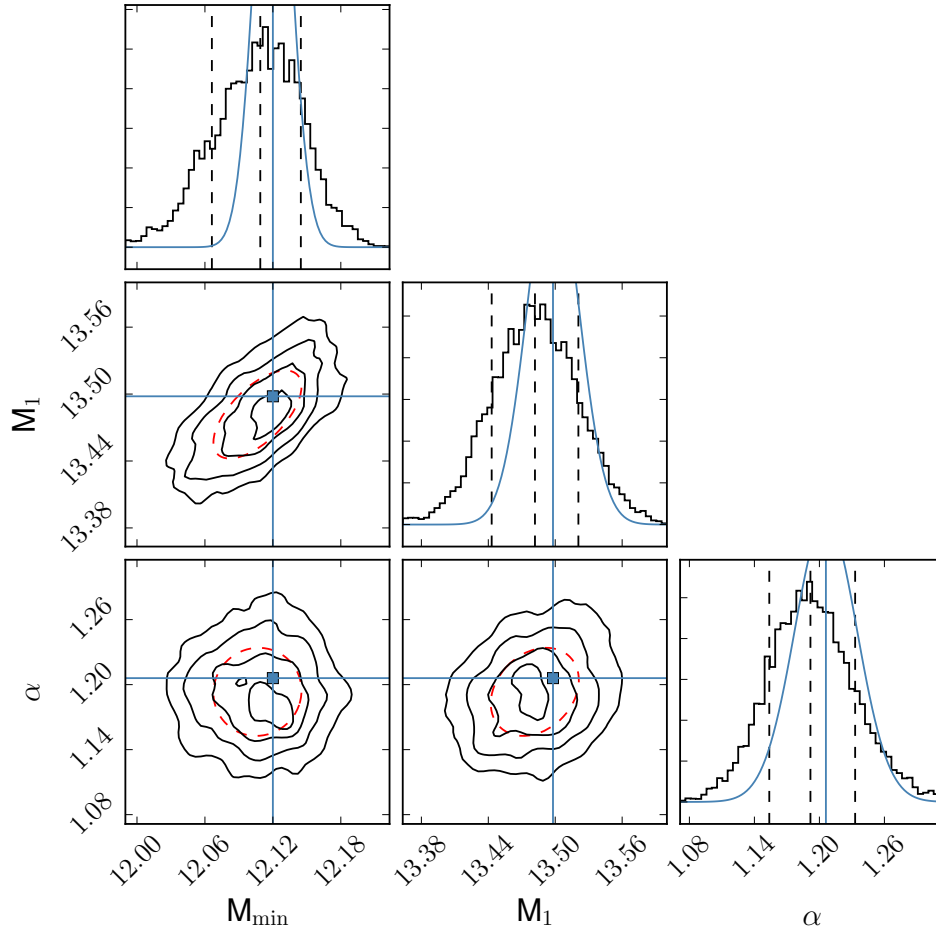


Figure 5.14: Marginalised joint-posteriors for the HOD parameters with the inclusion of both covariant cosmology and nuisance parameters from the HMF, bias and concentration. Blue curves show the distributions derived in the case with covariant cosmology only.

**WEB-INTERFACE.** The release of HMF was accompanied by a web-application exposing much of the functionality of the underlying engine. We are in the process of developing an updated version of this tool incorporating the HALOMOD code. Along with the increased range of functions, the updated application will feature a more dynamic and intuitive interface.

**ADDITIONAL AND HIGHER-ORDER STATISTICS.** Though the most popular usage of the halo model has been to analyse the two-point clustering of galaxies, there has also been a range of studies with other quantities. In particular, galaxy-galaxy weak lensing, as an independent observable, has been shown to be able to break HOD degeneracies (Leauthaud et al., 2012). Future versions of HALOMOD will include these quantities in addition to higher-order statistics such as the bispectrum and 3-point correlation function.

**BETTER DEFAULT CHOICES AND IMPROVED USABILITY.** At this point, all default parameter choices are static. Future versions will include more intelligent default choices, such as bias models chosen based on the HMF model and vice versa. This will come with a range of general improvements to the code, including documentation, examples, tests, exception handling, and introspective aspects.

**SUPPORT FOR FITTING MULTIPLE OBSERVABLES** There is a growing interest in using multiple observables when fitting the HOD (Leauthaud et al., 2011; Leauthaud et al., 2012; Tinker et al., 2013), enabling the breaking of degeneracies and the simultaneous constraint of cosmological parameters. We aim to support these kinds of analyses with arbitrary combinations of observables in the future.

## 5.6 SUMMARY

We have presented a new code, HALOMOD, for the calculation of quantities within the halo model framework. This code aims to meet five criteria: simplicity, flexibility/extendibility, comprehensiveness, efficiency and openness; we have shown how these criteria have shaped the architecture of our application.

HALOMOD is designed to be a valuable resource for the burgeoning community of researchers applying the halo model to current and next-generation surveys. To this end, we have presented an illustrative application which closely follows common usage, showing the ease with which common tasks can be performed with this new tool.

In the future we will expand the range of quantities that HALOMOD can generate, simplify the usage, and release an updated web-application to enable simple access and an intuitive interface.



---

## CONNECTING GALAXIES TO WARM DARK MATTER HALOS: CLUSTERING SIGNATURES

---

### ABSTRACT

We explore signatures of warm dark matter (WDM) on galaxy clustering, as expressed in the halo model formalism with a simple halo occupation distribution (HOD). We focus on WDM candidates roughly consistent with a recent potential detection of the decay of sterile neutrinos as a 3.5 keV X-ray line, and examine three regimes – very small, mid and BAO scales – for a set of fiducial HODs which act as a proxy for survey depth, and a range of redshifts. We find that for a set of scales  $0.01 < r < 40h^{-1}\text{Mpc}$ , a 3.3 keV WDM particle only exceed 20% maximum deviation from a best-fit CDM model at  $z > 5$ , while lowering the minimum scale to  $10^{-3}h^{-1}\text{Mpc}$  enables the same at  $z = 4$ . While the signatures that we find are not competitive with results from the Lyman- $\alpha$  forest, we conclude that with more accurate modelling and in combination with other observables, they may form a valuable independent probe.

### 6.1 INTRODUCTION

The concept of a Universe filled with dark matter (DM) has been on the table for decades (Zwicky, 1933; Rubin and Ford, 1970). With the simultaneous advent of particle physics, high-resolution  $N$ -body simulations and the detection of the *cosmic microwave background* (CMB), the story has become both more well-defined and mysterious. A wealth of evidence (Frenk and White, 2012) now supports both a dominant *dark energy* component (Schmidt et al., 1998; Perlmutter et al., 1999), and a gravitationally dominant DM component which is non-baryonic (Bergström, 2000), non-interacting, and massive: otherwise known as *cold dark matter* (CDM).

One of the remarkable successes of this model (hereafter  $\Lambda$ CDM) is its ability to describe both the large-scale density field of the early Universe reflected in the CMB (Hinshaw et al., 2013; Planck Collaboration, 2014a) and the medium-to-large-scale structure at the present day reflected in modern galaxy surveys (Percival et al., 2001; Eisenstein et al., 2005). This success at large scales has entrenched  $\Lambda$ CDM as the standard cosmological model, and ensured that any prospective alternatives must first agree on these scales.

Despite this, there exist at least three well-known tensions between  $\Lambda$ CDM and observations on small scales: (i) the “missing-satellite” problem, in which CDM over-predicts

the number of satellite galaxies in Milky-Way-sized halos (Moore et al., 1999), (ii) the “core-cusp” controversy, in which CDM predicts a divergent central density for DM halos, whereas observations have claimed centrally flattened cores (Moore et al., 1999; Gentile et al., 2009), and (iii) the predicted abundance of dwarf galaxies in cosmic voids is higher than observed (Peebles, 2001). While a great deal of effort has been applied to understanding the effects of baryonic physics on these problems, with some success, the tensions have naturally led to alternative theories on the nature of DM, which preserve the large-scale properties of CDM, but suppress the small-scale structure in some way.

The archetype of such alternatives is so-called *warm dark matter* (WDM) (Bond and Szalay, 1983; Dodelson and Widrow, 1994; Bode, Ostriker, and Turok, 2001). In this model, the DM particle is lighter and faster than its CDM counterpart, and is therefore able to ‘free-stream’ out of perturbations below a given threshold in the early Universe. This effectively eliminates small-scale perturbations in the initial fluctuation spectrum, resulting in a delay in structure formation compared to CDM, and a corresponding suppression of small-scale structure at the present day.

WDM models have garnered a great deal of attention due to their apparent ability to alleviate the aforementioned tensions of CDM, and also because suitable particles exist in minimal extensions to the Standard Model of particle physics, namely the sterile neutrino (Dodelson and Widrow, 1994; Boyarsky, Ruchayskiy, and Shaposhnikov, 2009) and the gravitino (Gorbunov, Khmelnitsky, and Rubakov, 2008). This has led to a slew of studies with the aim of constraining the WDM particle mass, under certain assumptions concerning the nature of the particle itself. A variety of methods have been used to do so, with the most common being the Lyman- $\alpha$  forest (Narayanan et al., 2000; Viel et al., 2005; Viel et al., 2006; Seljak et al., 2006; Abazajian, 2006; Viel et al., 2008; Boyarsky et al., 2009; Viel et al., 2013), but also include the effects on reionization (Barkana, Haiman, and Ostriker, 2001; Dayal et al., 2015), weak lensing (Inoue et al., 2015), abundance of high-redshift sources (Souza et al., 2013; Steinhardt et al., 2015) and Milky-Way satellite abundance (Macciò and Fontanot, 2010; Polisensky and Ricotti, 2011; Kennedy et al., 2014). The current lower-bound for a thermal particle is  $m_{\text{WDM}} \geq 3.3$  keV at the  $2\text{-}\sigma$  level, derived using the Lyman- $\alpha$  forest in Viel et al. (2013).

Adding to the interest has been the recent tentative discovery of a 3.5 keV X-ray line (Bulbul et al., 2014; Boyarsky et al., 2014), which may be a result of the decay of sterile neutrinos with  $m_s = 7.1$  keV. This would correspond to a thermal mass of  $m_{\text{WDM}} \sim 2.5 - 5$  keV dependent on the production mechanism of the particle (Schneider, 2014).

Though it has been shown that pure WDM cannot account for the core-cusp problem given current constraints on its mass (Macciò et al., 2012; Schneider et al., 2014), and is unlikely to fully account for other small-scale tensions, it is still a valuable exercise to accurately describe its nonlinear effects. Even CDM must free-stream at some scale, so the general methodology of accounting for free-streaming in the large-scale structure analysis is theoretically important. Furthermore, while it may not fully account for the tensions for which it was proposed, this does not preclude some or all of the DM from being warm, with supplementary physical processes filling the gap.

One promising avenue for independent constraints has been the recent development of the halo model in the context of WDM (Markovič et al., 2010; Smith and Markovič,



2011; Schneider et al., 2012; Benson et al., 2012; Schneider, Smith, and Reed, 2013; Schneider et al., 2014). The halo model is a semi-analytic description of the nonlinear DM density field, based on analytic and empirical models for DM halo properties, coupled with the assumption that all matter has collapsed into halos at some scale (Neyman, Scott, and Shane 1953; Ma and Fry 2000; Peacock and Smith 2000; for a review see Cooray and Sheth 2002). In principle, it is able to reproduce clustering statistics of DM at any order and on any scale, though common application has been to the two-point statistics. The WDM halo model has received several adjustments from the standard CDM case, with the goal of these adjustments in previous studies to constrain WDM through weak lensing, which is dependent on the nonlinear DM power spectrum.

In this chapter, we implement the WDM halo model within the `HALOMOD`<sup>1</sup> software framework (cf. Chapter 5), and extend its application to *galaxy* two-point clustering via a simple halo occupation distribution (HOD) model. Our goal is to determine possible WDM signatures in the clustering, so as to provide the foundation for tests in future large galaxy surveys, focussing on three fiducial WDM candidates roughly consistent with the recent 3.5 keV X-ray line: 2.0, 3.3 and 5 keV.

The chapter is set out as follows: §6.2 describes the halo model formalism, especially in the context of WDM, and introduces our simple test statistic, §6.3 describes the signatures of WDM on the correlation function at different scales, and the results of our test statistic, and we discuss future developments and conclude in §6.5.

## 6.2 METHODS

We set out to pair simple HOD models with halo model prescriptions for clustering, to determine expected galaxy clustering signatures in WDM. In this section we briefly introduce the standard halo model formalism (for a more thorough review, see eg. Cooray and Sheth (2002) and for updated methodology see Tinker and Weinberg 2005), and then describe the modifications made to it in recent studies (Smith and Marković, 2011; Schneider, Smith, and Reed, 2013; Schneider, 2014) to extend the model to WDM cosmologies. Finally we present the simple test that we perform to determine conditions of detection.

### 6.2.1 The halo model formalism

The halo model operates on the assumption that all matter exists within virialized structures called haloes. Under this assumption, the full density field can be reconstructed as the sum of contributions from all haloes. This is tractable since haloes have empirically been shown to obey simple universal models for their number density, profile and biasing with respect to the underlying matter from high-resolution  $N$ -body simulations. These empirical relations enable us to probe scales well into the nonlinear regime, past the reach of traditional perturbation theory techniques (Smith et al., 2003).

As an extension to the basic halo model, it is common to determine the *galaxy* distribution by the additional assumption that the distribution of galaxy numbers within

<sup>1</sup> <https://github.com/steven-murray/halomod.git>

haloes is simply dependent on the halo mass. This is termed the *halo occupation distribution* (HOD) (Berlind et al., 2003), and is generally modelled as the sum of a Poisson component describing *satellite* galaxy abundance and a Bernoulli component describing *central* galaxy abundance (Kauffmann et al., 2004; Zheng et al., 2005).

The combination of the four basic ingredients mentioned – a mass function  $n(m)$ , halo profile  $\rho(r|m)$ , bias function  $b(m)$  and HOD  $N(m)$  – provides a complete description of the galaxy density field, and therefore (in principle) any clustering statistic on any scale. In this work, we focus on the galaxy-galaxy autocorrelation function (2PCF), which has had extensive application over the past decade (Moustakas and Somerville, 2002; Bullock, Wechsler, and Somerville, 2002; Magliocchetti and Porciani, 2003; Yan, Madgwick, and White, 2003; Zheng, 2004; Zehavi et al., 2005; Hamana et al., 2006; Zheng, Coil, and Zehavi, 2007; Blake, Collister, and Lahav, 2008; Brown et al., 2008; Quadri et al., 2008; Wake et al., 2008a; Zheng et al., 2009; Ross, Percival, and Brunner, 2010; Zehavi et al., 2011; Wake et al., 2011; Beutler et al., 2013; Mostek et al., 2013; Palamara et al., 2013; Dolley et al., 2014; Guo et al., 2014; Skibba et al., 2014; Skibba et al., 2015; Kim et al., 2015; McCracken et al., 2015), but not in the context of WDM.

Customarily, the calculation of the 2PCF is broken into two regimes, the so-called 1-halo and 2-halo terms, which arise from correlations between galaxies within a single halo, and galaxies in separate haloes respectively. The final 2PCF is thus

$$\xi_{gg}(r) = [\xi_{gg}^{1h}(r) + 1] + \xi_{gg}^{2h}(r). \quad (6.1)$$

#### 6.2.1.1 Two-halo term

The 2-halo term is typically calculated in Fourier space by

$$P_{gg}^{2h}(k, r) = \frac{1}{n_g^2} \int \int \prod_{i=1}^2 dm_i n(m_i) N(m_i) u(k|m_i) P_{hh}^c(k, r, m_1, m_2), \quad (6.2)$$

where  $u(k|m)$  is the normalised Fourier-transform of the halo profile,  $n_g$  is the mean galaxy density

$$n_g = \int n(m) N(m) dm, \quad (6.3)$$

and  $P_{hh}^c(k, r, m_1, m_2)$  is the power spectrum of halo centres, which is in general a complicated function of scale and mass.

On reasonably large scales, we expect halo centres to be linearly biased with respect to the dark matter, and so it is customary to write

$$P_{hh}^c(k, r, m_1, m_2) \simeq b(m_1, r) b(m_2, r) P_m(k), \quad (6.4)$$

where  $b(m, r)$  is the (possibly scale-dependent) first-order deterministic bias of halos with respect to dark matter, and  $P_m(k)$  is the matter power. We follow usual practice, and calculate the linear power using CAMB<sup>2</sup> (Lewis, Challinor, and Lasenby, 2000),

<sup>2</sup> <http://camb.info>

applying the nonlinear corrective fit of Smith et al. (2003) with updated parameters from Takahashi et al. (2012) to generate  $P_m(k)$ .

An added complication with the 2-halo term is that of *halo exclusion*, which is the condition that we should not double-count correlations between galaxies in the same halo by allowing them in the 2-halo term. There are several methods which aim to account for this behaviour (Zheng, 2004; Tinker and Weinberg, 2005; Smith, Desjacques, and Marian, 2011) and we use the so-called ‘ $n_g$ -matched’ method of Tinker and Weinberg (2005), which approximates the effect of ellipsoidal exclusion in a distribution of triaxial haloes.

With this method Eq. 6.2 becomes separable, and we arrive at

$$P_{gg}'^{2h}(k, r) = P_m(k) \left[ \frac{1}{n_g'} \int_0^{m_{\text{lim}}} dm n(m) N(m) b(m, r) u(k|m) \right]^2, \quad (6.5)$$

where  $n_g'$  is the number density of galaxies in halos with  $m < m_{\text{lim}}$ , which is defined by the  $n_g$ -matched method.

To convert the power spectrum to the desired correlation function, we perform a Hankel transform:

$$\zeta(r) = \frac{1}{2\pi^2} \int k^2 P(k) j_0(kr) dk, \quad (6.6)$$

where  $j_0 = \sin(kr)/kr$  is the zeroth-order spherical Bessel function. The final 2-halo term is given by

$$\zeta_{gg}^{2h}(r) = \left( \frac{n_g'}{n_g} \right)^2 \left[ 1 + \zeta_{gg}'^{2h}(r) \right] - 1. \quad (6.7)$$

### 6.2.1.2 One-halo term

Following the division of the HOD into central and satellite components, the 1-halo term is typically calculated as the sum of a central-satellite ( $c-s$ ) and satellite-satellite ( $s-s$ ) term, quantifying the correlations of the central galaxy with its satellites, and between satellites respectively. These can be written

$$\zeta_{gg}^{1h,c-s}(r) = \frac{2}{n_g^2} \int n(m) N_c(m) N_s(m) \frac{\rho(r|m)}{m} dm, \quad (6.8)$$

$$\zeta_{gg}^{1h,s-s}(r) = \frac{1}{n_g^2} \int n(m) N_c(m) N_s^2(m) \lambda(r|m) dm \quad (6.9)$$

where  $N_c$  and  $N_s$  are the mean central and satellite occupation functions respectively, and  $\rho/m$  and  $\lambda$  are the mass-normalised halo density profile and its self-convolution (Ma and Fry, 2000).

Finally we have

$$\zeta_{gg}^{1h}(r) = \zeta_{gg}^{1h,c-s}(r) + \zeta_{gg}^{1h,s-s}(r) - 1, \quad (6.10)$$

where the “-1” arises because  $1 + \xi$  is proportional to the pair-counts which are measured in Eqs. 6.8.

### 6.2.1.3 HOD Model

We do not expect the broad characteristics of the HOD to change in WDM, though the precise parameters may. While a dedicated study of the expected HOD from numerical simulations would be interesting, it is beyond the scope of this chapter. Thus, we conservatively choose to use a standard CDM parameterisation. The effects of our choice of parameters are explored further in §6.3.2.2.

Though many parameterisations have been proposed, for the purpose of this study we are most interested in capturing the basic effects of the HOD, which are characterised by the formulation of Zehavi et al. (2004), in which the mean central occupation is a step function:

$$N_c(m) = \begin{cases} 0 & m < M_{min} \\ 1 & m \geq M_{min} \end{cases}, \quad (6.11)$$

and the satellite occupation is a power-law:

$$N_s(m) = \left( \frac{m}{M_1} \right)^\alpha. \quad (6.12)$$

$M_{min}$  controls the lowest possible halo mass that contains galaxies in the given sample,  $M_1$  is the halo mass at which there is on average one satellite galaxy, and  $\alpha$  controls the growth of satellite galaxy occupation with halo mass.

To ensure that the ‘central condition’ holds (ie. that if a halo contains a galaxy, it contains a central galaxy), we set the total occupation to be

$$N(m) = N_c(m)(1 + N_s(m)). \quad (6.13)$$

### 6.2.2 WDM halo model

A first-principles extension of the halo model framework to account for WDM has been developed in Smith and Markovič (2011). The primary goal of this extension is to account for a non-negligible fraction of mass that cannot be treated hierarchically since it is never virialized, which arises in WDM models because of the smoothing of small-scale structure. However, in the treatment of galaxies, we need not be concerned with this smooth component, as we will still assume that all galaxies are bound to haloes, so no modifications are necessary to the framework we have already presented.

However, there are considerable effects of the WDM on the components of the framework, which we summarise here.

### 6.2.2.1 Transfer function

The primary effect of WDM is to introduce a cutoff in the power spectrum on small scales. This is typically modelled as an extra transfer function applied to the standard CDM spectrum:

$$T_{\text{WDM}}(k) = \sqrt{\frac{P_{\text{WDM}}(k)}{P_{\text{CDM}}(k)}}. \quad (6.14)$$

Though this is properly calculated by solving the full Einstein-Boltzmann equations with the relevant species, we follow usual practice (Smith and Markovič, 2011; Dunstan et al., 2011; Schneider et al., 2012; Murray, Power, and Robotham, 2013b; Bose et al., 2015) and use the fitting function of Viel et al. (2005), which is similar to that of Bode, Ostriker, and Turok 2001:

$$T_{\text{WDM}} = (1 + (\alpha k)^{2\nu})^{-5/\nu}, \quad (6.15)$$

with  $\nu = 1.12$  (compared to 1.2 in Bode, Ostriker, and Turok (2001)) and

$$\alpha = 0.049 \left( \frac{m_x}{1 \text{ keV}} \right)^{-1.11} \left( \frac{\Omega_c}{0.25} \right)^{0.11} \left( \frac{h}{0.7} \right)^{1.22}. \quad (6.16)$$

We also follow Schneider et al. (2012) and define a “half-mode” scale and corresponding mass, where the amplitude of the WDM transfer function is 1/2 (and thus the power is suppressed by a factor of 4):

$$\lambda_{\text{hm}} = 2\pi\alpha \left( 2^{\mu/5} - 1 \right)^{-0.5/\mu} \quad (6.17)$$

$$m_{\text{hm}} = \frac{4\pi}{3} \bar{\rho}(z) \left( \frac{\lambda_{\text{hm}}}{2} \right)^3. \quad (6.18)$$

These values give some insight into the scales at which the WDM has an appreciable effect.

We plot the CDM spectrum along with those for three choices of WDM mass in figure 6.1. Vertical lines in each colour correspond to the half-mode scale. There is a clear break in the power at large  $k$  in each case, moving to the right for heavier particles.

Note that the nonlinear corrections of Smith et al. (2003) are applied to the linear power spectrum after being modified by the WDM transfer function when calculating  $P_{hh}^c$ . The validity of this correction has received some attention (Viel et al., 2005; Markovič et al., 2010; Smith and Markovič, 2011; Schneider et al., 2012), and on large scales (where effects of  $P_{hh}^c$  are non-negligible in the halo model framework) it performs remarkably well.

### 6.2.2.2 Halo Mass Function

In the context of the halo model, the halo mass function (HMF) is typically (but see eg. Hahn and Paranjape, 2013) derived via the extended Press-Schechter formalism (Press and Schechter, 1974; Bond et al., 1991; Sheth and Tormen, 1999), which provides a semi-analytic estimate that corresponds tightly to simulations. This formalism connects the

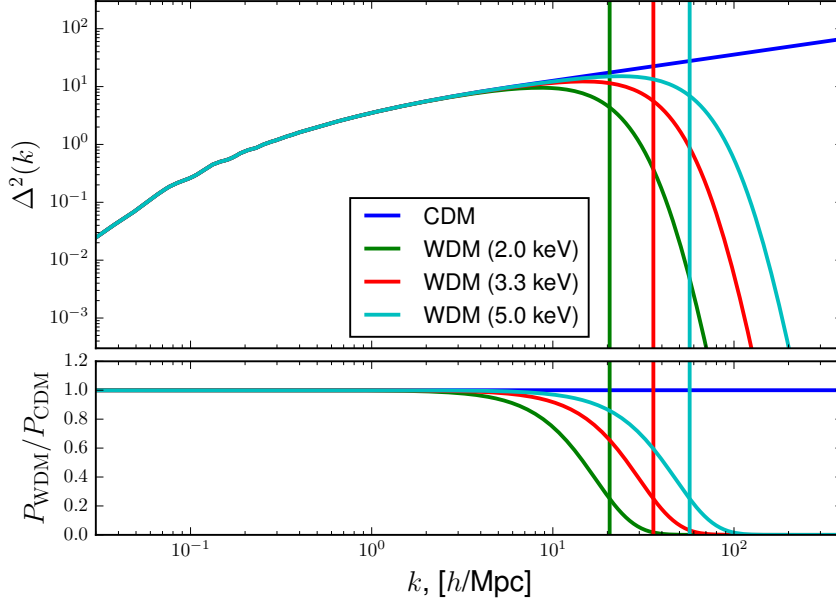


Figure 6.1: Comparison of  $z = 0$  CDM power spectrum with those for our three fiducial WDM candidates. There is clear suppression of small-scale structure in WDM. Vertical lines correspond to half-mode scales in each case.

abundance of haloes of a given mass to peaks in the Gaussian linear density field, by assuming that there is a critical density threshold above which a region is certain to collapse. Thus, the all-important quantity is the rms of the density fluctuations,  $\sigma(m, z)$ , which in the Gaussian context completely specifies the fraction of mass that is collapsed on a given scale.

The calculation of this mass variance inherently involves a choice of window (or filter) function, which defines the mass at each point. Common choices for this window function have been a top-hat in real space (‘tophat’), Gaussian in real-space, and top-hat in Fourier space (‘sharp- $k$ ’), with the tophat being the option of choice for most empirical studies of the HMF.

However, in the context of WDM, it has become apparent that it is advantageous to choose the sharp- $k$  window (Barkana, Haiman, and Ostriker, 2001; Benson et al., 2012; Schneider, Smith, and Reed, 2013; Schneider, 2014). This is due to the low-mass asymptotic limit of the HMF, which for the tophat and Gaussian windows is divergent towards infinity, while for the sharp- $k$  window goes to zero, as we expect for WDM. This effect is made possible by the fact that the mass variance in WDM converges to a constant at low mass, as seen in figure 6.2.

The disadvantage of using the sharp- $k$  window function is that it does not afford a clear prescription of the enclosed mass in a given radius. However, it has generally been accepted (Lacey and Cole, 1993; Benson et al., 2012; Schneider, Smith, and Reed, 2013) that the mass should follow the standard cubic relationship with radius, with a free parameter which can be chosen to fit the results of simulations:

$$m = \frac{4\pi}{3} \bar{\rho}(z) [cR]^3. \quad (6.19)$$

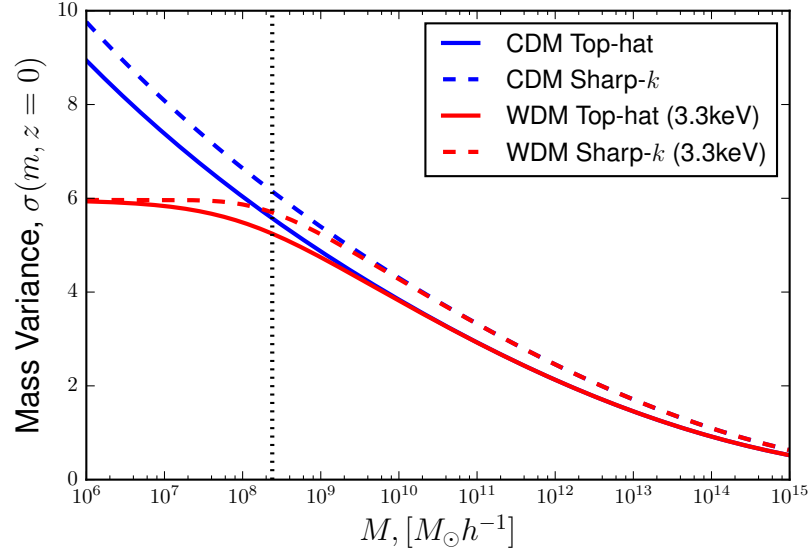


Figure 6.2: Mass variance in CDM and 3.3 keV WDM (blue and red) using tophat and sharp- $k$  window functions (solid and dashed lines). The half-mode mass is shown by the dotted vertical line. For a given mass, the sharp- $k$  window enhances the derived variance compared to the usual tophat window. The curves are normalised at  $R = 8h^{-1}\text{Mpc}$ , but this scale corresponds to different masses in the two windows.

For the standard tophat,  $c \equiv 1$ , but for the sharp- $k$  window function, both Benson et al. (2012) and Schneider (2014) find that  $c = 2.5$  provides an excellent fit to the HMF from simulations.

The extended Press-Schechter formalism predicts that the HMF has the form

$$n(m) = \frac{dn}{dm} = \frac{\rho_0}{M^2} f(v) \left| \frac{d \ln \sigma}{d \ln m} \right|, \quad (6.20)$$

where  $v = (\delta_c/\sigma)^2$  is the peak-height, with  $\delta_c \simeq 1.686$  the critical density for collapse, and  $f(v)$  is universal with respect to cosmology and redshift (but see eg. Tinker et al., 2008; Watson et al., 2013). The form of  $f(v)$  relevant for the collapse of triaxial halos is (Sheth and Tormen, 1999)

$$f(v) = A \sqrt{\frac{2qv}{\pi}} (1 + (qv)^{-p}) \exp(-qv/2), \quad (6.21)$$

where  $A = 0.3222$  is a normalisation which enforces the criterion that all mass should be in halos at some scale,  $p$  is a free parameter to be fit to simulations (Sheth, Mo, and Tormen (2001) prescribe the value  $p = 0.3$ ), and  $q$  is predicted to be 1 by the ellipsoidal collapse theory, but provides tighter fits to empirical results with  $q = 0.707$ .

However, Schneider, Smith, and Reed (2013) find that when using the sharp- $k$  window function, using the originally predicted  $q \equiv 1$  matches the tophat-derived HMF extremely well. In effect, the parameter  $q$  is replaced by the mass-association parameter  $c$ .

We show the HMF of CDM and three choices of WDM, using both the tophat and sharp- $k$  windows, in figure 6.3. Though there is a clear suppression of the HMF at low mass in the tophat WDM scenarios, it still diverges rather than being completely

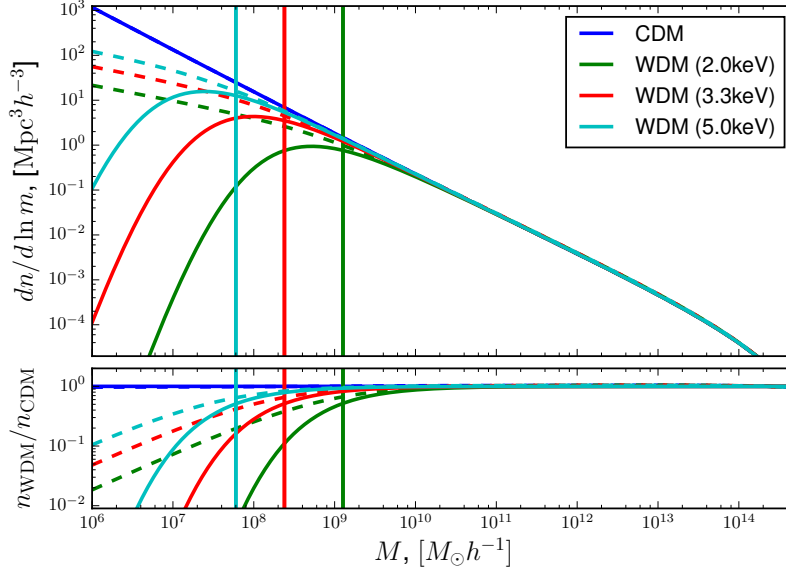


Figure 6.3: Mass functions for CDM and our 3 fiducial WDM candidates, using both the tophat (dashed) and sharp- $k$  (solid) window functions. Vertical lines again represent the half-mode mass for each WDM scenario. All methods and cosmologies align at high mass. At low mass, WDM has suppressed density, though in the case of the tophat filter, the density still diverges to infinity, while the sharp- $k$  filter has the expected behaviour of being truncated towards the free-streaming scale.

truncated as we would expect from free-streaming. The sharp- $k$  model corrects this behaviour with a sharp turnover below the half-mode mass (indicated by the vertical lines). All curves agree remarkably well at high mass.

We use the sharp- $k$  window function (with  $c = 2.5$ ) for all calculations involving the HMF (and bias) in this study.

### 6.2.2.3 Halo bias

To first order, the density field of halos can be related to that of dark matter by a deterministic linear bias:

$$\delta_h(\mathbf{x}|m) = b(m)\delta_m(\mathbf{x}). \quad (6.22)$$

The linear bias is predicted along with the HMF by the extended Press-Schechter formalism, and for the case of ellipsoidal collapse, gives (Sheth, Mo, and Tormen, 2001)

$$b(v) = 1 + \frac{qv - 1}{\delta_c(z)} + \frac{2p}{\delta_c(z)(1 + (qv)^p)}, \quad (6.23)$$

where the values of  $(p, q, A)$  are the same as in Eq. 6.21.

Though updated HMF/bias models have appeared in the literature (eg. Tinker and Weinberg, 2005; Warren et al., 2006; Tinker et al., 2008; Tinker et al., 2010; Watson et al., 2013), the uncertainty in the relation within WDM simulations to date motivates the usage of the more physically motivated forms of Sheth, Mo, and Tormen (2001) in this study.



In addition, we apply scale-dependent corrections to the bias after Tinker and Weinberg (2005), using

$$b^2(m, r) = b^2(m) \frac{[1 + 1.17\xi_m(r)]^{1.49}}{[1 + 0.69\xi_m(r)]^{2.09}}. \quad (6.24)$$

#### 6.2.2.4 Halo Profile

The shape of the 1-halo term is predominantly influenced by the average density profile of halos. We adopt two profiles in this analysis. Our first model, used by default throughout the chapter, is that of Navarro, Frenk, and White (1997, hereafter NFW), which has been shown to be an excellent fit to the dark matter distribution within haloes. However, we also consider that the galaxy number-density distribution within haloes does not trace the dark matter, but is rather less centrally-concentrated (Gao et al., 2004; Giocoli et al., 2010). Indeed, high-resolution numerical simulations such as the Aquarius project (Springel et al., 2008) have found that the substructure profile follows a shallow Einasto profile (Einasto, 1965), which we take as our second model.

The Einasto profile has two free parameters, the concentration  $c$  and the logarithmic slope  $\alpha$ , and has the form

$$\rho(r) = \rho_s(c) \exp\left(-\frac{2}{\alpha}(x^\alpha - 1)\right). \quad (6.25)$$

A value of  $\alpha = 0.18$  provides a good approximation to the NFW profile. However, our choice to use the Einasto profile is primarily motivated by a desire to predict the *galaxy* correlations, rather than the pure matter correlations. The Aquarius simulations reveal that  $\alpha = 0.678$  provides a good fit to the substructure profile, and we follow this prescription when using the Einasto model.

We note that the Aquarius results are based on the standard CDM cosmology, and we might expect the WDM substructure profile to differ due to a dearth of low-mass satellites. To accurately predict the substructure profile of WDM haloes is itself an interesting project, but beyond the scope of this paper. We consider that using the shallower CDM substructure profile is a step in the right direction, and its comparison to the standard assumption of galaxies tracing the dark matter will provide the insight for which we have employed it.

#### 6.2.2.5 Concentration-Mass Relation

The concentration parameter has been shown to correlate with halo mass (Bullock et al., 2001), and thus we may, via the mean concentration-mass-redshift relation, transform the profile to depend solely on halo mass.

Various models have been proposed for the  $\bar{c}(m, z)$  relation, both semi-analytic (Bullock et al., 2001; Ludlow et al., 2014; Correa et al., 2015c; Ludlow et al., 2016) and empirical (Macciò et al., 2007; Duffy et al., 2008; Macciò, Dutton, and Bosch, 2008). While the empirical models are commonly employed in applications of the HOD due to their computational efficiency (eg. Cooray and Sheth, 2002; Zehavi et al., 2005; Zehavi et al., 2011; Beutler et al., 2013; Mohammed et al., 2014), it is unclear how relevant they are to

WDM cosmologies. Bose et al. (2015) proposed an empirical correction for an empirical  $\bar{c}(m, z)$  relation, but we follow the more recent work of Ludlow et al. (2016), who use a physically motivated model to explore the  $\bar{c}(m, z)$  relation in WDM cosmologies.

Their model closely follows the original idea of NFW, who assumed that a halo's characteristic density is proportional to  $\rho_c$ , the critical density of the universe at the redshift at which it collapsed. NFW proposed that the characteristic density be given by  $\delta_c$  and the collapse redshift be defined when half the virial mass,  $M_0$ , of the halo was collapsed into haloes with mass  $m > fM_0$ . Ludlow et al. (2016) modify this scheme slightly by suggesting that the characteristic density be given by the mean density of the halo within the scale radius  $r_s$  and the characteristic mass as the mass within the scale radius  $M(< r_s)$ .

This results in a coupled set of equations to solve for  $c$  and the formation redshift  $z_f$ :

$$\frac{\rho_s}{\rho_0} = C \times \left( \frac{H(z_f)}{H(z_0)} \right)^2 \quad (6.26)$$

$$\frac{M(< r_s)}{M_0} = \text{erfc} \left( \frac{\delta_c / D^+(z_f) - \delta_c / D^+(z_0)}{\sqrt{2(\sigma_0^2(fM_0) - \sigma_0^2(M_0))}} \right), \quad (6.27)$$

where  $D^+$  is the linear growth factor and  $\sigma_0^2$  is the mass variance at  $z = 0$ . The model has two free parameters, similar to that of NFW:  $C$  which relates the background and characteristic densities at formation, and  $f$  which is the fraction of the final mass which has collapsed at  $z_f$ . The left-hand side of each equation incorporates the profile shape, and thus the concentration, while the right-hand side incorporates the formation redshift.

Using the language of §5.2.8, we note that for an arbitrary profile  $f(x)$  where  $x = r/r_s$  and

$$h(c) = \int_0^c x^2 f(x) dx, \quad (6.28)$$

we have the following definitions:

$$\frac{\rho_s}{\rho_0} \equiv \frac{\Delta_h c^3 h(1)}{h(c)} \quad (6.29)$$

$$\frac{M(< r_s)}{M_0} \equiv \frac{h(1)}{h(c)}. \quad (6.30)$$

In figure 6.4, we show the effects of WDM on  $\bar{c}(m, z = 0)$ . In pure hierarchical CDM, since larger halos form later, they have a lower mean concentration. However, in WDM cosmologies, structure formation is not purely hierarchical, but is a mixture of top-down and bottom-up (Paduroiu, Revaz, and Pfenniger, 2015), so we expect the smallest scales to be collapsing later as well. This is quantified by the turnover in the WDM relations in figure 6.4. Note also that the turnover is more than an order of magnitude above the half-mode mass in each case, rendering the distinguishable scales of WDM more accessible.

We finally note that while the concentration-mass-redshift relation we have chosen is the most robust available, it is specified in terms of the dark matter profile, rather than

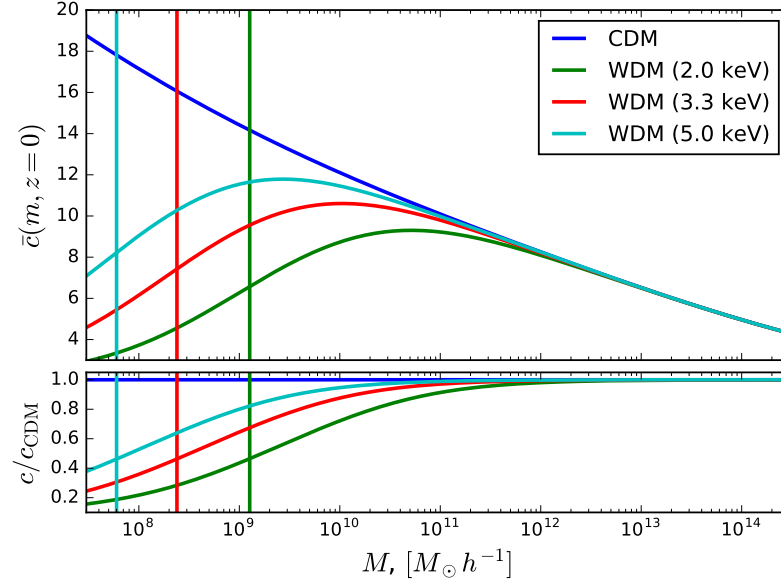


Figure 6.4: Concentration-mass relations for CDM and our 3 fiducial WDM candidates. Vertical lines again represent the half-mode mass for each WDM scenario. At high mass, the models are indistinguishable. At low mass, the WDM concentrations undergo a turnover. Note that the turnover is more than an order of magnitude higher than the half-mode mass.

the galaxy number-density profile. We do not expect these to be the same in general, but we leave this problem to future work.

#### 6.2.2.6 Implementation

We implement all of the components defined here within the `HALOMOD`<sup>3</sup> Python code (cf. Chapter 5). This code provides a general framework for integrating halo model components into clustering calculations in a simple, flexible and efficient manner.

The code also provides native support for running Markov Chain Monte Carlo (MCMC) or downhill-gradient fits with any parameter of the halo model for any quantity, which we utilise in this work.

#### 6.2.3 Test Statistics

We identify three regimes of interest in the correlation function: large-scales affected only by the two-halo term, small-scales affected only by the 1-halo term, and transition scales. Each region we treat (at first) separately, first modelling expected signatures, and secondly applying a statistical test of the likelihood of detection of WDM, for our three fiducial candidates. In this section we describe our methodology concerning this simple statistic.

We note that the very large scales have different properties, and are treated in a different way to the small scales. We leave the description of this regime to §6.3.1.

<sup>3</sup> <https://github.com/steven-murray/halomod.git>

We expect the largest differences between WDM and CDM to occur on the smallest scales, so the results presented here are quite sensitive to the smallest scale probed. Typically, at low redshift, scales below  $r \sim 0.1h^{-1}\text{Mpc}$  are extremely difficult to probe due to systematic effects such as fibre collisions. Furthermore, scales below this begin to encroach on galactic-size scales, at which the halo model formalism surely breaks down. We thus choose to use  $r_{\min} = 0.1h^{-1}\text{Mpc}$  at our minimum redshift of  $z = 0.05$ , and keep a constant angular resolution for higher redshifts, according to the angular diameter distance. This results in minimum scales of  $r_{p,\min} = 8.5 \times 10^{-1}$  and  $6.2 \times 10^{-1}h^{-1}\text{Mpc}$  for  $z = 1$  and  $5$  respectively.

#### 6.2.3.1 Observable

On small-to-medium scales, the quantity that is typically measured is the *projected correlation function* (Beutler et al., 2013)

$$w_p(r_p) = 2 \int_{r_p}^{\infty} \frac{r\zeta(r)}{\sqrt{r^2 - r_p^2}} dr. \quad (6.31)$$

Our analyses (other than in the purely 2-halo regime) are all performed in terms of this quantity.

#### 6.2.3.2 Parameter space

Our main aim is to establish the strength of detected signatures with various survey specifications. We explore two aspects: survey depth and redshift. We perform our analysis over a range of combinations of these variables, for each scale regime and particle mass.

In principle, the constraints will also vary depending on the value of the HOD parameters, which would introduce three extra dimensions. However, for simplicity, we use the parameter  $M_{\min}$  as a proxy for survey depth, and use approximate empirical relations between  $(M_1, \alpha)$  and  $M_{\min}$  to keep the total number of dimensions to four (including WDM particle mass, cf. §6.2.3.4).

To calculate an approximately valid range of  $M_{\min}$  for a given redshift, we use a fiducial stellar mass limit of  $5 \times 10^5 h^{-1} M_{\odot}$ <sup>4</sup> at  $z = 0.05$  and calculate the appropriate stellar mass limit at a new redshift via the luminosity distance (we use the cosmology of Planck Collaboration (2014a) and the `ASTROPY`<sup>5</sup> package (Robitaille et al., 2013) for these calculations). This stellar mass limit is converted to a halo mass limit via the stellar-mass-halo-mass relation of Behroozi, Conroy, and Wechsler (2010), which includes redshift-dependence. The corresponding limits are  $M_{\lim} \sim 10^{9.6} - 10^{12.1} h^{-1} M_{\odot}$  at  $z = (0.05, 5)$ . At each redshift, we cover  $M_{\min}$  in the range  $(M_{\lim}, M_{\lim} + 3)$ .

The redshift range is ultimately limited by galaxy/halo number density, and clearly  $z = 5$  is very high compared to current standards. However, there is sufficient halo number density above  $M_{\lim}$  to expect  $\sim 25,000$  halos in a full hemisphere between  $z \in (4.5, 5.5)$ , and so we optimistically include it.

<sup>4</sup> This is a very low stellar mass limit for current surveys, but future surveys such as WAVES will extend to  $\sim 10^6 h^{-1} M_{\odot}$ , so it is not inconceivable for the next-generation

<sup>5</sup> <http://www.astropy.org>

Finally, as mentioned, we focus on three WDM particle mass candidates: 2.0, 3.3 and 5.0 keV. The latter two are consistent both with current constraints and theoretical expectation from the possible detection of 7.1 keV sterile neutrinos (Bulbul et al., 2014). The first is slightly below current constraints, but will serve as an illustration of the effects of WDM.

### 6.2.3.3 Test Statistic

A robust test would be to derive the marginalised posterior of  $m_{\text{WDM}}$  using Monte Carlo Markov Chain (MCMC) methods, for each point in the parameter space. However, this is infeasibly inefficient, since there are four dimensions ( $M_{\text{min}}, z, r_{p,\text{min}}, m_{\text{WDM}}$ ) over two regimes, and each MCMC run is relatively time-consuming. We thus opt for a simpler test.

Our test effectively accounts for degeneracies between the HOD and WDM effects in a more efficient manner than a full MCMC analysis. At each choice of underlying/survey parameters, we calculate  $w_p(r_p)$  for WDM. We then calculate the best-fit CDM HOD model using simple downhill-gradient methods (we find that the L-BFGS-B method, provided in the Scipy<sup>6</sup> package, performs adequately).

We vary only two parameters,  $M_1$  and  $\alpha$ , with  $M_{\text{min}}$  constrained by the mean galaxy density. The likelihood is then given by the  $\chi^2$  value of the fitted curve, where we use a constant independent uncertainty of 10% for all bins (this value should not affect the value of the best fit, though the assumption of independent uncertainties will). Finally, our test statistic is the maximum relative difference between the best-fit curve and the data.

This statistic takes account of the correlations between HOD and WDM through the downhill-gradient method, but the final value is independent of the precise scale range used (which would not be the case if the final  $\chi^2$  of the best fit model was used).

### 6.2.3.4 Empirical HOD relationship

We have mentioned that we use  $M_{\text{min}}$  as a proxy for survey depth in this analysis. However, we also expect that the value of  $M_1$  for a given sample will move along with  $M_{\text{min}}$ , since a deeper survey should see more satellites as well as centrals. This motivates us to specify a fiducial relationship between the two parameters ( $M_1, \alpha$ ) and our “free” parameter  $M_{\text{min}}$ , so that our fiducial models remain relatively physically appropriate, and we keep the parameter space as small as possible.

The studies of Zehavi et al. (2011) and Skibba et al. (2015) imply a constant relationship of the ratio  $M_1/M_{\text{min}}$ , with a small redshift dependence

$$M_1/M_{\text{min}} = 1.2 - z/10.0, \quad (6.32)$$

and a constant  $\alpha \approx 1$ . It is of interest whether these scaling relations, measured assuming a CDM universe, are valid in WDM. Certainly, one could imagine that  $M_1$  would increase, and  $\alpha$  decrease in WDM due to a dearth of low-mass subhaloes. However, assessing the validity of this reasoning is beyond the scope of this chapter, and would

<sup>6</sup> <http://www.scipy.org/>

require detailed  $N$ -body simulations. We do however show the impact of these shifts on our results in §6.3.2.2. We otherwise fiducially use these CDM relations throughout the paper.

### 6.3 WDM SIGNATURES IN VARYING REGIMES

Our aim is to explore the possibility of WDM signatures in  $\xi_{gg}(r)$ , and determine some measure of their detectability according to the methods in §6.2.3. To this end, we examine each regime individually, identifying the possible signatures, and analysing their origins (particularly their reliability).

#### 6.3.1 Large-scale 2-halo term

The 2-halo term, on large scales, is merely

$$\xi_{gg}^{2h}(r) = b_{\text{eff}}^2 \xi_m(r), \quad (6.33)$$

where

$$b_{\text{eff}} = \frac{1}{n'_g} \int n(m) N_t(m) b(m) dm. \quad (6.34)$$

This reduction occurs because at large scales,  $u(k|m) \rightarrow 1$ , halo exclusion effects are non-existent, and the scale-dependence of the bias (Eq. 6.24) is negligible, since  $\xi_m(r) \ll 1$ .

Thus all effects of the HOD are sequestered into a single parameter (and indeed the effects of WDM on the halo model components). This is why HOD analyses are typically not performed solely on large scales – the parameters are entirely unidentifiable.

Conversely, WDM also affects  $\xi_m(r)$ , and so we may be able to identify features in the ratio of the WDM to CDM matter correlation. We show the correlations and this ratio for our three fiducial WDM candidates, at  $z = 3$ , in figure 6.5. Clearly, the effect of reducing the mass of the WDM particle is to increase the distance between trough and peak of the BAO feature. This effect is amplified with increasing redshift, as we shall see. Interestingly, reducing the WDM mass to very low values ( $\sim 0.1$  keV) begins to *reduce* the distance between the trough and peak, and gives a broadened peak. This corresponds the known effect on the BAO from smoothing the power spectrum at a given scale (see eg. Avila et al., 2015).

An appropriate test in this context then, given that the scale-independent bias can be reproduced by some combination of HOD parameters, is merely to assess the ratio of the peak-trough distance,  $\Delta\xi_{\text{BAO}}$ , in WDM to that in CDM for our fiducial candidates, over a range of redshifts. We show the results of this test in figure 6.6. None of our fiducial models are detectable at the 10% level below  $z \sim 4$ . However, at  $z = 4$ , the extreme-case 2 keV model becomes detectable at the 10% level, and likewise for the 3.3 keV model at  $z = 5$ .

Clearly, robust measurement of a BAO peak in galaxy correlations at  $z \geq 4$  will be a challenging feat, so this signature is far from competitive with more commonly used

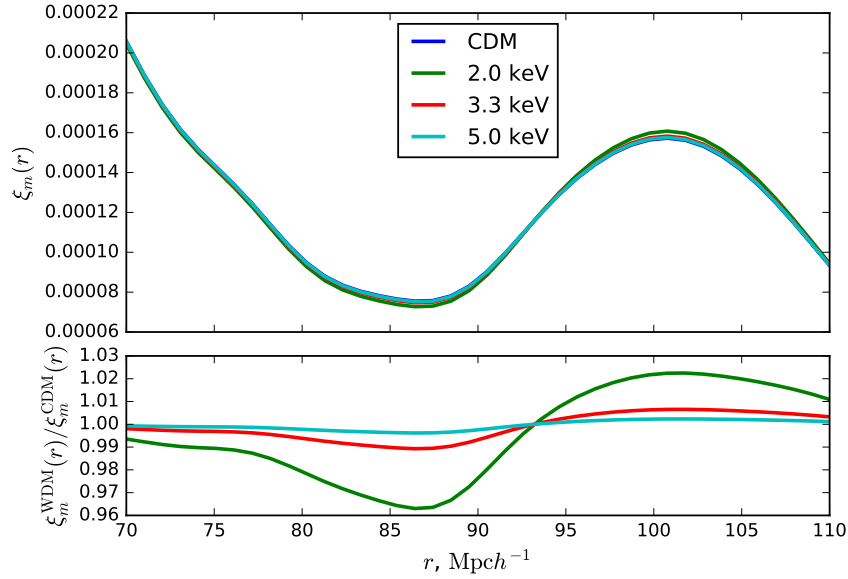


Figure 6.5: Matter correlation functions around the BAO scale for the three fiducial WDM models, and their ratio to CDM at  $z = 3$ . Clearly, decreasing the WDM mass increases the the distance between trough and peak of the BAO feature.

observables (eg. Lyman- $\alpha$  or Milky-Way satellite abundances). It is nevertheless an interesting result worthy of further consideration.

Note however that these results derive from the approximation of Eq. 6.4. As we have mentioned, the use of the WDM-modified nonlinear power from HALOFIT, though tested to some degree (Viel et al., 2012) and found to be appropriate for  $k < 1 h^{-1} \text{ Mpc}$ , is not a fundamentally WDM-motivated form. It remains to be seen if it is accurate on BAO scales when transformed to real-space. Potential future tests of this kind must take this into consideration.

### 6.3.2 Small-scale 1-halo term

In the 1-halo term, there are two components which are directly affected by the WDM model:  $n(m)$  and  $\rho(r|m)$  or  $\lambda(r|m)$  via  $\bar{c}(m)$ . For reasons that we outline below, while both of these will contribute to a potential WDM signature, we expect the dominant contribution to be through the halo profile.

We note that throughout this section we show modelling for  $r$  extending far below  $0.1 h^{-1} \text{ Mpc}$ , which we have previously mentioned is physically tentative at best. However, we find that doing so provides better intuition concerning the effects of WDM, and we restrict our final analysis to physically plausible scales.

In particular, the truncation of the HMF due to WDM particle mass should be mimicked by the HOD parameter  $M_{\min}$ , which controls truncation of the integral via the galaxy abundance. Therefore, we should not expect that the WDM HMF will solely contribute to a detection of WDM in this context, since it will be degenerate with the HOD. This is not the case for the halo profile, however.



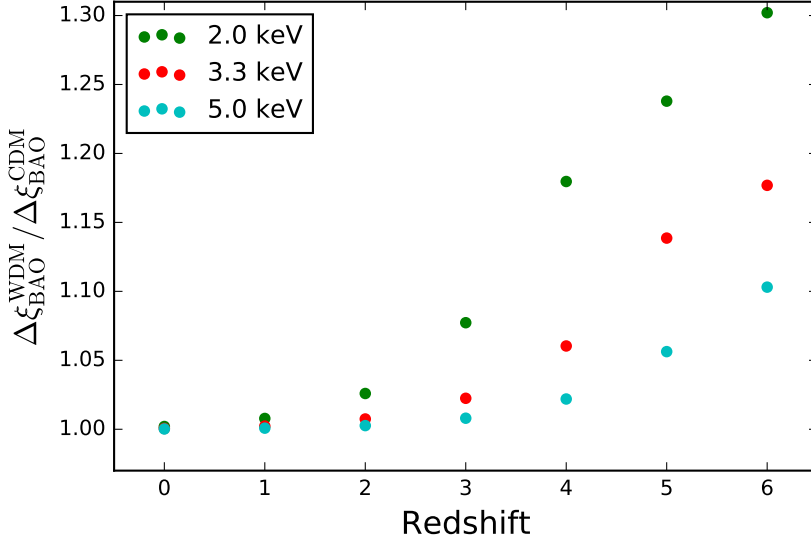


Figure 6.6: The ratio of WDM to CDM peak-trough distances for our fiducial WDM candidates, over a range of redshifts. A measurement at  $z = 5$  could delineate a 3.3keV model to the 10% level.

On average we expect WDM haloes to exhibit lower concentrations than their CDM counterparts, with the largest discrepancy at small halo masses (see §6.2.2.5, and in particular, figure 6.4). If the integrals Eq. 6.8 extend to low enough mass scales, this discrepancy will begin to be non-negligible (in general we might expect that ‘low enough’ refers to scales  $\approx m_{\text{hm}}$ ).

Note that since the turnover for  $\bar{c}(m)$  is an order of magnitude above  $m_{\text{hm}}$ , while the turnover in the HMF is slightly below  $m_{\text{hm}}$ , there is still sufficient number density of haloes present at low mass to reduce the “effective concentration”:

$$c_{\text{eff}} = \frac{1}{n_g} \int n(m) N_t(m) c(m) dm. \quad (6.35)$$

We plot the effective concentration for the HOD defined by our scaling relations (cf. §6.2.3.4), for a range of  $M_{\text{min}}$ , in figure 6.7. It has a turnover much like  $\bar{c}(m)$ , but turns over at small  $M_{\text{min}}$  because of the truncation of the HMF and the change of  $M_1$  with  $M_{\text{min}}$ . We find that this effective concentration becomes the defining characteristic of the very small scale correlations, as outlined below.

### 6.3.2.1 WDM modelling

Figure 6.8 shows the dependence of  $\rho(r)$  (left) and  $\lambda(r)$  (right) on  $m$ , over all  $m > M_{\text{min}}$ , for 4 values of  $r$ , which represent four different regions. In each case, we normalise the curves at  $m = 10^{17} M_{\odot} h^{-1}$ . The high-mass slope with  $m$  for each  $r$  is identical, for both  $\rho$  and  $\lambda$ .

Additionally shown as thick lines, as a visual guide, is the rest of the integrand:  $n(m) N_s^{[2]}(m) N_c(m) m$  (arbitrarily normalised), where the multiplication by  $m$  is to aid interpretation, since we plot in log space, and the squaring of  $N_s(m)$  only occurs in the  $s - s$  term. Ignoring the constant factor of the mean galaxy density in each case, the total



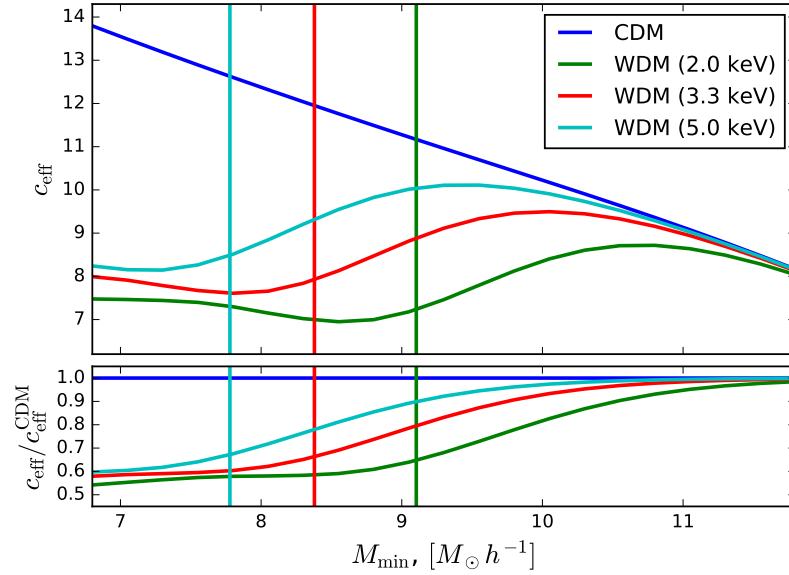


Figure 6.7: The effective concentration (Eq. 6.35) for our fiducial WDM models over a range of  $M_{\min}$  (other HOD parameters kept constant). The shape is similar to that of  $\bar{c}(m)$ , but asymptotes to a constant at low  $M_{\min}$  due to the truncation of the HMF.

integral can be interpreted as the sum of  $\rho(r|m_i)$  or  $\lambda(r|m_i)$  (thin solid lines) weighted by the thick lines. The value of this weighted sum in each case is shown by horizontal dashed lines (increasing amplitude corresponds to decreasing  $r$ ).

Importantly, in both the  $c-s$  and  $s-s$  terms, the highest weight (if valid for a given  $r$ ) is given to the knee of the HMF,  $m \sim M_*$ . The relative strength of this weight will be determined by the HOD slope  $\alpha$  combined with the power-law slope of  $n(m)$ . In principle, the slope of  $n(m)$  cannot be steeper than -2, so  $\alpha \geq 1$  will ensure this effect.

The primary observation to make is that at any distance scale  $r$ , the range of valid  $m$  in each case is well-defined: for the  $c-s$  term by  $m_{\text{vir}}(r)$ , and for the  $s-s$  term by  $m_{\text{vir}}(r/2)$ . The interplay of this scale with  $m_{\text{hm}}$ ,  $M_{\min}$  and  $M_*$  introduces five regions of interest. We consider each of these regions in turn.

At large scales (in the 1-halo term), the entire range of valid  $m$  is at a sufficiently high mass that the concentration difference between WDM and CDM is negligible. The correlation ratio for the same HOD parameters will only differ due to a reduction in mean galaxy density, induced by the HMF turnover. Thus we expect a constant ratio greater than unity between the two (in reality, the mean galaxy density is forced to match, and these large scales become identical, being interpreted primarily as a different value of  $M_{\min}$ ).

As the scale decreases, at some point, the concentration of WDM diverges from CDM (eg.  $r \sim 0.112$  in figure 6.8). Where  $r \sim r_{\text{vir}}(m)$ , this tends to increase the density, but when  $r \ll r_{\text{vir}}(m)$  it has the opposite effect. Firstly we note that the  $s-s$  term should have quite a simple dependence in this region. Since scales at  $M_*$  have such a high relative weight, the marginal effects of the increase in concentration at the smallest valid mass scales for a given  $r$  are completely mitigated, and thus the correlation ratio remains constant. For the  $c-s$  term, there is the possibility that the correlation ratio will marginally

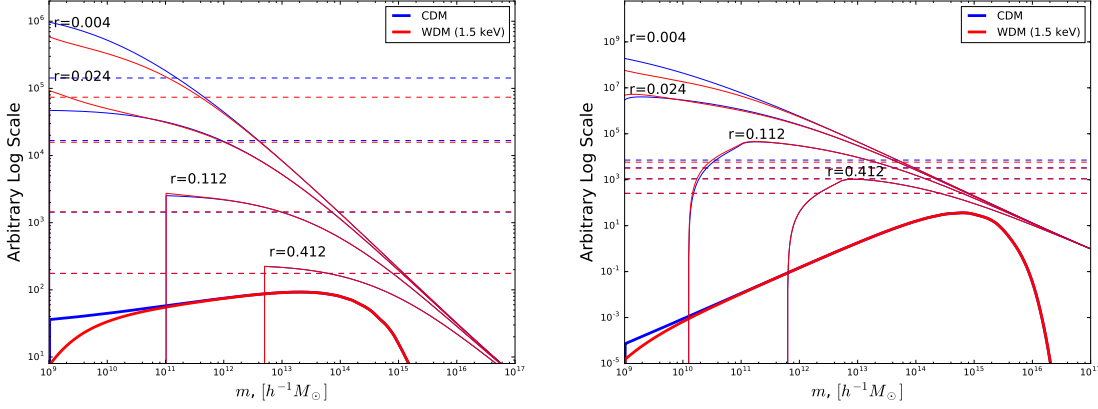


Figure 6.8: The dependence of  $\rho(r)$  (left) and  $\lambda(r)$  (right) on  $m$ , over all  $m > M_{\min}$ , for 4 values of  $r$  as marked. In each case, we normalise the curves at  $m = 10^{17} M_{\odot} h^{-1}$ . The left and right figures illustrate the calculation of the  $c-s$  and  $s-s$  parts of the 1-halo term respectively. Thick lines show the rest of the integrand in each case (renormalised for clarity). The total correlation in each case is proportional to the weighted sum of the thin lines, weighted by the thick lines. The value of the weighted sum is shown by horizontal dashed lines. The HOD parameters used are  $M_{\min} = 10^9 M_{\odot} h^{-1}$ ,  $M_1 = 10^{10.2} M_{\odot} h^{-1}$  and  $\alpha = 1.0$ .

increase through this region, since the relative weight of  $M_{\star}$  scales is reduced. However, this window is small, since at still smaller scales the increase in concentration is countered by a reduction in the weight due to the turnover in the HMF.

As the scale decreases further (eg. the  $r = 0.024$  curve in figure 6.8, which corresponds to a virial mass of exactly  $M_{\min}$ ), though the suppressed concentration of WDM has a larger effect on the density (especially at the virial radius), the corresponding turnover in the HMF completely mitigates the expected effects. Indeed, the first mass scales which have a significant contribution become so large that  $r \ll r_{\text{vir}}(m)$ , and the effect of the reduction in concentration is reversed (this effect is less noticeable in the  $s-s$  term, as it is still very highly weighting yet larger mass scales at which there is no concentration difference).

Beyond this scale, the integral becomes "saturated", ie. the range of valid masses extends below  $M_{\min}$ , which becomes the constant truncation mass. In effect, the decrease we have already described continues, but with a constant threshold. For both terms, an asymptotic limit of this decrease is reached as the inner logarithmic slope of the NFW profile reaches its asymptotic limit of -1 (and the self-convolution reaches a constant value). At this point, the behaviour of both WDM and CDM is identical, though the WDM will have a consistently lower density/correlation (modulo multiplication by the mean galaxy density ratio), set by the effective concentration.

We show a schematic of all five regions in figure 6.9. The general shape of both correlation ratios is an "S"-curve, with plateaus at both scale extremes. The value of the upper plateau is fully specified by the ratio  $(n_g^{\text{CDM}}/n_g^{\text{WDM}})^2$ , whereas the value of the lower plateau is affected both by this ratio and the effective concentration (for the  $s-s$  term, this plateau is much higher, since inner parts of the halo are not probed, but the 1-halo term is dominated by the  $c-s$  term here).

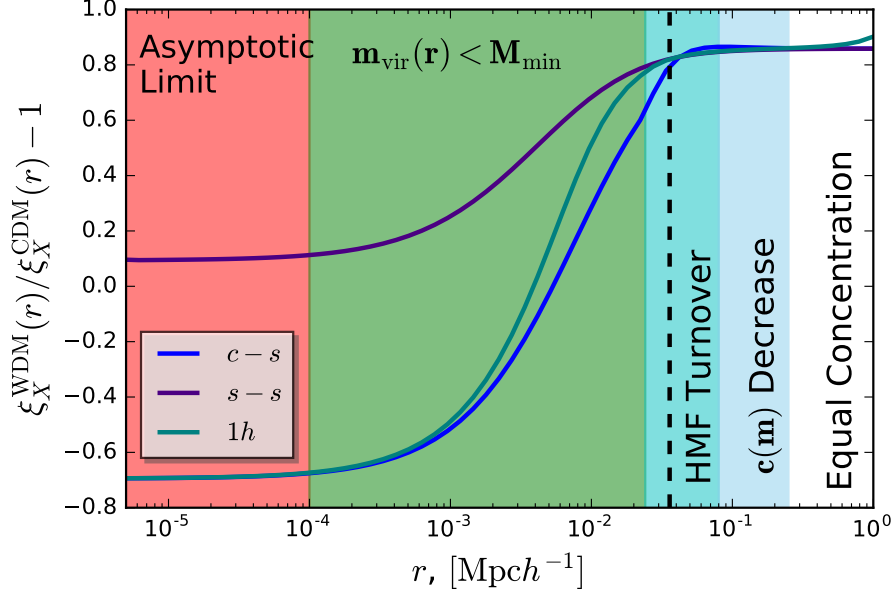


Figure 6.9: Schematic of 1-halo correlations showing the five regions outlined in the text. Red is the region in which the NFW profile reaches an asymptotic limit on its logarithmic slope. Green is the “saturated” region, in which the virial mass of the relevant scale is below  $M_{\text{min}}$ . Cyan is the region in which the effects of the HMF turnover dominate the reduction in concentration. Sky blue is the region in which, for the  $c-s$  term, there is a possibility of the reduction in concentration affecting the correlation ratio. White is the region in which all valid masses have negligibly different concentrations. The dashed vertical line is the virial radius of the half-mode mass. The blue, purple and green curves are the  $c-s$ ,  $s-s$  and total 1-halo correlation ratios respectively.

Between these plateaus is a somewhat smooth interpolation, however in the case of the  $c-s$  term, the drop is sharper, due to the increased effects of the HMF turnover. Furthermore, there is a discontinuity in the first derivative at  $r = r_{\text{vir}}(M_{\text{min}})$ .

In summary, the effect of WDM on the 1-halo term is, assuming a matched mean density, to reduce the correlations on small scales (as long as  $M_{\text{min}} < \sim 10m_{\text{hm}}$ ), which is expected. The point of departure from CDM is approximately the half-mode mass (shown as the dashed vertical line in figure 6.9). For a minimum  $r$  of  $0.1h^{-1}\text{Mpc}$ , it is clear that the differences between WDM and CDM will be very small.

### 6.3.2.2 Relation to effects of HOD

We would also like to understand the interplay between the WDM effects and the HOD parameters, not least because we may expect the underlying HOD parameters to be different in a WDM universe. For illustrative purposes, we plot an example of each of these cases in figure 6.10. In this figure, we take as a “fiducial” model a WDM mass of 1.5 keV and a HOD with parameters  $(\log_{10} \frac{M_{\text{min}}}{M_{\odot} h^{-1}}, \log_{10} \frac{M_1}{M_{\odot} h^{-1}}, \alpha) = (9, 10.2, 1.0)$ . In each alternative case we change one of the HOD parameters in both CDM and WDM, and match the mean density by adjusting  $M_{\text{min}}$ . We summarise the effects in this section.

Firstly, we note that at “large” scales (in the context of the 1-halo term), the correlation ratio is unity in all cases, which is simply due to the matching of mean densities. At the smallest scales, all ratios tend to be less than unity, as expected from the WDM suppres-

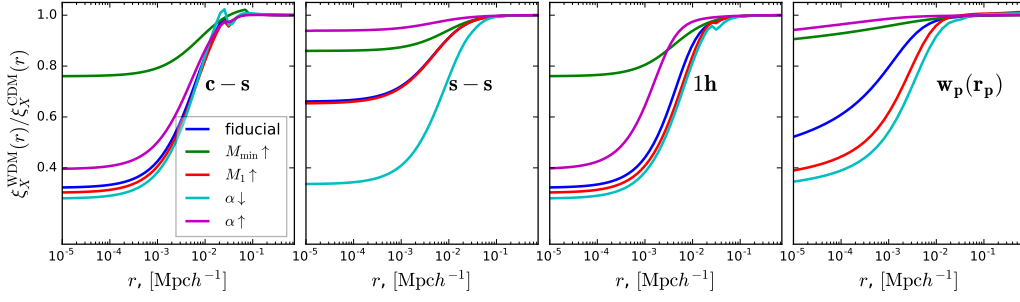


Figure 6.10: Dependence of the correlation ratio with  $m_{\text{WDM}} = 1.5$  keV on HOD parameters. From left to right panels are the  $c-s$  term,  $s-s$  term, full 1-halo term and projected correlations. The typical “S” curve is preserved in all cases, though the amplitude of the small-scale limit is modified.

sion of small-scale structure. We are interested in the effects of the HOD parameters on the magnitude and scale of this suppression.

Secondly, we note that the last panel denotes our quantity of interest: the projected correlation function (Eq. 6.31). As this is an integral, all scales  $r > r_p$  contribute to the value. This is especially noticeable in the purple curve, which shows the effect of increasing  $\alpha$ . The 1-halo term has a significant suppression at small scales, but since the drop is at quite a low scale, the projected correlation has no significant suppression at any scale.

The most dramatic effect is from  $M_{\text{min}}$  itself. While it has a somewhat complicated effect in detail, clearly raising  $M_{\text{min}}$  increases the effective concentration (see figure 6.7), and therefore inhibits suppression. In the case of the  $c-s$  term, it also changes the position of the discontinuity, but this doesn’t affect the final 1-halo term a great deal.

In principle, an increase in  $M_1$  for a given  $\alpha$  and mean density just increases the ratio of centrals to satellites. However, in WDM, the scales at which centrals dominate are suppressed, and so we see a modest enhancement of the suppression by increasing  $M_1$ . This has a weaker effect on  $s-s$ .

Increasing  $\alpha$  tilts the balance of satellite density to higher mass haloes, at which the suppression due to WDM is negligible, so we see an inhibition of suppression for increased  $\alpha$ . This effect is enhanced in the  $s-s$  term due to the squaring of  $N_s(m)$ . Additionally, lowering  $\alpha$  increases the effects of both the “ $c(m)$  decrease” and “HMF Turnover” regions, which causes the visible discontinuities in the  $c-s$  term.

In summary, the largest differences are to be found at the smallest scales, which are generally beyond the limit of spatial resolution of galaxy surveys. The extent of the disparity is entirely set by the small-scale limit of the  $c-s$  term, which depends on  $c_{\text{eff}}$ . The scale at which the transition begins (as scales are probed smaller and smaller) however, is affected also by the small-scale limit of the  $s-s$  term. Thus, a population with higher  $M_1/M_{\text{min}}$  or lower  $\alpha$  will be beneficial for WDM constraints. Interestingly, this is the direction we would expect these parameters to move in a WDM universe.

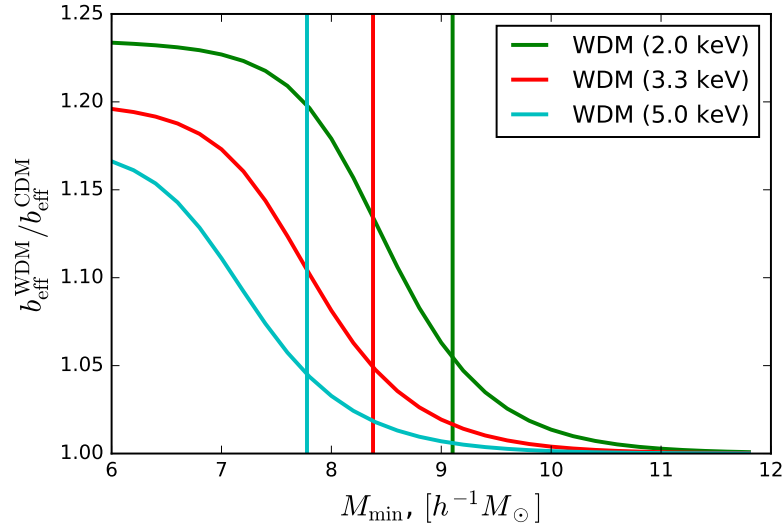


Figure 6.11: Effective bias ratios of the fiducial WDM candidates, as a function of  $M_{\min}$ . In each case, the  $M_{\min}$  value in CDM is chosen to match the mean density, and we use our fiducial relations for the HOD parameters with  $M_{\min}$ .

### 6.3.3 Transition scales

Possibly the most difficult regime to accurately model, the transition scales ( $\sim 1 - 4 h^{-1}\text{Mpc}$ ) are characterised by the crossover of 1-halo and 2-halo terms, and are greatly affected by the chosen method of halo exclusion (cf. §6.2.1.1). It is these scales that introduce a characteristic deviation from power-law behaviour in the galaxy correlations (Zehavi et al., 2004).

In terms of the behaviour of WDM compared to CDM, we have already been able to characterise both scale extremes of the region: in 1-halo-dominant parts of the transition, given we can match the mean galaxy density, the correlation ratio is very close to unity; on large scales, where halo exclusion is no longer effective, the correlation ratio is

$$\frac{\xi_{gg}^{\text{WDM}}(r)}{\xi_{gg}^{\text{CDM}}(r)} = \frac{b_{\text{eff, WDM}}^2 \xi_m^{\text{WDM}}(r)}{b_{\text{eff, CDM}}^2 \xi_m^{\text{CDM}}(r)}. \quad (6.36)$$

Since the matter correlation ratio should be near unity at large scales (except for around the BAO peak), this reduces to the ratio of effective biases. This quantity need not be unity when the mean density is, since the bias function is defined in terms of the mass variance, which is suppressed in WDM.

We show the effective bias ratio for our three fiducial candidates, as a function of  $M_{\min}$ , in figure 6.11. In each case, the  $M_{\min}$  value in CDM is chosen to match the mean density, and we use our fiducial relations for the HOD parameters with  $M_{\min}$ . We note that  $b_{\text{eff}}$  has less than 5% deviation from CDM for  $M_{\min} > m_{\text{hm}}$ , so we expect reasonable samples to have very similar effective bias.

Nevertheless, even if this is not the case (eg. if  $m_{\text{WDM}}$  happened to be small, and we had a very deep survey), we can easily match  $b_{\text{eff}}$  by adjusting the two remaining free HOD parameters. Thus, if we assume that the final correlation ratio will be a smooth interpolation between two scale extrema, we have reason to believe that it will be very

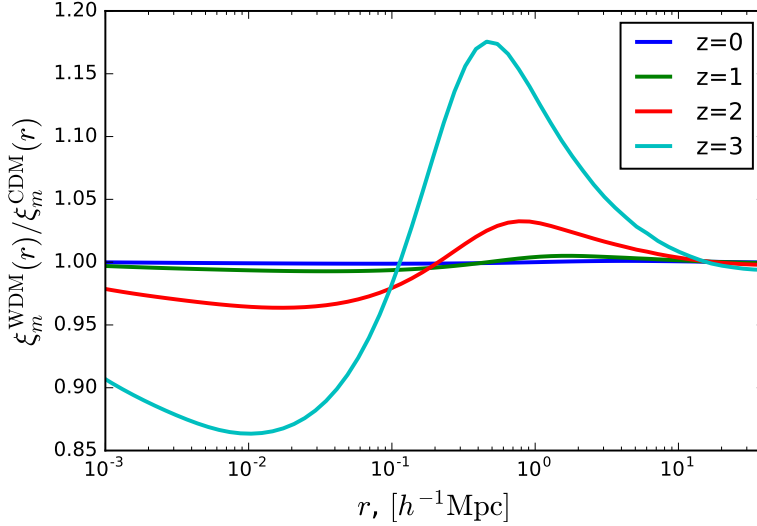


Figure 6.12: Ratios of WDM to CDM matter correlation functions (derived from HALOFIT) over small-to-medium scales, for redshifts up to  $z = 3$ . WDM exhibits excess correlation around the scale of the transition, and suppressed correlation at smaller scales. The peak excess moves to smaller scales with increasing redshift.

close to unity on all scales, once appropriate values of the HOD are chosen. Indeed, even having set two of the HOD parameters in this manner, the third is still free to absorb some of the residual.

However, at scales small enough to be truly within the transition, the matter correlation ratio can exhibit deviations from unity. This is especially true at higher redshift, where gravitational nonlinearities have not yet restored small-scale power to the WDM density field. We plot examples of this ratio for redshifts up to  $z = 3$  in figure 6.12, and note that WDM has excess correlation around the scale of the transition, and suppressed correlation at smaller scales. Also, the peak excess moves to smaller scales with increasing redshift. We stress again that using  $\xi_m(r)$  from HALOFIT in the context of WDM may be inaccurate, and further modelling with more robust and resolved  $N$ -body simulations will be necessary before concrete conclusions may be drawn.

Besides the effects of the matter correlation, there are two other effects which must be considered: that of the scale-dependence of the bias, and halo exclusion. Though these are important for modelling of the total galaxy correlation, we find they have very little effect on the WDM to CDM correlation ratio. The scale-dependence of the bias becomes important for scales below the transition region, in which the dominant term is the 1-halo term, so they have little effect. Furthermore, the scale dependence is modelled identically in both WDM and CDM, with differences arising only through the relative value of  $1 + \xi_m(r)$ , which is very close to unity throughout the 2-halo dominance.

The halo exclusion has much the same limitation – it is modelled identically between WDM and CDM, so differences that arise are minimal. We show the extent of both of these effects at  $z = 3$  in figure 6.13, and note that they are negligible in comparison to the effects of the matter correlation ratio.

Despite the expectation that the galaxy correlation ratio should be unity in the transition other than a difficult-to-model contribution from the mid-scale matter correlation function, including the transition and mid-scale regimes into our analysis has the po-

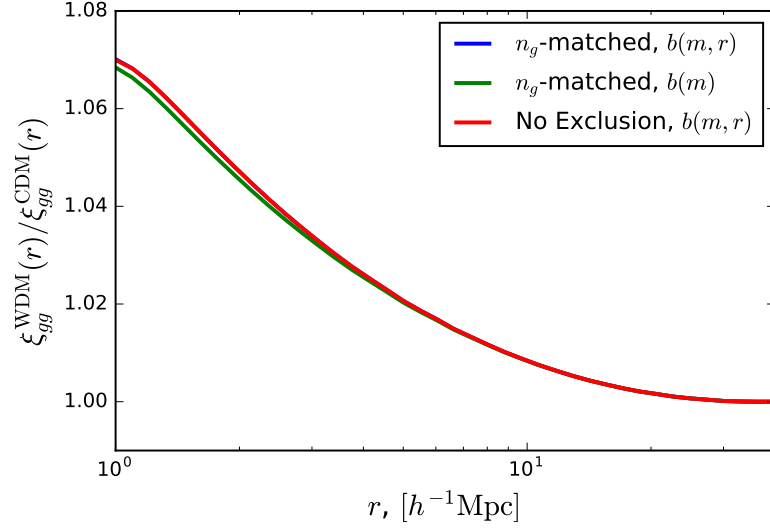


Figure 6.13: Galaxy correlation ratio at  $z = 3$  and  $M_{\min}^{\text{WDM}} = 10.5$ , for a fiducial model with  $n_g$ -matched exclusion and scale-dependent bias, and two variants in which each of these are relaxed: to scale-independent bias, and no halo exclusion modelling respectively. The broad deviation from unity comes primarily from the matter correlation ratio, and the adjustment to this deviation from the small-scale modelling is negligible.

tential advantage that an extra parameter may be constrained by the large-scale bias, effectively leaving only one parameter to absorb the residual at small scales, which may improve model differentiation.

## 6.4 RESULTS

### 6.4.1 Model I: NFW

The results of our test statistic over the entire parameter space outlined in §6.2, using the fiducial NFW model, are shown in figure 6.14. In this figure, colours represent particle mass, and are the same as figure 6.11, and marker type indicates redshift. The  $x$ -axis is our proxy for survey depth,  $M_{\min}$ , and the left and right panels are the “pure” 1-halo term and 1-halo+transition regions respectively<sup>7</sup>.

As expected, there is clearly an inverse relationship between the maximum deviation and both  $m_{\text{WDM}}$  and  $M_{\min}$ . The effect of including the transition region is relatively small, especially for  $M_{\min} \leq 10^{12.5} h^{-1} M_{\odot}$ . This is because the effective bias ratio is very close to unity when the galaxy density is matched, and certainly its deviation from unity is dominated by small-scale effects.

This point may be somewhat misleading, as we find that the maximum deviation in many cases arises from within the transition region. The limits of this region change with redshift and  $M_{\min}$ , and so often our upper limit of  $1 h^{-1} \text{Mpc}$  for the 1-halo term is far too large, and transition effects appear in the “pure” 1-halo term. These transition

<sup>7</sup> The distinction is somewhat vague, especially as the redshift and HOD parameters are changed, so we arbitrarily choose a cut at  $r = 1 h^{-1} \text{Mpc}$  as elsewhere in this work.



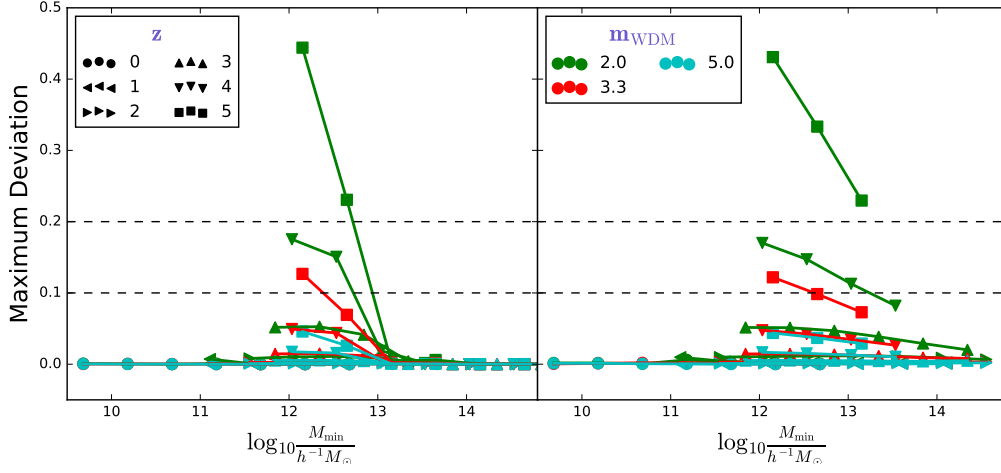


Figure 6.14: Maximum deviation from WDM of best-fit CDM models with matched mean galaxy density, over our four-dimensional parameter space, using the DM-tracing NFW profile. Left and right panels are results for scales  $r_{p,\min} < r_p < 1h^{-1}\text{Mpc}$  and  $r_{p,\min} < r_p < 40.0h^{-1}\text{Mpc}$  respectively (i.e. “pure” 1-halo and 1-halo + transition). Colours represent WDM particle mass, with the same scheme as in other plots. Markers represent redshift and the  $x$ -axis is the “survey depth”, identified by  $M_{\min}$ . Horizontal dashed lines indicate 10 and 20% deviations.

effects are predominantly affected, at least at high redshift, by the matter correlation ratio, as discussed in §6.3.3.

Unfortunately, as we have mentioned, the reliability of the nonlinear WDM correlation function on these scales is unclear. In figure 6.15, we show a very similar set of results as in figure 6.14, but where the test statistic is the deviation at the minimum scale. This is the most conservative estimate of the maximum deviation available. The effects of the nonlinear WDM correlation function are still present to some degree through the best-fit CDM model, but the dominant effect is the suppression of structure on small scales. The results are similar in form to our default test statistics, but significantly reduced in amplitude (by a factor of  $\sim 10$ ).

In summary, for the base NFW model, dependent on the reliability of the HALOFIT-corrected matter correlations in WDM, the dominant signature of WDM on the projected correlation function may be an *excess* at scales within the 1-halo – 2-halo transition (for a given redshift and  $M_{\min}$ ). If this modelling is quite inaccurate, then the dominant signature is a slight suppression at the smallest resolved scales.

#### 6.4.2 Model II: Einasto

We now turn to a galaxy abundance profile which differs from the underlying dark matter – parameterised by our second model, the Einasto profile. For our fiducial model, we use  $\alpha = 0.678$ , corresponding to a suppression of abundance towards the centre of haloes compared to the underlying field. However, we also make a change to our analysis in this case, leaving  $\alpha$  free (normal prior with mean 0.678 and standard deviation 0.1) when finding a best-fit CDM model. We show the results of this analysis in figures 6.16 and 6.17 for the maximum deviation and conservative deviation at  $r_{r,\min}$  respectively.



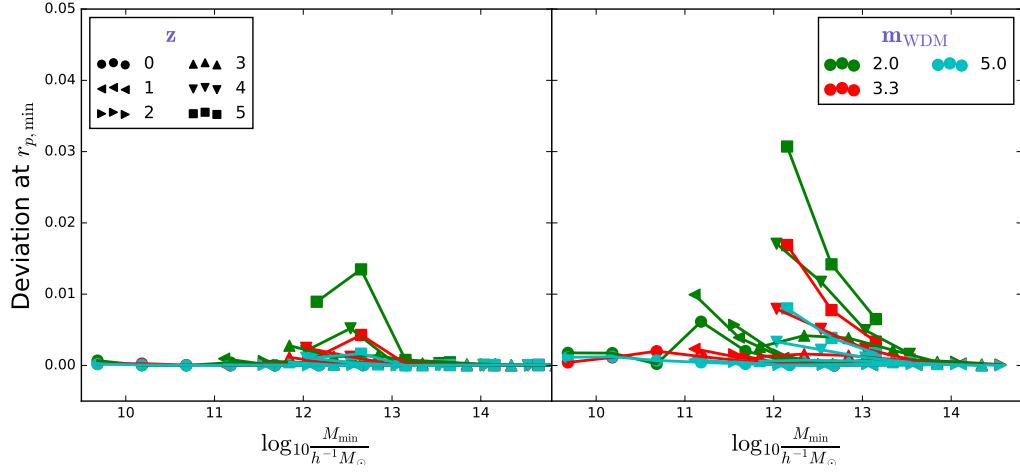


Figure 6.15: Same as figure 6.14, but replacing the maximum deviation with the deviation at  $r_{p,min}$ .

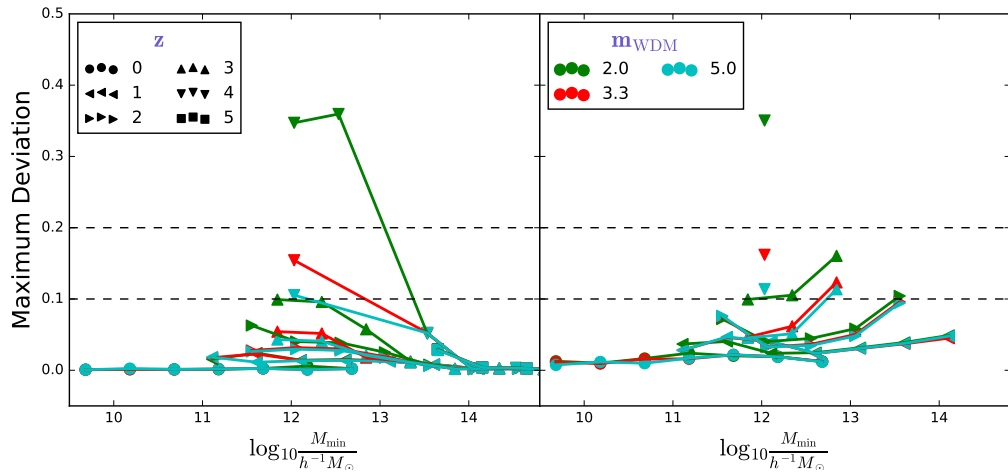


Figure 6.16: The same plot as 6.14, but using an Einasto profile for galaxy abundance derived from the Aquarius simulations.

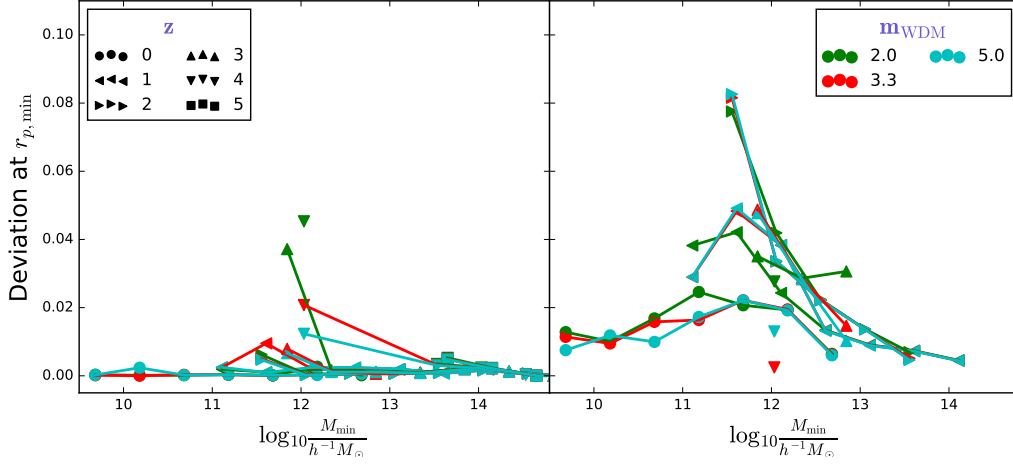


Figure 6.17: The same plot as 6.15, but using an Einasto profile for galaxy abundance derived from the Aquarius simulations.

With respect to the DM-tracing NFW profile, altering the galaxy distribution tends to accentuate the difference between WDM and CDM. On small scales, this results in the ability to differentiate at the 10% level for a 3.3keV candidate at  $z = 4$ , rather than requiring  $z = 5$  as in the NFW model. Otherwise, the trends are very much the same.

In the more conservative case, results are again significantly reduced in magnitude, however they are a factor of  $\sim 2$  more differentiated than in the NFW model.

We note that due to numerical issues in Fourier-transforming the Einasto profile (for which no analytic transform exists), the results for this model are less robust. Consequently some of the analysed points in our parameter space did not converge to a solution, and are thus not displayed. This issue presents a major obstacle for widespread usage of the Einasto profile in halo model calculations.

The amplitude of the deviations, in either case, is typically well below 10%, which will remain undetectable for many years. This is only overcome at high redshift, where the corresponding measurement error will also be larger. Thus there is limited opportunity with the galaxy correlation function to detect the signatures we have identified.

#### 6.4.3 Analysis of WAVES

The WAVES (Wide Area VISTA Extra-galactic Survey; Driver et al., 2015) survey, due to commence in 2021, is a large-scale survey focussed on galaxy evolution and the nature of dark matter. It will observe  $\sim 10^6$  galaxies out to  $z \approx 0.2$ , for the first time probing representative samples of Milky-Way-mass haloes ( $\sim 10^{12} h^{-1} M_\odot$ ). This depth makes it an ideal survey for the type of analysis we have outlined in this chapter.

WAVES-Wide will cover about 1350 square degrees, which equates to a volume of  $V = 0.0255 h^{-3} \text{Gpc}^3$ . With this volume, we can define an expected value of  $M_{\min}$  for the survey via

$$V \int_{M_{\min}} n(m) N_t(m) dm = 10^6, \quad (6.37)$$

where we again employ our scaling relations for the HOD parameters to calculate  $N_i(m)$ . We find  $M_{\min} = 10^{11.14} h^{-1} M_{\odot}$ . Using  $r_{p,\min} = 0.1 h^{-1} \text{Mpc}$  as we have done throughout (and which roughly corresponds to the mean projected separation of galaxies in the survey), we perform the same analysis as in the previous section, and find that the maximum deviation for the three (increasing-mass) WDM candidates is (0.56, 0.14, 0.048)% for the NFW model.

Setting the profile to follow substructure, and relaxing the substructure profile parameter to be fit within the Einasto model, we find (2.33, 1.56, 1.53)%.

Again, we find that allowing the substructure profile to vary with respect to the underlying field enhances the effects of WDM. However, the differences are very small – on the order of 1%. Given the systematic uncertainties inherent in such a survey analysis, this is well below any reasonable detection threshold.

## 6.5 CONCLUSIONS

In this study we have performed the first detailed study of the effects of WDM on the galaxy correlation function using HOD models. To do so we have relied largely on the halo model of Smith and Marković (2011) in conjunction with components from Benson et al. (2012) and Schneider, Smith, and Reed (2013). We have focussed on three WDM particle candidates which roughly coincide with expectations from a recent 3.5 keV X-ray line discovery (Bulbul et al., 2014), and its effects on three different scale regimes and varying redshift.

The major disadvantage of using luminous galaxies as opposed to a statistic that depends solely on the DM distribution (eg. weak lensing) is that there is an inefficiency of star formation in low-mass haloes (Behroozi, Conroy, and Wechsler, 2010), thus limiting the observed mass scales to those where the effects of WDM are weakest. On the other hand, galaxy surveys are currently ubiquitous, and will be expanded by an order of magnitude in the decade to come (Ivezic et al., 2014), making them the ideal testbed if detection is at all possible.

### 6.5.1 Summary of results

Given the physical restrictions of our approach, we do not expect that any signatures will be easy to detect. However, we summarise our results as follows.

ON LARGE SCALES, the clearest signature is at the BAO peak, which is affected by the smoothing of the power spectrum on small scales. Independent of the halo modelling and halo occupation, the distance between trough and peak of the BAO feature is modified by WDM. The extent of the effect is however quite small, and a 10% difference for a 2keV candidate requires a measurement at  $z \geq 4$ , and for the more interesting 3.3 keV candidate,  $z \geq 5$ .

IN THE 1-HALO TERM, we found that large scales ( $r \sim 1 h^{-1} \text{Mpc}$ ) are completely unaffected by WDM for reasonable candidates, since the masses probed are too large.

Very small scales see a constant suppression of structure: constant due to the asymptotic logarithmic limit of the NFW profile, and a suppression due to the difference in effective concentration. Intervening scales are a smooth interpolation between the small and large scale limits, with the highest transition scale strongly affected by the small-scale limit of the satellite-satellite term.

THE TRANSITION REGION, while difficult to model absolutely, has negligible modelling effects from halo exclusion and scale-dependent bias on the WDM to CDM correlation ratio. We find that the ratio is unity at both scale extrema for the region if both the mean galaxy density and effective bias are matched. This is always possible, and effectively leaves (at least) one free HOD parameter. In practice, the effective bias ratio is so close to unity that this behaviour does not aid much in resolving degeneracies. The matter correlation derived from HALOFIT does exhibit some noticeable differences on the small-scales of this region, via the 2-halo term, at high redshift.

THE DOMINANT SIGNATURE is dependent on the accuracy of the HALOFIT-derived matter correlation function in WDM. If it is relatively reliable, then at high-redshifts, the dominant signature is an excess correlation around the scales of transition. If so, deep surveys at  $z \geq 4$  are able to differentiate 2.0 keV models at the 15% level, and 3.3 keV models at the 10% level. If not, then the dominant signature is a suppression of structure at the smallest resolved scales. In this case, no candidate is able to be resolved at any redshift above the 5% level, due primarily to angular resolution limitations.

USING MODIFIED PROFILES to account for a different distribution of galaxies with respect to the underlying dark matter within haloes tends to enhance the effects of WDM, though it is a numerically unstable calculation. In this case, if one trusts the accuracy of the HALOFIT corrections in WDM, one needs only to reach  $z = 3$  to gain a 10% discrepancy for the 2.0 keV candidate, while in the more conservative assumption, still no candidate is able to be resolved to the 10% level.

### 6.5.2 Future work

Our results depend on the best analytic modelling available in the context of WDM at this time. However, there are several components that may benefit from more extensive modelling with more robust simulations.

HALO-CENTRE POWER SPECTRUM. We have remarked several times that the results both at BAO scales and  $r \sim 1h^{-1}\text{Mpc}$  are most dependent on the choice of halo-centre power spectrum. In CDM, the nonlinear fitted correction from Smith et al. (2003, i.e. HALOFIT) is typically used. This choice is also common in WDM modelling (eg. Schneider, Smith, and Reed, 2013), but it is unclear whether this is appropriate on small scales (for a limited discussion see Viel et al. 2012, and for a comparison of the halo model presented here to a naive application of HALOFIT see Smith and Markovič (2011)). A

more finely tuned correction in the context of WDM may be a beneficial tool in future analyses.

**HALO PROFILE.** We have employed the Einasto profile as a means to marginalise over the effects of galaxy abundance not tracing the underlying dark matter distribution. It is however unclear whether the value of  $\alpha$  we have fiducially set is appropriate in the context of WDM, and certainly highly-resolved WDM simulations may help to answer this question. Furthermore, an advancement in terms of numerical handling of the Einasto profile would be highly advantageous if it is to be widely used in the future.

**CONCENTRATION-MASS RELATION.** While the relation of Ludlow et al. (2016) is the most sophisticated relation to date, especially when considering WDM haloes, it is specifically designed for dark matter concentrations. In the context of galaxy-galaxy correlations, it would be helpful to have a similar model defined for the radial galaxy (or substructure) abundance with which a completely self-consistent halo model of galaxy positions could be created. Such a work would however be highly non-trivial.

**HALO EXCLUSION METHOD.** We found that the halo exclusion method did not greatly affect our results, however this was primarily because the method was applied identically in WDM and CDM. The  $n_g$ -matched methodology empirically accounts for the distribution of halo triaxiality, and some results suggest that halos may be less triaxial in WDM (Bullock, 2001). To accurately model the transition region, this may require recalibration.

These various improvements require highly resolved and robust WDM simulations, which have been notoriously difficult to produce due to discreteness effects (Wang and White, 2007). Several attempts have been made to post-process results in order to clean the ubiquitous spurious haloes (Schneider et al., 2012; Lovell et al., 2014), or to modify the force softening (Angulo, Hahn, and Abel, 2013) to prohibit their formation. One promising development is a novel adaptive softening (Hobbs et al., 2015), a technique that may allow the kinds of robust studies necessary to specify the semi-analytic models accurately.

We have also mentioned that our test statistic is quite simplistic – a more thorough future analysis will use full MCMC to derive robust parameter constraints. Furthermore, it may be beneficial to include several observables simultaneously (eg. Leauthaud et al., 2012; Reddick et al., 2014) to break degeneracies between the HOD and WDM particle mass. We leave such analyses to future studies.

In summary, the prospects of constraining the WDM particle mass via galaxy clustering are slim, as we may have expected. There are however at least two signatures which may be probed with future high-redshift surveys, and in conjunction with other probes it may form a useful independent contribution. Furthermore, understanding the effects of WDM on galaxy clustering properties is a valuable exercise in itself, and may be useful for further studies of damped perturbation models such as mixed or self-interacting DM.



---

## SUMMARY AND CONCLUSIONS

---

### 7.1 SUMMARY

This thesis has presented a statistical framework for the design and exploitation of next generation galaxy surveys. It is comprised of two major new codes (chapters 2 and 5), which utilise our current best theory for cosmological structure formation to model the dark matter distribution and the halo model to capture the distribution of galaxies. In addition, we present a novel and robust statistical method to apply the halo mass function to arbitrarily uncertain measurements using a new functional form called the MRP (chapter 4), as well as several tailored applications of these tools in chapters 3, 4 and 6.

A brief summary of the two codes is as follows.

`HMFCALC`, presented in chapter 2, is an online tool dedicated to simple and efficient calculation of the halo mass function (HMF). The HMF is ubiquitous in studies of structure formation, important directly in cluster cosmology analyses, but also necessary for analytic halo models of galaxy clustering. Its wide applicability and the broad cross-section of potential users make it especially suited to implementation as a web-application, easily accessible and intuitive.

The backend implementation, `HMF`, is a Python package whose main design goals are to provide a comprehensive, intuitive and extendible framework for the efficient calculation of the HMF and other related quantities. It achieves this by including a large fraction of the HMF models from the literature (limited to those within the Press-Schechter framework). These models are simple to switch between, their parameters can be modified (or fit!), and custom models can easily be injected without touching the source code. The inclusion also of several different models for the transfer function, including on-the-fly calculations with the `CAMB` Boltzmann code (and an interface for adding more), along with out-of-the-box support for warm dark matter (WDM) models, makes it the most comprehensive and authoritative freely available HMF code.

The web-application `HMFCALC` exposes a great deal of the functionality of `HMF` within a visually intuitive interface, and enables the quick download of data files containing many relevant quantities, along with automatically produced figures and meta-data.

Along with the presentation of these codes, chapter 2 contains a description of the relevant theory (both of the HMF and WDM modifications), and also two simple worked examples: firstly a calculation of the simulation box size required to expect a single halo

heavier than  $m$ , useful for simulators planning simulations targeted at a specific mass range of halos. And secondly, an exploration of the effects of finite simulation box size and resolution, which can alter the amplitude and shape of the measured HMF.

`HALOMOD`, presented in chapter 5, is a Python package which builds upon the HMF framework and extends its application to the full analytic halo model. This framework has been the subject of much attention in the past decade, frequently used to interpret the galaxy clustering (and other statistics, such as weak lensing) in galaxy redshift surveys. In conjunction with a halo occupation distribution (HOD) model, it provides an efficient way to model galaxy bias, and thus test both the dark and light sides of structure formation.

The underlying HMF code provides a couple of the required components of the halo model (transfer function and HMF), and `HALOMOD` introduces the halo bias, halo profiles, concentration-mass-redshift relation and halo exclusion components. Importantly these components are implemented with a well-defined structure, rendering them easily interchangeable and extendible. The components are brought together in frameworks that facilitate the calculation of spatial statistics. These frameworks also have a well-defined structure, enabling lazy-loading of desired quantities, and an efficiency-boosting caching mechanism.

The most common task in halo model studies is to fit HOD parameters to observed correlation functions (or other spatial statistics). This functionality is included in `HALOMOD` (and HMF) via a very flexible command-line script, through which any included quantity of `HALOMOD` may be fit, with any parameters kept free. Indeed, we use this functionality to provide a worked example in which a correlation function measured in 6dFGS provides fits to ever-expanding sets of free parameters, including the HOD, cosmology, and empirical parameters of the various components.

The third tool presented in this thesis was a new analytic formula to efficiently characterise the HMF directly as a function of mass. The new formula, which we called MRP, is in essence a truncated generalised gamma distribution (also being a generalisation of the popular Schechter function), and for good choices of its three parameters, is in excellent agreement with both Press-Schechter-derived and simulated HMFs. In fact, we showed that the deviation between MRP and standard HMFs is less than the deviation introduced by choices of halo-finding technique or uncertainties involved with including baryons.

We presented various helpful properties of our form, including statistical moments, useful re-parameterisations and an idealised analytic model for its parameter covariance. We focussed on presenting the machinery required to use the MRP form to fit halo catalogues. Simulated catalogues are simple, but observed catalogues provide a significant challenge in terms of measurement error. We tackle these problems in terms of a hierarchical Bayesian model with thousands of nuisance parameters, which we showed performs remarkably well even with mass measurement uncertainties that span  $1/4$  of the logarithmic mass range of the sample. We provide a simple `STAN` model for the general implementation of this method.



Using the MRP, we also explored two applications. Firstly, we derived fitting functions for the dependence of the MRP on the physical parameters  $z$ ,  $\Omega_m$  and  $\sigma_8$ , with the resulting parameterisations yielding  $\sim 20\%$ -level accuracy over 4 decades of mass for redshifts extending past the epoch of reionisation ( $z = 10$ ). Secondly, based on the analytic nature of MRP, we derived a precise parameterisation of the stellar-mass-halo-mass (SMHM) relation, improving the accuracy over previous models by  $\sim 50\%$  for the same number of parameters.

As applications of these tools, we focused statistical questions, especially pertaining to the nature of dark matter.

With HMF/HMFCALC, in chapter 3, we asked the question “How well do we know the halo mass function?”. The answer traced the improvement of uncertainty in the analytic HMF given uncertainty in cosmological parameters over the past decade of CMB mission results. We found that the key cosmological parameter that will influence further precision is  $\sigma_8$ , with  $\Omega_m$  also playing a major role. Interestingly, we concluded that the current estimated uncertainty of  $\sim 6\%$  at low mass and  $\sim 20\%$  at high mass is comparable to the uncertainty introduced by arbitrary choice of fitting function from the literature. This motivates the continued modelling of analytic fits for improved constraints in the future.

We also performed a rudimentary examination of prospects for delineating dark matter models (i.e. CDM from WDM and WDM-like models) with the HMF, finding that detection of reasonable candidates at masses reachable by future surveys such as WAVES-Wide is only possible with a  $\sim 300\%$  increase in precision on our knowledge of  $\sigma_8$ . We concluded that in the context of the HMF, surveys that are specifically targeted at the smallest halos are the most likely to provide insights into the nature of dark matter.

In chapter 6, using HALOMOD, we again addressed the problem of WDM, this time by exploring expected signatures in the galaxy correlation function. Using recent extensions to the halo model framework to characterise the effects of WDM, we examined three regions: very small scales dominated by intra-halo galaxy pairs, very large scales dominated by inter-halo pairs, and transition scales. On the large scales, the clearest signature arises at the baryon acoustic oscillation (BAO) peak, whose trough-to-peak amplitude is modified by WDM. This is independent of the galaxy bias, which on these scales is largely scale-independent. On smaller scales, the same is not true, and the effects of WDM can be correlated with effects from the HOD parameters. We account for this by fitting the HOD parameters in each case, and taking the residual between the fitted curve and the fiducial WDM case. Using this methodology, for intra-halo pairs, we found that only the very smallest scales exhibit effects from WDM, being suppressed by a constant factor, dependent on the “effective concentration”. The transition region is primarily affected at small scales by the power spectrum of halo centres, which is involved in the inter-halo pairs. We conclude that the overall dominant signature is dependent on the accuracy of the modelling of this halo-centre power spectrum. If reliable, then the dominant signature is an excess correlation around the scales of transition, especially at high redshift, resulting in a potential differentiation of a 2 keV model at  $z \geq 4$  at the

20% level and a 3.3 keV model at the 10% level. If not, then the dominant signature is a suppression of the correlations at the smallest resolvable scales, leading to a 10% deviation for the 2 keV model at  $z \geq 4$  with an angular resolution capable of resolving separations of  $r = 0.01h^{-1}\text{Mpc}$  at  $z = 0.05$ .

## 7.2 CONCLUSIONS

### 7.2.1 Next-generation tools

Traditionally, astronomy has had a penchant for large and complex codes written in languages only matched in their efficiency by their opacity. The reason for this is clear enough; astronomy typically engenders large datasets and involved mathematical procedures, while astronomers are by-and-large not professional programmers. Projecting into the data-laden future, these factors will not greatly change, and so there will conceivably always be a role for the super-efficient “black box” codes. Unfortunately, in conjunction with this trend, the majority of smaller codes (often using the larger ones) are written for personal use by individuals, without the slightest intention of their release to the wider community.

As is growing increasingly noted, this practice has several drawbacks which undermine the quality of science both present and future: (i) it makes the exact reproduction of results impossible; (ii) it heightens the chances of coding errors affecting published results; and (iii) it unnecessarily increases the workload of the individual who is unknowingly ‘re-inventing the wheel’<sup>1</sup>. The antidote to these undesirable effects is well-written, reliable, easily accessible, *open-source* code.

Such code<sup>2</sup> has been the cornerstone of this thesis. We have found that using modern programming tools can facilitate the development of code that is not only performant, but democratised.

In particular, interpreted languages, such as our adopted Python, have primarily been the domain of simple plotting scripts and data manipulation. However, with appropriate usage of the excellent SciPy suite of packages, which interface with optimized C-code in performance-critical routines, Python becomes an extremely efficient calculation engine. Indeed, we have found little performance benefit in codes that are purely written in compiled languages such as Fortran or C over our implementations in Python. An additional reason for this is the ease of locating and rectifying *algorithmic* issues with performance in the easy-to-follow Python source.

Except under extreme conditions, the small efficiency losses are completely outweighed by the dramatic increase in source readability, documentation generation capabilities, community interaction, code *writing* efficiency, extendibility and flexibility afforded by an interpreted language. In an age of rapid model development and international collaboration, these features are indispensable.

<sup>1</sup> The counter-point to this is of course that it is while implementing a method that one often gains a level of understanding of the subject otherwise out of reach.

<sup>2</sup> At least, code with these goals.

As a first example, we developed HMF and HALOMOD (both open-source codes) with both a dynamic caching system that intelligently eliminates the need for a potentially large fraction of calculations without user intervention, and an inherently plug-and-play modelling system, allowing new custom models to be rapidly defined and injected into the overall framework without modifying the source code. These features, only feasible with interpreted languages, provide an enhanced incentive to the broad user base to both run large statistical analyses with myriad iterations of the models, and also to explore new models without the effort of coding up new structures.

The democratisation of software is possibly best-realised by web-applications, which are typically instantly available, without the need for complex installations, from anywhere. Web-applications are reaching a level of near-ubiquity in the general software context, due to the ease of creating engaging and intuitive visual interfaces, and the option of accessing additional resources over the web, along with the impressive current array of tools and frameworks enabling web-application creation. Their adoption into astronomical applications has been a slower process, but is beginning to gain traction<sup>3</sup>.

Indeed, a variety of astronomical and cosmological applications are very well-suited to the web-application format. Key indicators of this suitability include (i) data-base driven application, (ii) relatively low calculation time (iii) usage made simpler by a graphical interface and (iv) a broad base of potential users. The latter three are especially characteristic of the HMF code, and we have presented our web-application incarnation of this, called HMF<sub>CALC</sub>, in this thesis.

We have found the web interface to be highly appreciated by the community for the ease with which reliable HMFs can be accessed for a range of particular applications. To illustrate this, we show an overview of the Google Analytics<sup>4</sup> report for HMF<sub>CALC</sub> over the past two years (initial publication of the application was in July 2013). In the past year, HMF<sub>CALC</sub> has averaged over 300 users per month (up from about 72 the previous year), totalling almost 6000 individual usages (compared to 2500 in the previous year). These users have hailed from 90 countries, reflecting the global nature of research today.

Clearly, there is a demand for web-applications within astronomy to perform some of the most common tasks and analyses. This represents a step towards the future of a globally networked research environment.

### 7.2.2 *Next-generation surveys*

What kinds of surveys are required to probe the nature of dark matter? This has been a recurring (if not dominating) theme throughout this thesis, and it here constitutes the final word.

There are many observations which can probe dark matter, and to date the most successful of these have been the Lyman- $\alpha$  forest due to its location at high redshift, and Milky-Way substructure counts, due to the extremely small scales probed. Despite the strength of these probes, there are significant systematic challenges with them (Garzilli,

<sup>3</sup> We note for example the excellent web-applications at [cosmocalc.icrar.org](http://cosmocalc.icrar.org) and [hyperfit.icrar.org](http://hyperfit.icrar.org).

<sup>4</sup> <http://www.google.com/analytics/>

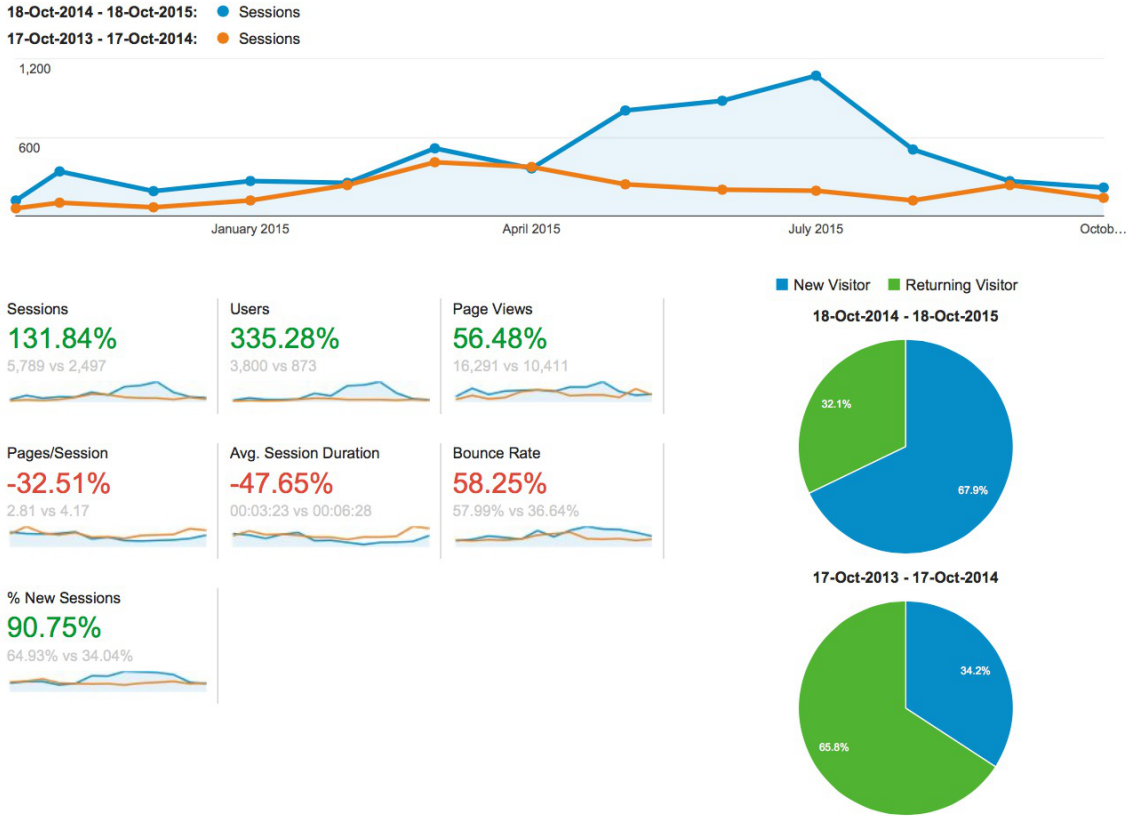


Figure 7.1: Analytics for HMFCALC over the past two years. Initial publication of the code was in July 2013.

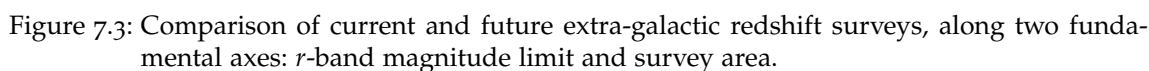
Boyarsky, and Ruchayskiy, 2015), thus justifying our focus on potential independent probes from extra-galactic surveys, and the structure formation they trace.

Cluster surveys are incredibly sensitive to the effects of dark energy, but their typical mass scales, of  $m > \sim 10^{14} h^{-1} M_{\odot}$ , are too large to be affected by the free-streaming of early-Universe particles, and therefore ineffective for dark matter studies. Similarly, large surveys tend to focus on cosmology measurements, mapping tracer galaxies which define the large-scale structure, following the growth of perturbations. These are dominantly influenced by the universal expansion, and sensitive to dark energy. Dark matter, on the other hand, is influential in the evolution of galaxies, surveys of which require detailed spatial and velocity clustering information for volume complete samples.

The most effective surveys, then, will be those that detect the lowest-mass galaxies, and therefore the lowest-mass halos. In observational parlance, what is required are surveys with a low limiting magnitude. Though such surveys will be eminently possible in the coming decade, they will be less common than their large-scale-focussed counterparts.

Figure 7.2 shows a schematic (taken from [compare.icrar.org](http://compare.icrar.org), Meyer et al. 2015) of current and future telescopes and instruments, colour-coded by a figure-of-merit derived from a combination of collecting area, field of view and number of fibres. This combination yields a rough guide to the power of the facility for constructing large surveys with the appropriate qualities for dark matter studies. In principal, any of these facilities could provide useful deep surveys, with darker bars indicating those more capable of doing so.





Through galaxy cluster and other strong probes, the nature of dark matter promises to be one mystery solved in the next decade and a half.

Strong foundational tools, such as those presented in this thesis, will form a bedrock, ensuring the coming generation of big-data surveys can be processed consistently and reproducibly. The time has come to move away from our familiar habits. More than ever, it is important that we collaborate on and open-source our next-generation toolset, just as we are, to great effectiveness, with our next-generation surveys. May this be a brick in that foundation.

---

## BIBLIOGRAPHY

---

- Abazajian, K. N. (2006). *Phys. Rev. D* 73.6, p. 063513.
- Allen, S. W., A. E. Evrard, and A. B. Mantz (2011). *Annu. Rev. Astron. Astrophys.* 49.1, pp. 409–470.
- Anderhalden, D. and J. Diemand (2011). *Mon. Not. R. Astron. Soc.* 414.4, pp. 3166–3172.
- Angulo, R. E., O. Hahn, and T. Abel (2013). *Mon. Not. R. Astron. Soc.* 434.4, pp. 3337–3347. arXiv: [arXiv:1304.2406v2](#).
- Angulo, R. E. et al. (2012). 19.March, pp. 1–19. arXiv: [arXiv:1203.3216v1](#).
- Avila, S. et al. (2015). *Mon. Not. R. Astron. Soc.* 450.December, pp. 1856–1867. arXiv: [1412.5228](#).
- Bagla, J. S. and S. Ray (2005). *Mon. Not. R. Astron. Soc.* 358.3, pp. 1076–1082.
- Baldry, I. K. et al. (2012). *Mon. Not. R. Astron. Soc.* 421.1, pp. 621–634. arXiv: [1111.5707](#).
- Bardeen, J. M. et al. (1986). *Astrophys. J.* 304, p. 15.
- Barkana, R., Z. Haiman, and J. P. Ostriker (2001). *Astrophys. J.* 558, pp. 482–496. arXiv: [0102304 \[astro-ph\]](#).
- Behroozi, P. S., C. Conroy, and R. H. Wechsler (2010). *Astrophys. J.* 717.1, p. 27. arXiv: [1001.0015](#).
- Behroozi, P. S., R. H. Wechsler, and C. Conroy (2013). *Astrophys. J.* 770.1, p. 57. arXiv: [1207.6105](#).
- Benson, A. J. et al. (2012). *Mon. Not. R. Astron. Soc.* 428.2, pp. 1774–1789.
- Bergström, L. (2000). *Reports Prog. Phys.* 63.5, pp. 793–841.
- Berlind, A. A. and D. H. Weinberg (2002). *Astrophys. J.* 575, pp. 587–616.
- Berlind, A. A. et al. (2003). *Astrophys. J.* 593, pp. 1–25.
- Beutler, F. et al. (2013). *Mon. Not. R. Astron. Soc.* 429.4, pp. 3604–3618. arXiv: [arXiv:1212.3610v1](#).
- Bhattacharya, S. et al. (2011). *Astrophys. J.* 732.2, p. 122. arXiv: [1005.2239](#).
- Blake, C., A. Collister, and O. Lahav (2008). *Mon. Not. R. Astron. Soc.* 385, pp. 1257–1269. arXiv: [0704.3377](#).
- Blas, D., J. Lesgourgues, and T. Tram (2011). *J. Cosmol. Astropart. Phys.* 2011.07, pp. 034–034.
- Bocquet, S. et al. (2015). *Arxiv Prepr.* February. arXiv: [arXiv:1502.07357v1](#).
- Bode, P., J. P. Ostriker, and N. Turok (2001). *Astrophys. J.* 556.1, pp. 93–107.
- Bond, J. R. and A. S. Szalay (1983). *Astrophys. J.* 274, p. 443.
- Bond, J. R. and G. Efstathiou (1984). *Astrophys. J.* 285, p. L45.
- Bond, J. R. et al. (1991). *Astrophys. J.* 379, p. 440.
- Bosch, F. C. van den et al. (2013). *Mon. Not. R. Astron. Soc.* 430.2, pp. 725–746.
- Bosch, F. C. van den, X. Yang, and H. J. Mo (2003). *Mon. Not. R. Astron. Soc.* 340, pp. 771–792.
- Bose, S. et al. (2015). *Arxiv Prepr.* 16.July, pp. 1–16. arXiv: [arXiv:1507.01998v2](#).
- Boyarsky, A. et al. (2014). *Phys. Rev. Lett.* 113.25, p. 251301.



- Boyarsky, A., O. Ruchayskiy, and M. Shaposhnikov (2009). *Annu. Rev. Nucl. Part. Sci.* 59.1, pp. 191–214.
- Boyarsky, A. et al. (2009). *J. Cosmol. Astropart. Phys.* 2009.05, pp. 012–012.
- Boylan-Kolchin, M., J. S. Bullock, and M. Kaplinghat (2011). *Mon. Not. R. Astron. Soc. Lett.* 415.1, pp. L40–L44.
- Brown, M. J. I. et al. (2008). *Astrophys. J.* 682, pp. 937–963. arXiv: [0804.2293](#).
- Bryan, G. L. and M. L. Norman (1998). *Astrophys. J.* 495.1, pp. 80–99.
- Bulbul, E. et al. (2014). *Astrophys. J.* 789.1, p. 13.
- Bullock, J. S. (2001). *Arxiv Prepr.* arXiv: [0106380 \[astro-ph\]](#).
- Bullock, J. S., R. H. Wechsler, and R. S. Somerville (2002). *Mon. Not. R. Astron. Soc.* 329, pp. 246–256.
- Bullock, J. S. et al. (2001). *Mon. Not. R. Astron. Soc.* 321.3, pp. 559–575.
- Buote, D. A. et al. (2007). *Astrophys. J.* 664.1, pp. 123–134.
- Cacciato, M. et al. (2013). *Mon. Not. R. Astron. Soc.* 430.2, pp. 767–786.
- Christodoulou, L. et al. (2012). *Mon. Not. R. Astron. Soc.* 425.2, pp. 1527–1548. arXiv: [1206.0943](#).
- Cirasuolo, M. et al. (2014). *Proc. SPIE* 9147, 91470N.
- Cole, S. and N. Kaiser (1989). *Mon. Not. R. Astron. Soc.* 237, p. 1127.
- Colless, M. et al. (2001). *Mon. Not. R. Astron. Soc.* 328, pp. 1039–1063.
- Comerford, J. M. and P. Natarajan (2007). *Mon. Not. R. Astron. Soc.* 379.1, pp. 190–200.
- Conroy, C., R. H. Wechsler, and A. V. Kravtsov (2006). *Astrophys. J.* 647.1, pp. 201–214.
- Contreras, S. et al. (2013). *Mon. Not. R. Astron. Soc.* 432.4, pp. 2717–2730.
- Cooray, A. (2006). *Mon. Not. R. Astron. Soc.* 365.3, pp. 842–866. arXiv: [0509033 \[astro-ph\]](#).
- Cooray, A. and M. Milosavljević (2005a). *Astrophys. J.* 627.2, pp. L85–L88. arXiv: [0503596 \[astro-ph\]](#).
- Cooray, A. and M. Milosavljević (2005b). *Astrophys. J.* 627.2, pp. L89–L92. arXiv: [0504580 \[astro-ph\]](#).
- Cooray, A. and R. Sheth (2002). *Phys. Rep.* 327.1, pp. 1–129. arXiv: [0206508v1 \[arXiv:astro-ph\]](#).
- Corasaniti, P. S. and I. Achitouv (2011). *Phys. Rev. D* 84.2, p. 023009.
- Correa, C. A. et al. (2015a). *Mon. Not. R. Astron. Soc.* 450.2, pp. 1514–1520. arXiv: [1409.5228](#).
- Correa, C. A. et al. (2015b). *Mon. Not. R. Astron. Soc.* 450.2, pp. 1521–1537. arXiv: [1501.0438](#).
- Correa, C. A. et al. (2015c). *Mon. Not. R. Astron. Soc.* 452.2, pp. 1217–1232. arXiv: [1502.0039](#).
- Coupon, J. et al. (2015). *Mon. Not. R. Astron. Soc.* 449.2, pp. 1352–1379. arXiv: [1502.02867](#).
- Courtin, J. et al. (2010). *Mon. Not. R. Astron. Soc.* 1931, no-no.
- Crocce, M. et al. (2010). *Mon. Not. R. Astron. Soc.* 403.3, pp. 1353–1367.
- Cuesta, A. J. et al. (2008). *Mon. Not. R. Astron. Soc.* 389.1, pp. 385–397.
- Cui, W., S. Borgani, and G. Murante (2014). *Mon. Not. R. Astron. Soc.* 441.2, pp. 1769–1782. arXiv: [1402.1493](#).
- Cui, W. et al. (2012). *Mon. Not. R. Astron. Soc.* 423.3, pp. 2279–2287. arXiv: [1111.3066](#).
- Curtis, H. D. (1921). *Bull. Natl. Resour. Counc.* 2, p. 171.
- Dave, R. et al. (2001). *Astrophys. J.* 547.2, pp. 574–589.



- Davis, M. et al. (1985). *Astrophys. J.* 292, p. 371.
- Dawson, K. S. et al. (2013). *Astrophys. J.* 145.1, p. 10.
- Dayal, P. et al. (2015). *Arxiv Prepr.*
- Diemer, B., S. More, and A. V. Kravtsov (2013). *Astrophys. J.* 766.1, p. 25.
- Dodelson, S. and L. M. Widrow (1994). *Phys. Rev. Lett.* 72.1, pp. 17–20.
- Dolag, K. et al. (2004). *Astron. Astrophys.* 416.3, pp. 853–864.
- Dolley, T. et al. (2014). *Astrophys. J.* 797.2, p. 125. arXiv: [1410.3489](#).
- Driver, S. P. et al. (2011). *Mon. Not. R. Astron. Soc.* 413.2, pp. 971–995.
- Driver, S. P. et al. (2015). *Arxiv Prepr.* arXiv: [1507.00676](#).
- Dubinski, J. and R. G. Carlberg (1991). *Astrophys. J.* 378, p. 496.
- Duffy, A. R. et al. (2008). *Mon. Not. R. Astron. Soc. Lett.* 390.1, pp. L64–L68.
- Dunstan, R. M. et al. (2011). *Phys. Rev. D*, pp. 1–15. arXiv: [arXiv:1109.6291v1](#).
- Durkalec, A. et al. (2015). *Astron. Astrophys.* 576, pp. L7–L10.
- Einasto, Å. (1965). *Tr. Astrofiz. Instituta Alma-Ata* 5, pp. 87–100.
- Einstein, A. (1915). *Sitzungsberichte der Königlich Preussischen Akad. der Wissenschaften*.
- Einstein, A. (1916). *Ann. Phys.* 354.7, pp. 769–822.
- Eisenstein, D. J. et al. (2005). *Astrophys. J.* 633, pp. 560–574. arXiv: [0501171 \[astro-ph\]](#).
- Eisenstein, D. J. and W. Hu (1999). *Astrophys. J.* 496.3, p. 20. arXiv: [9709112 \[astro-ph\]](#).
- Eke, V. R. et al. (2006). *Mon. Not. R. Astron. Soc.* 370.3, pp. 1147–1158.
- Eke, V. R., S. Cole, and C. S. Frenk (1996). *Mon. Not. R. Astron. Soc.*
- Firmani, C. and V. Avila-Reese (2013). *Mon. Not. R. Astron. Soc.* 432.3, pp. 2420–2432.
- Francis, M. J., G. F. Lewis, and E. V. Linder (2009). *Mon. Not. R. Astron. Soc. Lett.* 393.1, pp. L31–L35.
- Frenk, C. and S. White (2012). *Ann. Phys.* 524.9-10, pp. 507–534.
- Friedmann, A. (1922). *Zeitschrift fur Phys.* 10.1, pp. 377–386.
- Friedmann, A. (1924). *Zeitschrift fur Phys.* 21.1, pp. 326–332.
- Gao, L. et al. (2004). *Mon. Not. R. Astron. Soc.* 355.3, pp. 819–834. arXiv: [0404589 \[astro-ph\]](#).
- Garzilli, A., A. Boyarsky, and O. Ruchayskiy (2015). *Arxiv Prepr.* P. 6. arXiv: [1510.07006](#).
- Geach, J. E. et al. (2012). *Mon. Not. R. Astron. Soc.* 426.1, pp. 679–689.
- Gentile, G. et al. (2009). *Nature* 461.7264, pp. 627–8.
- Giacconi, R. et al. (1972). *Astrophys. J.* 178, p. 281.
- Giocoli, C. et al. (2010). *Mon. Not. R. Astron. Soc.* 408.1, pp. 300–313.
- Goodman, J. (1995). *Phys. Rev. D* 52.4, pp. 1821–1827. arXiv: [9506068 \[astro-ph\]](#).
- Gorbunov, D., A. Khmelnitsky, and V. Rubakov (2008). *J. High Energy Phys.* 2008.12, pp. 055–055.
- Greene, J. E. et al. (2012). *Tech. rep.* arXiv: [9809069v1 \[arXiv:gr-qc\]](#).
- Guo, H. et al. (2014). *Mon. Not. R. Astron. Soc.* 441.3, pp. 2398–2413.
- Guth, A. H. (1981). *Phys. Rev. D* 23.2, pp. 347–356.
- Haan, T. de et al. (2016). arXiv: [1603.06522](#).
- Hahn, O. and A. Paranjape (2013). *Mon. Not. R. Astron. Soc.* 438.1, pp. 878–899.
- Hamana, T. et al. (2006). *Mon. Not. R. Astron. Soc.* 369.4, pp. 1929–1938. arXiv: [0508536 \[astro-ph\]](#).
- Hasselfield, M. et al. (2013). *J. Cosmol. Astropart. Phys.* 2013.07, pp. 008–008.
- Hernquist, L. (1990). *Astrophys. J.* 356, p. 359.

- Hinshaw, G. et al. (2013). *Astrophys. J. Suppl. Ser.* 208.2, p. 25. arXiv: [arXiv:1212.5226v1](#).
- Hobbs, A. et al. (2015). *Arxiv Prepr.* March. arXiv: [arXiv:1503.02689v1](#).
- Hu, W. and A. V. Kravtsov (2003). *Astrophys. J.* 584, pp. 702–715. arXiv: [0203169 \[astro-ph\]](#).
- Hubble, E. (1929). *Proc. Natl. Acad. Sci.* 15.3, pp. 168–173.
- Huchra, J. P. et al. (1988). *Large Scale Struct. Universe Proc. 130th Symp. Int. Astron. Union.*
- Hunter, J. D. (2007). *Comput. Sci. Eng.* 9, pp. 90–95.
- Ibata, R. A. et al. (2013). *Nature* 493.7430, pp. 62–5.
- Inoue, K. T. et al. (2015). *Mon. Not. R. Astron. Soc.* 448.3, pp. 2704–2716.
- Ishiyama, T. et al. (2015). *Publ. Astron. Soc. Japan* 67, p. 61. arXiv: [arXiv:1412.2860v2](#).
- Ivezic, Z. et al. (2014). *Arxiv Prepr.* P. 39. arXiv: [0805.2366](#).
- Jenkins, A. et al. (2001). *Mon. Not. R. Astron. Soc.* 321.2, pp. 372–384.
- Jing, Y. P. (1998). *Astrophys. J.* 503.1, pp. L9–L13.
- Jing, Y. P. (1999). *Astrophys. J.* 515, pp. L45–L48. arXiv: [9901138 \[astro-ph\]](#).
- Jing, Y. P. and Y. Suto (2002). *Astrophys. J.* 574.2, pp. 538–553.
- Jones, D. H. et al. (2009). *Mon. Not. R. Astron. Soc.* 399.2, pp. 683–698.
- Jose, C., C. G. Lacey, and C. M. Baugh (2015). *Arxiv Prepr.* arXiv: [1509.06715](#).
- Kauffmann, G., A. Nusser, and M. Steinmetz (1997). *Mon. Not. R. Astron. Soc.* 286.4, pp. 795–811.
- Kauffmann, G. et al. (2004). *Mon. Not. R. Astron. Soc.* 353.3, pp. 713–731.
- Kennedy, R. et al. (2014). *Mon. Not. R. Astron. Soc.* 442.3, pp. 2487–2495.
- Kim, H.-S. et al. (2011). *Mon. Not. R. Astron. Soc.* 414.3, pp. 2367–2385.
- Kim, J.-W. et al. (2015). *Astrophys. J.* 806.2, p. 189.
- Klypin, A. et al. (1999). *Astrophys. J.* 522.1, pp. 82–92.
- Knebe, A. et al. (2013). *Mon. Not. R. Astron. Soc.* 435.2, pp. 1618–1658.
- Knebe, A. et al. (2011). *Mon. Not. R. Astron. Soc.* 415.3, pp. 2293–2318. arXiv: [1104.0949](#).
- Komatsu, E. et al. (2009). *Astrophys. J. Suppl. Ser.* 180.2, pp. 330–376.
- Krause, E. et al. (2012). *Mon. Not. R. Astron. Soc.* 428.3, pp. 2548–2564.
- Kuhlen, M. et al. (2005). *Mon. Not. R. Astron. Soc.* 357.1, pp. 387–400.
- Lacey, C. G. and S. Cole (1993). *Mon. Not. R. Astron. Soc.* 262, pp. 627–649. arXiv: [9402069 \[astro-ph\]](#).
- Lacey, C. G. and S. Cole (1994). *Mon. Not. R. Astron. Soc.* 271.
- Lagos Álvarez, B., G. Ferreira, and M. Valenzuela Hube (2011). *Proyecciones (Antofagasta)* 30.3, pp. 415–439.
- Lapi, A. and L. Danese (2014). *J. Cosmol. Astropart. Phys.* 7, p. 44. arXiv: [1407.1137](#).
- Lapparent, V. de, M. J. Geller, and J. P. Huchra (1986). *Astrophys. J.* 302, p. L1.
- Laureijs, R. et al. (2011). *arXiv.org* 1110.July, p. 3193. arXiv: [1110.3193](#).
- Lawless, J. F. (1982). *Can. J. Stat.* 10.4, pp. 316–317.
- Le Brun, A. M. C. et al. (2014). *Mon. Not. R. Astron. Soc.* 441.2, pp. 1270–1290. arXiv: [1312.5462](#).
- Leauthaud, A. et al. (2011). *Astrophys. J.* 738.1, p. 45.
- Leauthaud, A. et al. (2012). *Astrophys. J.* 744.2, p. 159.
- Lemaître, G. (1927). *Ann. la Soc. Sci. Bruxelles*, pp. 49–59.
- Lemaître, G. (1931). *Mon. Not. R. Astron. Soc.* 91.
- Lewis, A., A. Challinor, and A. Lasenby (2000). *Astrophys. J.* 538.2, pp. 473–476.

- Liddle, A. R. and D. H. Lyth (2000). Cambridge: Cambridge University Press.
- Lovell, M. R. et al. (2014). *Mon. Not. R. Astron. Soc.* 439.1, pp. 300–317.
- Ludlow, A. D. et al. (2014). *Mon. Not. R. Astron. Soc.* 441.1, pp. 378–388. arXiv: [1312.0945](#).
- Ludlow, A. D. et al. (2016). *Mon. Not. R. Astron. Soc.* 000.0000, pp. 0–0. arXiv: [1601.02624](#).
- Lukić, Z. et al. (2007). *Astrophys. J.* 671.2, pp. 1160–1181.
- Ma, C. and J. N. Fry (2000). *Astrophys. J.* 10, pp. 503–513.
- Ma, C. et al. (2011). *Mon. Not. R. Astron. Soc.* 411.4, pp. 2644–2652.
- Macciò, A. V., A. A. Dutton, and F. C. van den Bosch (2008). *Mon. Not. R. Astron. Soc.* 391.4, pp. 1940–1954.
- Macciò, A. V. and F. Fontanot (2010). *Mon. Not. R. Astron. Soc. Lett.* 404.1, pp. L16–L20.
- Macciò, A. V. et al. (2007). *Mon. Not. R. Astron. Soc.* 378.1, pp. 55–71.
- Macciò, A. V. et al. (2012). *Mon. Not. R. Astron. Soc.* 424.2, pp. 1105–1112.
- Magliocchetti, M. and C. Porciani (2003). *Mon. Not. R. Astron. Soc.* 346.1, pp. 186–198. arXiv: [0304003 \[astro-ph\]](#).
- Mandelbaum, R. et al. (2005). *Mon. Not. R. Astron. Soc.* 362.4, pp. 1451–1462. arXiv: [0410711 \[astro-ph\]](#).
- Mandelbaum, R. et al. (2006). *Mon. Not. R. Astron. Soc.* 368.2, pp. 715–731. arXiv: [0511164 \[astro-ph\]](#).
- Manera, M., R. K. Sheth, and R. Scoccimarro (2010). *Mon. Not. R. Astron. Soc.* 402.1, pp. 589–602. arXiv: [0906.1314](#).
- Manera, M. et al. (2012). *Mon. Not. R. Astron. Soc.* 428.2, pp. 1036–1054.
- Mantz, A. et al. (2010). *Mon. Not. R. Astron. Soc.* 406.3, no–no.
- Marković, K. et al. (2010). *Arxiv Prepr.* P. 26. arXiv: [1009.0218](#).
- Martizzi, D. et al. (2014). *Mon. Not. R. Astron. Soc.* 440.3, pp. 2290–2299. arXiv: [1309.4094](#).
- McCarthy, I. G. et al. (2013). *Mon. Not. R. Astron. Soc.* 440.March, pp. 3645–3657. arXiv: [1312.5341](#).
- McConnachie, A. et al. (2016), p. 210. arXiv: [1606.00043](#).
- McCracken, H. J. et al. (2015). *Mon. Not. R. Astron. Soc.* 449.1, pp. 901–916.
- McGaugh, S. S., M. K. Barker, and W. J. G. de Blok (2003). *Astrophys. J.* 584.2, pp. 566–576.
- Mehrtens, N. et al. (2012). *Mon. Not. R. Astron. Soc.* 423.2, pp. 1024–1052.
- Meszaros, P. (1974). *Astron. Astrophys.* 37, pp. 225–228.
- Meyer, Â. et al. (2015). *Proc. Adv. Astrophys. with Sq. Km. Array (AASKA14)*. 9–13 June.
- Mo, H. J. and S. D. M. White (1996). *Mon. Not. R. Astron. Soc.* 282.2, pp. 347–361. arXiv: [9512127 \[astro-ph\]](#).
- Mohammed, I. et al. (2014). *Arxiv Prepr.* October, p. 16. arXiv: [1410.6826](#).
- Moore, B. et al. (1998). *Astrophys. J.* 499.1, pp. L5–L8.
- Moore, B. et al. (1999). *Mon. Not. R. Astron. Soc.* 310.4, pp. 1147–1152.
- Moore, B. et al. (2000). *Astrophys. J.* 535.1, pp. L21–L24.
- More, S. (2013). *Astrophys. J.* 777.2, p. L26.
- More, S. et al. (2015). *Astrophys. J.* 806.1, p. 2.
- Mostek, N. et al. (2013). *Astrophys. J.* 767.1, p. 89.
- Moster, B. P. et al. (2010). *Astrophys. J.* 710, pp. 903–923. arXiv: [0903.4682](#).

- Moustakas, L. A. and R. S. Somerville (2002). *Astrophys. J.* 577, pp. 1–10. arXiv: [0110584 \[astro-ph\]](#).
- Murray, S. G., C. Power, and A. S. G. Robotham (2013a). *Astron. Comput.* 3-4, pp. 23–34.
- Murray, S. G., C. Power, and A. S. G. Robotham (2013b). *Mon. Not. R. Astron. Soc. Lett.* 434.1, pp. L61–L65. arXiv: [arXiv:1306.5140v1](#).
- Mutch, S. J., D. J. Croton, and G. B. Poole (2013). *Mon. Not. R. Astron. Soc.* 435.3, pp. 2445–2459. arXiv: [1304.2774](#).
- Narayanan, V. K. et al. (2000). *Astrophys. J.* 543.2, pp. L103–L106.
- Navarro, J. F., C. S. Frenk, and S. D. M. White (1997). *Astrophys. J.* 490.2, pp. 493–508.
- Navarro, J. F. et al. (2004). *Mon. Not. R. Astron. Soc.* 349.3, pp. 1039–1051.
- Neto, A. F. et al. (2007). *Mon. Not. R. Astron. Soc.* 381.4, pp. 1450–1462.
- Neyman, J., E. L. Scott, and C. D. Shane (1953). *Astrophys. J.* 117, p. 92.
- Noh, Y. and J. D. Cohn (2012). *Am. Astron. Soc.*
- Ogata, H. (2005). *Publ. RIMS, Kyoto Univ.* 41, pp. 949–970.
- Pacucci, F., A. Mesinger, and Z. Haiman (2013). *Mon. Not. R. Astron. Soc. Lett.* 435.1, pp. L53–L57.
- Paduroiu, S., Y. Revaz, and D. Pfenniger (2015). *Arxiv Prepr. June*. arXiv: [arXiv:1506.03789v1](#).
- Palamara, D. P. et al. (2013). *Astrophys. J.* 764.1, p. 31.
- Paranjape, A., R. K. Sheth, and V. Desjacques (2013). *Mon. Not. R. Astron. Soc.* 431.2, pp. 1503–1512.
- Paranjape, A. et al. (2013). *Mon. Not. R. Astron. Soc.* 436.1, pp. 449–459.
- Paranjape, A. (2014). *Phys. Rev. D* 90.2, p. 023520.
- Parkinson, D. et al. (2012). *Phys. Rev. D* 86.10, pp. 1–23. arXiv: [1210.2130](#).
- Peacock, J. A. (2007). *Mon. Not. R. Astron. Soc.* 379.3, pp. 1067–1074.
- Peacock, J. A. and A. F. Heavens (1990). *Mon. Not. R. Astron. Soc. (ISSN 0035-8711)* 243, pp. 133–143.
- Peacock, J. A. and R. E. Smith (2000). *Mon. Not. R. Astron. Soc.* 318.4, pp. 1144–1156.
- Peebles, P. J. E. (1980). Princeton: Princeton University Press.
- Peebles, P. J. E. (2001). *Astrophys. J.* 557.2, pp. 495–504.
- Percival, W. J. et al. (2001). *Mon. Not. R. Astron. Soc.* 327, pp. 1297–1306.
- Perlmutter, S. et al. (1999). *Astrophys. J.* 517.2, pp. 565–586.
- Pierre, M. et al. (2011). *Mon. Not. R. Astron. Soc.* 414.2, pp. 1732–1746.
- Piffaretti, R. et al. (2011). *Astron. Astrophys.* 534, A109.
- Pillepich, A., C. Porciani, and O. Hahn (2010). *Mon. Not. R. Astron. Soc.* 402.1, pp. 191–206. arXiv: [0811.4176](#).
- Pillepich, A., C. Porciani, and T. H. Reiprich (2012). *Mon. Not. R. Astron. Soc.* 422.1, pp. 44–69.
- Planck Collaboration (2014a). *Astron. Astrophys.* 571, A16. arXiv: [1303.5076](#).
- Planck Collaboration (2014b). *Astron. Astrophys.* 571. Ed. by J. Tauber, A29.
- Planck Collaboration (2015). *Astron. Astrophys.* P. 20. arXiv: [1502.0158](#).
- Polisensky, E. and M. Ricotti (2011). *Phys. Rev. D* 83.4, p. 043506.
- Poole, G. B. et al. (2015). *Mon. Not. R. Astron. Soc.* 449.2, pp. 1454–1469.
- Power, C. (2013). *Publ. Astron. Soc. Aust.* 30, e053.

- Power, C. and A. Knebe (2006). *Mon. Not. R. Astron. Soc.* 370.2, pp. 691–701.
- Prada, F. et al. (2006). *Astrophys. J.* 645.2, pp. 1001–1011.
- Press, W. H. and P. Schechter (1974). *Astrophys. J.* 187, pp. 425–438.
- Primack, J. R. and D. B. Cline (2009). *AIP Conf. Proc.* Vol. 1166. AIP, pp. 3–9.
- Quadri, R. F. et al. (2008). *Astrophys. J.* 685, pp. L1–L4. arXiv: [0808.0911](#).
- Reddick, R. M. et al. (2014). *Astrophys. J.* 783.2, p. 118.
- Reed, D. S. et al. (2003). *Mon. Not. R. Astron. Soc.* 346.2, pp. 565–572.
- Reed, D. S. et al. (2007). *Mon. Not. R. Astron. Soc.* 374.1, pp. 2–15.
- Reichardt, C. L. et al. (2013). *Astrophys. J.* 763.2, p. 127.
- Robertson, H. P. (1935). *Astrophys. J.* 82, p. 284.
- Robertson, H. P. (1936a). *Astrophys. J.* 83, p. 187.
- Robertson, H. P. (1936b). *Astrophys. J.* 83, p. 257.
- Robitaille, T. P. et al. (2013). *Astron. Astrophys.* 558, A33.
- Robotham, A. S. G. et al. (2011). *Mon. Not. R. Astron. Soc.* 416.4, pp. 2640–2668.
- Rocha, M. et al. (2013). *Mon. Not. R. Astron. Soc.* 430.1, pp. 81–104.
- Ross, A. J., W. J. Percival, and R. J. Brunner (2010). *Mon. Not. R. Astron. Soc.* 407.1, pp. 420–434. arXiv: [1002.1476](#).
- Rubin, V. C. and W. K. Ford (1970). *Astrophys. J.* 159, p. 379.
- SDSS Collaboration (2008). *Astrophys. J. Suppl. Ser.* 175, pp. 297–313.
- SDSS Collaboration (2009). *Astrophys. J. Suppl. Ser.* 182.2, pp. 543–558.
- Sartoris, B. et al. (2015). 000.May. arXiv: [1505.02165](#).
- Schechter, P. (1976). *Astrophys. J.* Pp. 203–297.
- Schmidt, B. P. et al. (1998). *Astrophys. J.* 507.1, pp. 46–63.
- Schneider, A. (2014). *Arxiv Prepr.* arXiv: [1412.2133](#).
- Schneider, A., R. E. Smith, and D. S. Reed (2013). *Mon. Not. R. Astron. Soc.* 433.2, p. 16. arXiv: [1303.0839](#).
- Schneider, A. et al. (2012). *Mon. Not. R. Astron. Soc.* 424.1, pp. 684–698.
- Schneider, A. et al. (2014). *Mon. Not. R. Astron. Soc. Lett.* 441.1, pp. L6–L10.
- Scoccimarro, R. et al. (2001). *Astrophys. J.* 546, pp. 20–34.
- Scrimgeour, M. I. et al. (2012). *Mon. Not. R. Astron. Soc.* 425.
- Seljak, U. (2000). *Mon. Not. R. Astron. Soc.* 213.
- Seljak, U. and M. S. Warren (2004). *Mon. Not. R. Astron. Soc.* 355.1, pp. 129–136.
- Seljak, U. et al. (2006). *Phys. Rev. Lett.* 97.19, p. 191303.
- Shapley, H. (1921). *Bull. Natl. Resour. Counc.* 2, p. 194.
- Shen, Y. et al. (2013). *Astrophys. J.* 778.2, p. 98.
- Sheth, R. K., H. J. Mo, and G. Tormen (2001). *Mon. Not. R. Astron. Soc.* 323.1, pp. 1–12.
- Sheth, R. R. K. et al. (2001). *Mon. Not. R. Astron. Soc.* 1302.4, pp. 1288–1302.
- Sheth, R. K. and G. Tormen (1999). *Mon. Not. R. Astron. Soc.* 308, pp. 119–126. arXiv: [9901122 \[astro-ph\]](#).
- Simons, D. A. et al. (2014). *Proc. SPIE*. Ed. by L. M. Stepp, R. Gilmozzi, and H. J. Hall. Vol. 9145, p. 914515.
- Skibba, R. A. et al. (2014). *Astrophys. J.* 784.2, p. 128. arXiv: [1310.1093](#).
- Skibba, R. a. and R. K. Sheth (2009). *Mon. Not. R. Astron. Soc.* 392.3, pp. 1080–1091.
- Skibba, R. a. et al. (2015). *Astrophys. J.* 807.2, p. 152.



- Slipher, V. M. (1915). *Pop. Astron.* 23, pp. 21–24.
- Smith, R. E. and K. Marković (2011). *Phys. Rev. D* 4.1, pp. 1–18. arXiv: [arXiv : 1103 . 2134v2](#).
- Smith, R. E. et al. (2003). *Mon. Not. R. Astron. Soc.* 341.4, pp. 1311–1332.
- Smith, R. E., V. Desjacques, and L. Marian (2011). *Phys. Rev. D* 83.4, p. 043526. arXiv: [arXiv:1009.5085v2](#).
- Smoot, G. F. et al. (1992). *Astrophys. J.* 396, p. L1.
- Souza, R. S. de et al. (2013). *Mon. Not. R. Astron. Soc.* 432.4, pp. 3218–3227.
- Spergel, D. N. et al. (2003). *Astrophys. J. Suppl. Ser.* 148.1, pp. 175–194.
- Spergel, D. N. et al. (2007). *Astrophys. J. Suppl. Ser.* 170.2, pp. 377–408.
- Springel, V., S. D. M. White, and A. Jenkins (2005). *Nature* 435.June, pp. 629–636.
- Springel, V. et al. (2008). *Mon. Not. R. Astron. Soc.* 391.4, pp. 1685–1711. arXiv: [0809.0898](#).
- Staveley-Smith, L. and T. Oosterloo (2015). *Proc. Adv. Astrophys. with Sq. Km. Array (AASKA14)*. 9 -13 June.
- Steinhardt, C. L. et al. (2015). *Arxiv Prepr.* P. 7. arXiv: [arXiv:1506.01377v1](#).
- Sunayama, T. et al. (2015). *Arxiv Prepr.*
- Sunyaev, R. A. and Y. B. Zeldovich (1972). *Comments Astrophys. Sp. Phys.* 4.
- Szapudi, I. et al. (2005). *Arxiv Prepr.* P. 4. arXiv: [0505389 \[astro-ph\]](#).
- Takada, M. and W. Hu (2013). *Phys. Rev. D* 87.12. arXiv: [1302.6994](#).
- Takada, M. et al. (2013). *Arxiv Prepr.* May 2009, pp. 1–42. arXiv: [1206.0737](#).
- Takahashi, R. et al. (2012). *Astrophys. J.* 761.2, p. 152. arXiv: [1208.2701](#).
- Taylor, J. E. (2011). *Adv. Astron.* 2011, pp. 1–17. arXiv: [arXiv:1008.4103v2](#).
- Tinker, J. L. and D. Weinberg (2005). *Astrophys. J.* 631, pp. 41–58.
- Tinker, J. L. et al. (2008). *Astrophys. J.* 688, pp. 709–728.
- Tinker, J. L. et al. (2010). *Astrophys. J.* 724.2, pp. 878–886.
- Tinker, J. L. et al. (2011). *Astrophys. J.* 16, p. 23. arXiv: [1104.1635](#).
- Tinker, J. L. et al. (2013). *Astrophys. J.* 778.2, p. 93.
- Viel, M. et al. (2012). *Mon. Not. R. Astron. Soc.* 421.1.
- Viel, M. et al. (2005). *Phys. Rev. D* 71.6, pp. 1–10. arXiv: [0501562 \[astro-ph\]](#).
- Viel, M. et al. (2006). *Phys. Rev. Lett.* 97.7, p. 071301.
- Viel, M. et al. (2008). *Phys. Rev. Lett.* 100.4, p. 041304.
- Viel, M. et al. (2013). *Phys. Rev. D* 88.4, pp. 1–20. arXiv: [1306.2314](#).
- Vikhlinin, A. et al. (1998). *Astrophys. J.* 502.2, pp. 558–581.
- Vikhlinin, A. et al. (2009). *Astrophys. J.* 692.2, pp. 1060–1074.
- Voit, G. M. (2005). *Rev. Mod. Phys.* 77.1, pp. 207–258.
- Wake, D. A. et al. (2008a). *Mon. Not. R. Astron. Soc.* 387.3, pp. 1045–1062. arXiv: [0802 . 4288](#).
- Wake, D. A. et al. (2008b). *Mon. Not. R. Astron. Soc.* 391.4, pp. 1674–1684. arXiv: [0810 . 1050](#).
- Wake, D. A. et al. (2011). *Astrophys. J.* 728.1, p. 46.
- Walker, A. G. (1935). *Mon. Not. R. Astron. Soc.* 95.
- Wang, J. and S. D. M. White (2007). *Mon. Not. R. Astron. Soc.* 380.1, pp. 93–103.
- Warren, M. S. et al. (2006). *Astrophys. J.* 646.2, pp. 881–885.

- Watson, W. A. et al. (2013). *Mon. Not. R. Astron. Soc.* 433.2, pp. 1230–1245. arXiv: [1212.0095](#).
- Wechsler, R. H. et al. (2002). *Astrophys. J.* 568.1, pp. 52–70.
- Wen, Z. L., J. L. Han, and F. S. Liu (2012). *Astrophys. J. Suppl. Ser.* 199.2, p. 34.
- White, S. D. M. (2002). *Astrophys. J. Suppl. Ser.* 143, pp. 241–255.
- White, S. D. M. and M. J. Rees (1978). *Mon. Not. R. Astron. Soc.* 183, pp. 341–358.
- Wu, K. K. S., O. Lahav, and M. J. Rees (1999). *Nature* 397.6716, pp. 225–230.
- Yan, R., D. S. Madgwick, and M. White (2003). *Astrophys. J.* 598, pp. 848–857.
- Yang, X., H. J. Mo, and F. C. Van den Bosch (2003). *Mon. Not. R. Astron. Soc.* 339.4, pp. 1057–1080. arXiv: [0207019 \[astro-ph\]](#).
- Yang, X. et al. (2005). *Mon. Not. R. Astron. Soc.* 358.1, pp. 217–232. arXiv: [0410114 \[astro-ph\]](#).
- Zaldarriaga, M. and U. Seljak (2000). *Astrophys. J.* 129, pp. 431–434.
- Zehavi, I. et al. (2004). *Astrophys. J.* 608.1, pp. 16–24.
- Zehavi, I. et al. (2005). *Astrophys. J.* 630.1, pp. 1–27.
- Zehavi, I. et al. (2011). *Astrophys. J.* 736.1, p. 59.
- Zemp, M. et al. (2011). *Astrophys. J. Suppl. Ser.* 197.2, p. 30.
- Zentner, A. R. (2007). *Int. J. Mod. Phys. D* 16, pp. 763–815. arXiv: [0611454 \[astro-ph\]](#).
- Zhao, D. H. et al. (2003). *Mon. Not. R. Astron. Soc.* 339.1, pp. 12–24.
- Zhao, D. H. et al. (2009). *Astrophys. J.* 707.1, pp. 354–369.
- Zheng, Z. (2004). *Astrophys. J.* 610, pp. 61–68. arXiv: [0307030 \[astro-ph\]](#).
- Zheng, Z., A. L. Coil, and I. Zehavi (2007). *Astrophys. J.* 667, pp. 760–779. arXiv: [0703457 \[astro-ph\]](#).
- Zheng, Z. and H. Guo (2015). *Arxiv Prepr.* arXiv: [arXiv:1506.07523v1](#).
- Zheng, Z. et al. (2005).
- Zheng, Z. et al. (2009). *Astrophys. J.* 707, pp. 554–572. arXiv: [0809.1868](#).
- Zwicky, F. (1933). *Helv. Phys. Acta* 6, pp. 110–127.





# Appendix



---

MRP SUPPLEMENTS

---

## A.1 RECURRENCE RELATION

A problem that arises in the calculation of the incomplete gamma function is that several libraries only implement  $\Gamma(z, x)$  for  $z > 0$ , and we often need it for  $z < 0$ . One way around this is to use the recurrence relation

$$\Gamma(z, x) = \frac{\Gamma(z+1, x) - x^z e^{-x}}{z}, \quad (\text{A.1})$$

In general, for arbitrary negative  $z$ , one can use the following extended relation to calculate the incomplete gamma function in terms of only positive shape parameter:

$$\Gamma(z, x) = \Gamma(z) \left( Q(z+n, x) - x^z e^{-x} \sum_{k=0}^{n-1} \frac{x^k}{\Gamma(z+k+1)} \right), \quad (\text{A.2})$$

where  $Q$  is the regularized incomplete gamma,  $Q(z, x) = \Gamma(z, x)/\Gamma(z)$ . To ensure the use of only positive shape parameters,  $n$  should be set to  $\lceil -z \rceil$ . Note that the accuracy of this formula will decrease for large  $|z|$  due to numerical precision, and likewise the evaluation time will increase due to the sum.

## A.2 RE-PARAMETERISATIONS

There are several re-parameterisations of the GGD to be found in the literature, which could potentially bring better covariance properties than vanilla MRP. In this section, we investigate three of these formulations, along with one that we have devised.

Table A.1 summarises the parameterisations we will investigate here. There is also at least one form in the literature which claims to have better properties than those listed here (Lawless, 1982), but is purely defined for  $z > 0$ , and therefore cannot be used.

The HT formulation has the immediate benefit that two of the parameters have physical meanings:  $\nu$  is the power-law slope, and  $\mu \equiv \mathcal{H}_T$  is the logarithmic mass mode. We expect this to be a very robust scale in any given data in which it exists (it will not exist when the slope is  $\leq -2$ ), and should be able to be fit almost independently of the other parameters.

In non-hierarchical contexts, we may simply test the expected covariance of any re-parameterisation by converting the identified covariance of the MRP form to any re-parameterisation for which we have the derivatives of the transform.

REF	$\mathcal{H}_s$	$\alpha$	$\beta$
GG2 (Lagos Álvarez, Ferreira, and Valenzuela Hube, 2011)	$\mu^{-1/\delta}$	$\nu - 1$	$\delta$
GG3 (Lagos Álvarez, Ferreira, and Valenzuela Hube, 2011)	$\mu$	$\delta\nu - 1$	$\delta$
HT	$\mu \left(\frac{\delta}{\nu+2}\right)^{1/\delta}$	$\nu$	$\delta$

Table A.1: Alternate parameterisations found in the literature, or presented here. Columns indicate the transformation from the new parameters  $(\mu, \nu, \delta)$  to the standard  $(\mathcal{H}_s, \alpha, \beta)$ . We note that in the HT formula,  $\mu$  is equivalent to the logarithmic mass mode  $\mathcal{H}_T$ .

Given a new vector of parameters,  $\vec{\psi}$ , and a set of three equations  $\vec{\theta}(\vec{\psi}) = \vec{\theta}$  relating the parameters, then the Jacobian at any point  $\vec{\psi}$  is given by

$$J_i^\psi = \vec{J}^\theta \cdot \frac{\partial \vec{\theta}}{\partial \vec{\psi}_i}. \quad (\text{A.3})$$

We can then calculate the Hessian as

$$H_{ij}^\psi = \vec{J}^\theta \cdot \frac{\partial \vec{\theta}}{\partial \vec{\psi}_i \partial \vec{\psi}_j} + \frac{\partial \vec{\theta}}{\partial \vec{\psi}_j} \cdot \mathbf{H}^\theta \cdot \frac{\partial \vec{\theta}}{\partial \vec{\psi}_i}. \quad (\text{A.4})$$

To test the covariance properties of these parameterisations with respect to the MRP form, we generate mock halo catalogues by sampling the MRP at a typical point of  $\vec{\theta} = (14.5, -1.85, 0.72)$  (corresponding roughly to P13 best-fit values). We do this for a series of differing truncation masses (from  $m_{\min} = 11$  to 13) and for each, we produce 20 realisations, calculating the Hessians at the input point using Eq. A.10. We convert these into the new parameterisation space using Eq. A.4, and then convert the result into covariance and correlation matrices.

Figure A.1 shows the mean results of this analysis, with errorbars. Each of the linestyles represents different parameterisations, while the colours represent different quantities. The top panel shows the ratios of standard deviations in the new parameterisation to MRP. The standard deviations are all normalised to their parameter values. This indicates which parameterisations are more effective at net constraints on the parameters. We find that, as expected,  $\delta \sim \beta$  is not improved for any of the alternate forms. On the other hand,  $\nu$  is estimated similarly for HT, but is outside the range of the plot for both GG2 and GG3, and so has poorer precision. However,  $\mu$  varies widely between the new forms. For GG2, it has a far poorer relative deviation, while for GG3, as expected, it remains consistent with unity. For HT, it is significantly more precise – for some truncation masses it has double the precision.

The double-parameter quantities in the lower panel are the ratios of the correlation coefficients between the new parameterisations and the MRP. While the  $\mu - \nu$  correlation swaps sign for GG2 and GG3, it perhaps surprisingly remains consistent with a parity of magnitude. Conversely, for HT, the correlation is rather dependent on  $m_{\min}$ , but is always significantly lower than MRP, as expected. We find much the same result for  $\mu - \delta$ , though HT only beats MRP for high  $m_{\min}$ . Finally,  $\nu - \delta$  is relatively unchanged in all parameterisations except for GG3, in which it performs poorly at low  $m_{\min}$ .

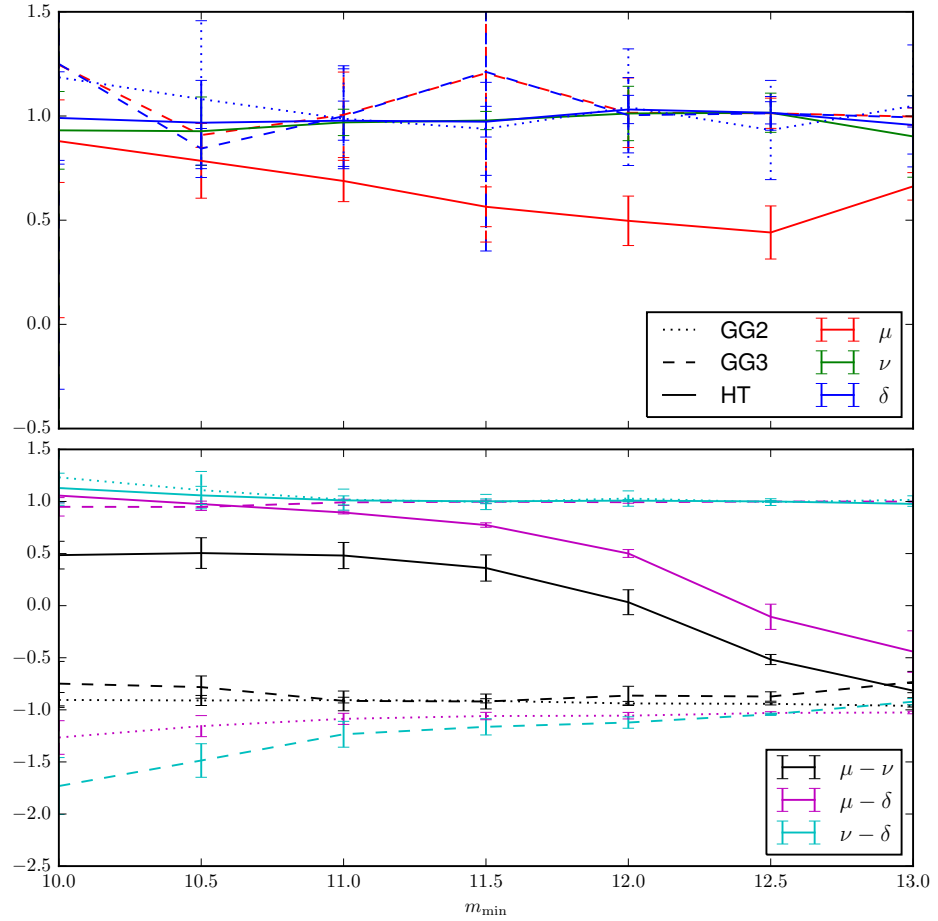


Figure A.1: Properties of three reparameterisations of the MRP compared to the form used in chapter 4.  $x$ -axis represents truncation mass. All dotted (dashed, solid) lines refer to GG2 (GG3, HT) respectively. The colours represent different quantities – the red (green, blue) line in the top-panel is the ratio of the standard deviation of  $\mu$  ( $\nu$ ,  $\delta$ ) to  $\mathcal{H}_s$  ( $\alpha$ ,  $\beta$ ) at the solution (where the standard deviation is relative to the value at the solution). The black (magenta, cyan) line in the lower panel is the ratio of the correlation coefficient between  $\mu - \nu$  ( $\mu - \delta$ ,  $\nu - \delta$ ) to the corresponding MRP correlations. All ratios are taken as the means of 20 realisations from ideal mock catalogues.

We conclude that neither GG2 or GG3 are useful alternatives for the MRP in the context of typical values for the HMF, having no beneficial properties. Conversely, HT is either beneficial or benign in every aspect, and provides an interesting re-formulation for future studies. Its major potential drawback is that it is undefined for  $\nu \leq 2$ , which may be transgressed in some rare cases for certain values of  $m_{\min}$ .

An interesting corollary is that the same situation applies to the popular Schechter function. In that case,  $\beta \equiv 1$ , so that  $\mathcal{H}_T = \mathcal{H}_s(\alpha + 2)$ . Since  $\alpha \approx -1$ ,  $\mathcal{H}_s$  has often been directly associated with the turnover. However, this is not strictly true, and for common measurements of  $\alpha$  between -1.4 and -1.1, there may be a corresponding benefit in using the true turnover scale.

### A.3 TRUNCATION MASS DEPENDENCE AND EXTENSIONS

Though the MRP is accurate within intrinsic HMF uncertainties (especially if fit over a smaller mass range), there are two main problems with it, both arising from the same source: i) it is not very precise when fit over a large mass range (small truncation mass), and ii) the parameters themselves depend on the truncation mass.

These problems both arise because there are systematic differences between the MRP and the intrinsic HMF distribution, which can be seen clearly in figure 4.1. An ideal model would have a similar level of deviation over any range of masses, and importantly would not have the parameters dependent on the unphysical truncation mass. In this section, we set out to determine the extent of the  $m_{\min}$ -dependence of the MRP parameters, and explore an extension which may reduce it.

To ascertain the dependence of the parameters on  $m_{\min}$ , we use our fiducial cosmology, P13, and fit the parameters for a wide range of  $m_{\min}$  values. We do this using both the I15 model and also simulated halos from  $\nu^2$ GC-M. We plot the ratio of the values to those found at the smallest applicable  $m_{\min}$  in figure A.2 and find that both  $\mathcal{H}_s$  and  $\alpha$  are very stable, while  $\beta$  can change by almost 15% when fit against I15 and almost 50% when fit to simulations. The latter figure is not surprising, since  $\beta$  is difficult to fit to point data where the density of high-mass objects is low.

What figure A.2 does seem to confirm is that the systematic differences between MRP and I15 are also present in a similar way between MRP and the simulated halos. This means that MRP is losing systematic information which the EPS formalism contains.

However, the effect is not a large one – the primary two parameters which are highly constrained are relatively unchanging over a wider range of masses than we will be able to fit to in the near future. Nevertheless, it is possible that an extension to the MRP may reduce this dependence, while simultaneously increasing accuracy.

We have tested a “double-MRP” form:

$$g_{dbl}(\mathcal{H}_s, \alpha, \beta, \gamma) \propto \beta \left( k \left( \frac{m}{\mathcal{H}_s} \right)^\gamma + \left( \frac{m}{\mathcal{H}_s} \right)^\alpha \right) e^{-\left( \frac{m}{\mathcal{H}_s} \right)^\beta}, \quad (\text{A.5})$$

and found that while it provides a tenfold improvement in accuracy when fit with  $m_{\min} = 10^{6.5} h^{-1} M_\odot$ , its additional parameters are incredibly sensitive to  $m_{\min}$ , and discontinuously in different regimes. Further improvement may yet be possible.

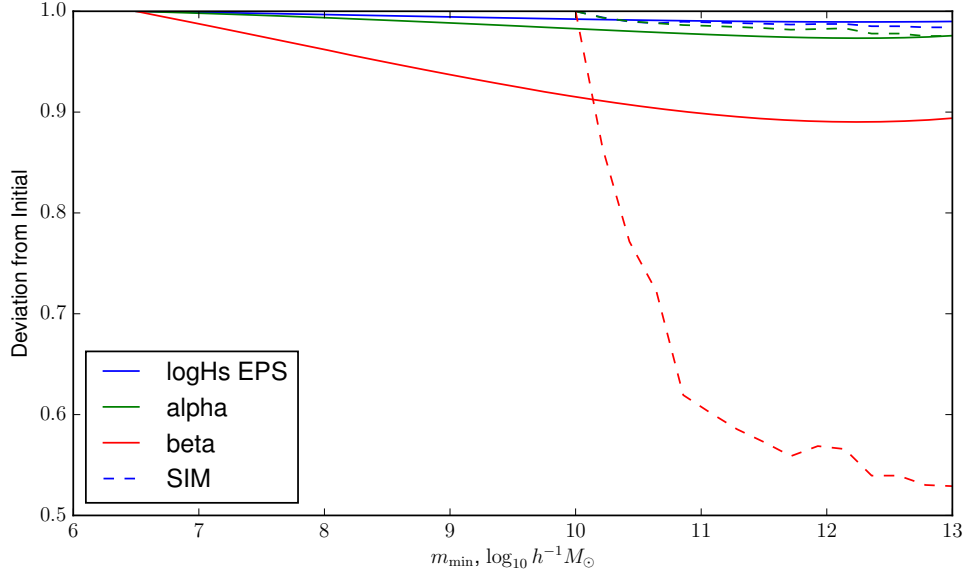


Figure A.2: The relative change in MRP parameters over truncation mass,  $m_{\min}$ , in P13 cosmology. Solid lines are fits to the analytic fit from I15, while dashed lines are fits to  $v^2$ GC-M simulated halos.

#### A.4 JACOBIANS AND HESSIANS OF MRP

It is often useful to have analytic expressions for the log-likelihood and its first and second derivatives, especially for downhill-gradient schemes and some MCMC step methods. These may be used to accelerate convergence and to estimate parameter covariances at the solution.

We here derive these quantities in the two generic cases which often arise: fitting to binned data (or theoretical curves), or the full halo mass distribution.

##### A.4.1 Method I: Binned data

The simplest method is to fit to binned data, or equivalently, a theoretical curve. We always perform this fit in log-space, so that the weighting of various regimes is similar. In this case, the log-likelihood can be expressed as

$$\ln \mathcal{L}_{\text{bin}} = -\frac{\chi^2}{2} = -\sum_i^N \frac{(\ln f(\vec{\theta}, m_i) - \ln d_i / \mathcal{D})^2}{2\sigma_i^2}, \quad (\text{A.6})$$

where  $f$  is defined as in Eq. 4.13,  $d_i$  is the value of the  $i^{\text{th}}$  bin of data,  $\mathcal{D}$  is the integral of the data with respect to  $m$ , and  $\sigma_i$  is the uncertainty in the  $i^{\text{th}}$  bin of data (this could be set to the value of the data point, unity, or some other value, depending on the specific example).

Setting the difference between logarithmic model and data to  $\Delta$ , we determine a Jacobian with the following form:

$$J_{\text{bin}}^j = \frac{-1}{\sigma^2} \sum \frac{d(\ln g_i - \ln q)}{d\vec{\theta}^j} \Delta. \quad (\text{A.7})$$

Similarly, the Hessian is

$$H_{bin}^{jk} = \frac{-1}{\sigma^2} \sum \frac{d^2(\ln g_i - \ln q)}{d\vec{\theta}^j d\vec{\theta}^k} \Delta + \frac{d(\ln g_i - \ln q)}{d\vec{\theta}^j} \frac{d(\ln g_i - \ln q)}{d\vec{\theta}^k}. \quad (\text{A.8})$$

We give the values of the necessary first- and second-derivatives of  $\ln g$  and  $\ln q$  in §A.4.3.

#### A.4.2 Method II: Per-object data

The second, and potentially more important method, is when the full distribution of halo masses is used to estimate the parameters. Here we retain the weighting factor  $s$  in the calculation, and the likelihood is given by Eq. 4.17.

Then the Jacobian is simply given by

$$J_{po}^j = \frac{N}{\sum m_i^s} \sum_{i=1}^{N_{data}} m_i^s \frac{d}{d\vec{\theta}^j} (\ln g_i - \ln q)_s, \quad (\text{A.9})$$

Here the subscript  $s$  indicates that  $\alpha + s$  is used throughout the calculation.

Clearly, the Hessian is given by

$$H_{po}^{jk} = \frac{N}{\sum m_i^s} \sum_{i=1}^{N_{data}} m_i^s \frac{d^2}{d\vec{\theta}^j d\vec{\theta}^k} (\ln g_i - \ln q)_s. \quad (\text{A.10})$$

These equations are true given that uniform priors are used for all parameters. Insertion of a non-uniform prior requires the addition of the derivatives of the logarithm of the prior distribution.

#### A.4.3 Derivatives of $g$ and $q$

Here we list the derivatives of  $\ln g$  and  $\ln q$  as coupled equations. Throughout the following we use  $h$  to represent  $\log_{10} \mathcal{H}_s$  for brevity, and denote partial derivatives with respect to any parameter by a subscript. Furthermore we use the following commonly occurring relations:

$$z = \frac{\alpha + 1}{\beta} \quad (\text{A.11})$$

$$y = m / \mathcal{H}_s \quad (\text{A.12})$$

$$x = y^\beta \cdot \Phi = yg / \Gamma(z, x). \quad (\text{A.13})$$



We begin with  $\ln g$ :

$$\ln g_h = \ln(10)(\beta x - \alpha) \quad (\text{A.14})$$

$$\ln g_\alpha = \ln y \quad (\text{A.15})$$

$$\ln g_\beta = \frac{1}{\beta} (1 - x \ln x) \quad (\text{A.16})$$

$$\ln g_{h,h} = -\ln^2(10)\beta^2 x \quad (\text{A.17})$$

$$\ln g_{\alpha,\alpha} = 0 \quad (\text{A.18})$$

$$\ln g_{\alpha,\beta} = 0 \quad (\text{A.19})$$

$$\ln g_{\beta,\beta} = -\frac{1}{\beta^2} (1 + x \ln^2 x) \quad (\text{A.20})$$

$$\ln g_{h,\alpha} = -\ln(10) \quad (\text{A.21})$$

$$\ln g_{h,\beta} = \ln(10)x(1 + \ln x) \quad (\text{A.22})$$

For the remaining, it is helpful if we first define a couple of functions based on the Meijer-G function  $G$ :

$$G1(z, x) = G_{2,3}^{3,0} \left( x \left| \begin{matrix} 1 & 1 \\ 0 & 0 & z \end{matrix} \right. \right) / \Gamma(z, x), \quad (\text{A.23})$$

$$G2(z, x) = G_{3,4}^{4,0} \left( x \left| \begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 & z \end{matrix} \right. \right) / \Gamma(z, x), \quad (\text{A.24})$$

and

$$\bar{G}(z, x) = G1(z, x) \ln x + 2G2(z, x). \quad (\text{A.25})$$

We note that everywhere in the derivatives of  $q$ , instances of  $m$  are taken to rather be  $m_{\min}$ . Keeping this in mind, for the first derivatives of  $\ln q$ , we have

$$\ln q_h = \ln(10)(1 + \Phi) \quad (\text{A.26})$$

$$\ln q_\alpha = \beta^{-1}(\ln x + G1) \quad (\text{A.27})$$

$$\ln q_\beta = -z \ln q_\alpha - \frac{\ln y}{\beta} \Phi. \quad (\text{A.28})$$

For the double-derivatives, we note that

$$\ln q_{\tilde{\theta}^i, \tilde{\theta}^j} = \frac{q_{\tilde{\theta}^i, \tilde{\theta}^j}}{q} - \ln q_{\tilde{\theta}^i} \ln q_{\tilde{\theta}^j} \quad (\text{A.29})$$

and present the functions  $\frac{q_{\vec{\theta}^i, \vec{\theta}^j}}{q} \equiv \tilde{q}_{\vec{\theta}^i, \vec{\theta}^j}$  as

$$\tilde{q}_{h,h} = \ln(10) (\ln q_h + \Phi \ln g_h) \quad (\text{A.30})$$

$$\tilde{q}_{\alpha,\alpha} = \beta^{-1} \left( \ln q_\alpha \ln x + \frac{\tilde{G}}{\beta} \right) \quad (\text{A.31})$$

$$\tilde{q}_{\alpha,\beta} = \beta^{-1} \left( -\ln q_\alpha + \ln q_\beta \ln x - \frac{z\tilde{G}}{\beta} \right) \quad (\text{A.32})$$

$$\tilde{q}_{\beta,\beta} = \frac{\Phi \ln y}{\beta} \left( \frac{1}{\beta} - \ln g_\beta \right) - z \left( \tilde{q}_{\alpha,\beta} - \frac{\ln q_\alpha}{\beta} \right) \quad (\text{A.33})$$

$$\tilde{q}_{h,\alpha} = \ln(10) (\ln q_\alpha + \Phi \ln g_\alpha) \quad (\text{A.34})$$

$$\tilde{q}_{h,\beta} = \ln(10) (\ln q_\beta + \Phi \ln g_\beta). \quad (\text{A.35})$$

## A.5 AN IDEALISED ANALYTIC MODEL

Since the MRP is analytic in form, it is possible (using idealised data) to determine the expected Fisher information matrix of a given solution. The Fisher information is an approximation of the inverse covariance (perfect in the case of a multivariate gaussian log-likelihood distribution). Such an analytical analysis has several expected benefits:

- Establish expected relationships between the parameters, including any severe degeneracies.
- Establish rule-of-thumb magnitudes of parameter constraints for given data specifications.
- Determine the best choice of the scaling  $s$  when fitting for a given dataset.

### A.5.1 Development of Model

We begin such an analysis by supposing that the data perfectly samples the MRP distribution with some  $\vec{\theta} = \vec{\theta}'$ , and we optionally weight the data with scale  $s$ . Throughout this analysis we assume that the data obeys the halo model assumption, so that the normalisation is  $A_{\rho_c}$ , though in principle this can be set arbitrarily without affecting the analysis. So the data is characterised as

$$\frac{dn'}{dm} = VA_{\rho_c} g(m|\vec{\theta}'), \quad (\text{A.36})$$

where  $V$  is the physical volume of the sample.

In the case of an ideal tracer, we convert the normalisation of the weighting (Eq. 4.16) to an integral, equating the total number density with the weighted number density to find

$$c = \frac{q'}{q'_s \mathcal{H}'_s}, \quad (\text{A.37})$$

where here, and throughout, a subscript  $s$  denotes replacement of  $\alpha$  with  $\alpha + s$  within the variable, and recall that the dash indicates the parameters of the data. From now on we will simply write  $g(m|\vec{\theta}')$  as  $g'$  and  $q(m_{\min}|\vec{\theta})$  as  $q$ .

Then the likelihood of each particle is

$$\ln \mathcal{L}_i = \ln \left( \frac{g_s^i}{q_s} \right) \frac{q'}{q'_s} \left( \frac{\mathcal{H}_s}{\mathcal{H}'_s} \right)^s. \quad (\text{A.38})$$

The total likelihood is the sum of the log-likelihoods of each particle, so we have

$$\ln \mathcal{L} = \frac{q'}{q'_s} \left( \frac{\mathcal{H}_s}{\mathcal{H}'_s} \right)^s \int_{m_{\min}}^{\infty} \frac{dn'}{dm} \ln(g_s/q_s) dm \quad (\text{A.39})$$

$$= B \int_{m_{\min}}^{\infty} g' \ln(g_s/q_s) dm, \quad (\text{A.40})$$

where

$$B = VA_{\rho_c} \frac{q'}{q'_s} \left( \frac{\mathcal{H}_s}{\mathcal{H}'_s} \right)^s. \quad (\text{A.41})$$

The integral can be separated into two parts initially, and the part with  $\ln q_s$  is simple:

$$\int_{m_{\min}}^{\infty} g' \ln q_s dm = q' \ln q_s. \quad (\text{A.42})$$

For the part with  $\ln g_s$ , we can use integration by parts. We note that

$$\int g' = \mathcal{H}'_s \gamma(z', x') \equiv \mathcal{H}'_s \Gamma(z') - q', \quad (\text{A.43})$$

where  $\gamma$  is the lower-incomplete gamma function, and use it to find

$$\int_{m_{\min}}^{\infty} g' \ln g_s dm = q' \ln g_{m,s} + \mathcal{H}'_s \int_{m_{\min}}^{\infty} -\gamma(z', x') \frac{d \ln g_s}{dm} dm, \quad (\text{A.44})$$

where the subscript  $m$  refers to the fact that here  $g$  is a function of  $m_{\min}$ , and we have

$$\frac{d \ln g_s}{dm} = \frac{\alpha_s}{m} - \frac{\beta x}{m}. \quad (\text{A.45})$$

Compiling everything thus far, we have

$$\int_{m_{\min}}^{\infty} g' \ln(g_s/q_s) dm = q' \ln(g_s/q_s) + \alpha_s t' - u, \quad (\text{A.46})$$

where

$$t = \int_{m_{\min}}^{\infty} (\Gamma(z', x') - \Gamma(z')) \frac{dm}{m} \quad (\text{A.47})$$

and

$$u = \int_{m_{\min}}^{\infty} (\Gamma(z', x') - \Gamma(z')) \frac{\beta x}{m} dm. \quad (\text{A.48})$$

For both, we use the substitution  $dm = m dx' / (\beta' x')$ , and for  $u$  we also use integration by parts to finally yield

$$\ln \mathcal{L} = B (q' \ln(g/q) + t - u), \quad (\text{A.49})$$

with

$$u = \left( \frac{\mathcal{H}'_s}{\mathcal{H}_s} \right)^\beta \left( \mathcal{H}'_s \Gamma(z' + \frac{\beta}{\beta'}, x') - x'^{\beta/\beta'} q' \right) \quad (\text{A.50})$$

and

$$t' = \frac{\mathcal{H}'_s}{\beta'} \Gamma(z') \left[ y'^{\alpha'+1} \Gamma(z') {}_pF_q \left( \begin{matrix} z' & z' \\ z'+1 & z'+1 \end{matrix}, -x' \right) + \text{Poly}\Gamma(z') - \beta' \ln y' \right], \quad (\text{A.51})$$

with  ${}_pF_q$  the generalised hypergeometric function.

We then want to find the hessian of the likelihood, with respect to the model parameters. We find

$$\frac{\partial^2 \ln \mathcal{L}}{\partial \vec{\theta}^j \partial \vec{\theta}^k} = B [q' ((\ln g)_{\vec{\theta}^j \vec{\theta}^k} - (\ln q)_{\vec{\theta}^j \vec{\theta}^k}) - u_{\vec{\theta}^j \vec{\theta}^k}] \quad (\text{A.52})$$

where  $t_{\vec{\theta}^j \vec{\theta}^k}$  has been removed since it is always zero. The expressions for the derivatives of  $\ln g$  and  $\ln q$  can be found in §A.4.3 (note that in this case, the derivatives of  $\ln g$  have  $m$  replaced with  $m_{\min}$  as well as the usual replacement in derivatives of  $\ln q$ ). The derivatives of  $u$  are:

$$u_h = -\ln(10)\beta u \quad (\text{A.53})$$

$$u_\alpha = 0 \quad (\text{A.54})$$

$$u_\beta = \frac{u \ln x + \mathcal{H}_s G1 \cdot \Gamma}{\beta} \quad (\text{A.55})$$

$$u_{h,h} = -\ln(10)\beta u_h \quad (\text{A.56})$$

$$u_{\alpha,\alpha} = 0 \quad (\text{A.57})$$

$$u_{\alpha,\beta} = 0 \quad (\text{A.58})$$

$$u_{\beta,\beta} = \frac{u_\beta \ln x + \mathcal{H}_s \bar{G} \cdot \Gamma / \beta}{\beta} \quad (\text{A.59})$$

$$u_{h,\alpha} = 0 \quad (\text{A.60})$$

$$u_{h,\beta} = -\ln(10)(u + \beta u_\beta) \quad (\text{A.61})$$

$$(\text{A.62})$$

where  $\Gamma$ ,  $G1$  and  $\bar{G}$  are here functions of  $(z+1, x)$ . Note that the formulae for the partial derivatives of  $u$  are only strictly correct at the solution (in general they involve  $\vec{\theta}'$  as well as  $\vec{\theta}$ ). The accompanying implementations calculate the more general form.

### A.5.2 Analysis of Simplest Model

For a fiducial model of  $\vec{\theta} = (14, -1.9, 0.8)$ , we calculate expected variances and correlations of and between the parameters in this simple model, assuming a  $L = 400h^{-1}\text{Mpc}$  box. Figure A.3 shows the expected variance, while figure A.4 shows the expected correlations.

A few points can be taken from this simple analysis. Firstly, the correlations between parameters are expected to be high, which as we have already seen can be problematic for parameter estimation. Secondly, a lower truncation mass is helpful in mitigating covariances, and improves variance. Thirdly, there is a tradeoff in choosing the value of  $s$ : higher values increase precision on  $\beta$ , but tend to decrease it for the other two parameters, while simultaneously increasing the correlations between them.

### A.5.3 Addition of Noise and Systematic Bias

Two simple adjustments can be made to the preceding analysis to bring it more into line with realistic datasets. The first adjustment is to introduce Poisson noise, which every dataset (whether ideal or not) will contain. Once Poisson noise is introduced, an analytic derivation is no longer possible. However, we can derive fast statistical estimates of the covariance numerically by introducing Poisson noise to the theoretical data as

$$N(m) \sim \text{Pois} \left( VA_{\rho_c} g(m | \mathcal{H}'_s, \alpha', \beta') \right). \quad (\text{A.63})$$

This can be numerically integrated and differentiated to find the Fisher information matrix. Doing so reveals the expected scatter in the determination of the parameters, due simply to Poisson noise. The expectation of the parameters remains the same, since the mean of the Poisson distribution is its parameter, and Fig. A.5 shows that the covariances stay very similar also.

Secondly, if real data is more tightly characterised by the EPS formalism fit, it is helpful to understand the expected uncertainty and *bias* of  $\vec{\theta}$  given the parameters of the fit (e.g.  $m_{\min}$ ,  $s$ ,  $V$ ). We can perform the same procedure, replacing  $N(m)$  by the Poisson-modified EPS expectation. We find that for low  $m_{\min}$ , in which the power-law regime dominates the fit, any slight bias in  $\alpha$  severely biases the estimation of  $\beta$  (see figure A.6). This effect is reduced by increasing  $s$ .

We see that the covariance estimate provided purely analytically (and the associated correlation) is quite robust to sampling noise and slight differences between the model and real data. This then becomes a useful tool to quickly explore expected covariances in the MRP parameters.

Furthermore, we see that a value of  $s = 1$  in general provides the tightest constraints across all parameters, both in the presence of noise and without it, and also serves to greatly reduce expected biasing when using realistic models.

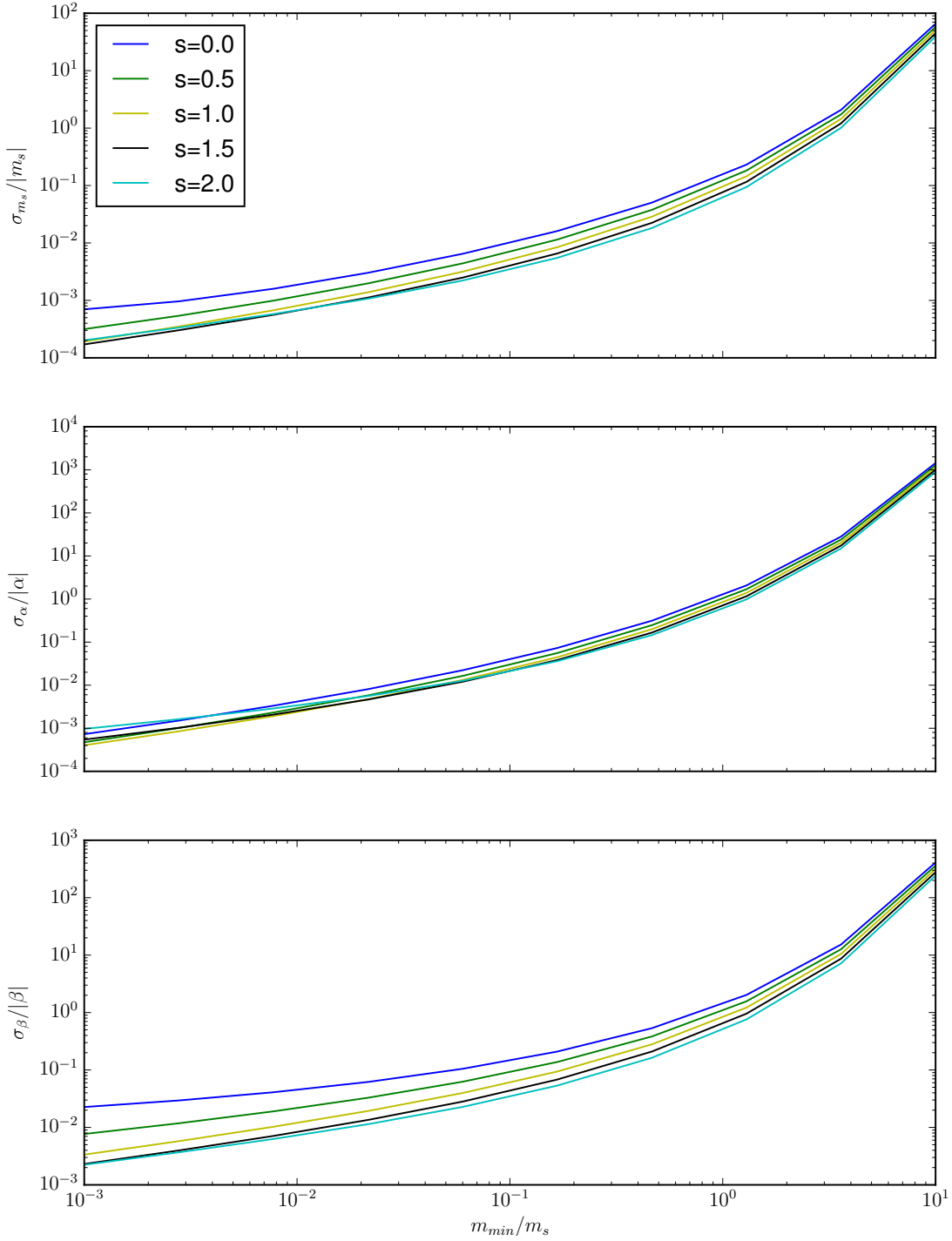


Figure A.3: Expected variance from simple fiducial analytical model. Different curves correspond to varying scale  $s$ , while the  $x$ -axis encodes the truncation scale. Noticeably, higher  $s$  increases precision for  $\beta$  (which is sensitive to high-mass halos), but decreases precision for  $\alpha$  (which is sensitive to low mass halos), but only for a low truncation scale. The overall trend is to decrease precision as the truncation scale increases, due to both decrease of information and numbers of halos. Of the parameters,  $\beta$  is the least constrained, which is to be expected, as it is probed by the fewest haloes.

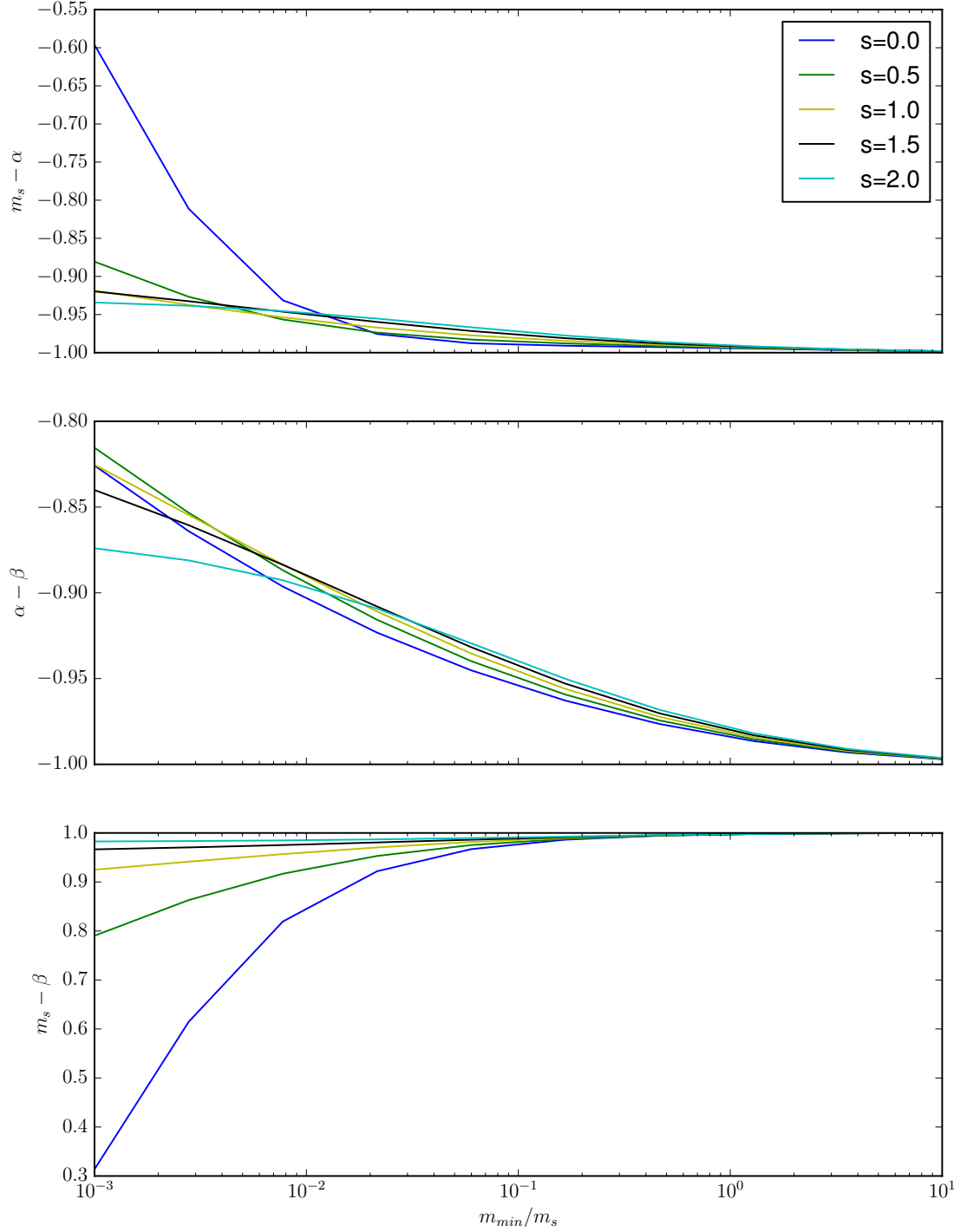


Figure A.4: Expected correlations between parameters for a simple fiducial analytical model. Line colours and  $x$ -axis are as in figure A.3. We see that lower values of  $s$  generally ‘disentangle’ parameters, so that the covariances, while still high, are somewhat reasonable. Furthermore, using lower truncation scales (compared to the turnover mass  $\mathcal{H}_s$ ), dramatically improves covariance. Even so, we expect  $\mathcal{H}_s - \alpha$  and  $\beta - \alpha$  to be quite anti-correlated, while  $\mathcal{H}_s - \beta$  is significantly correlated ( $> 0.9$ ).

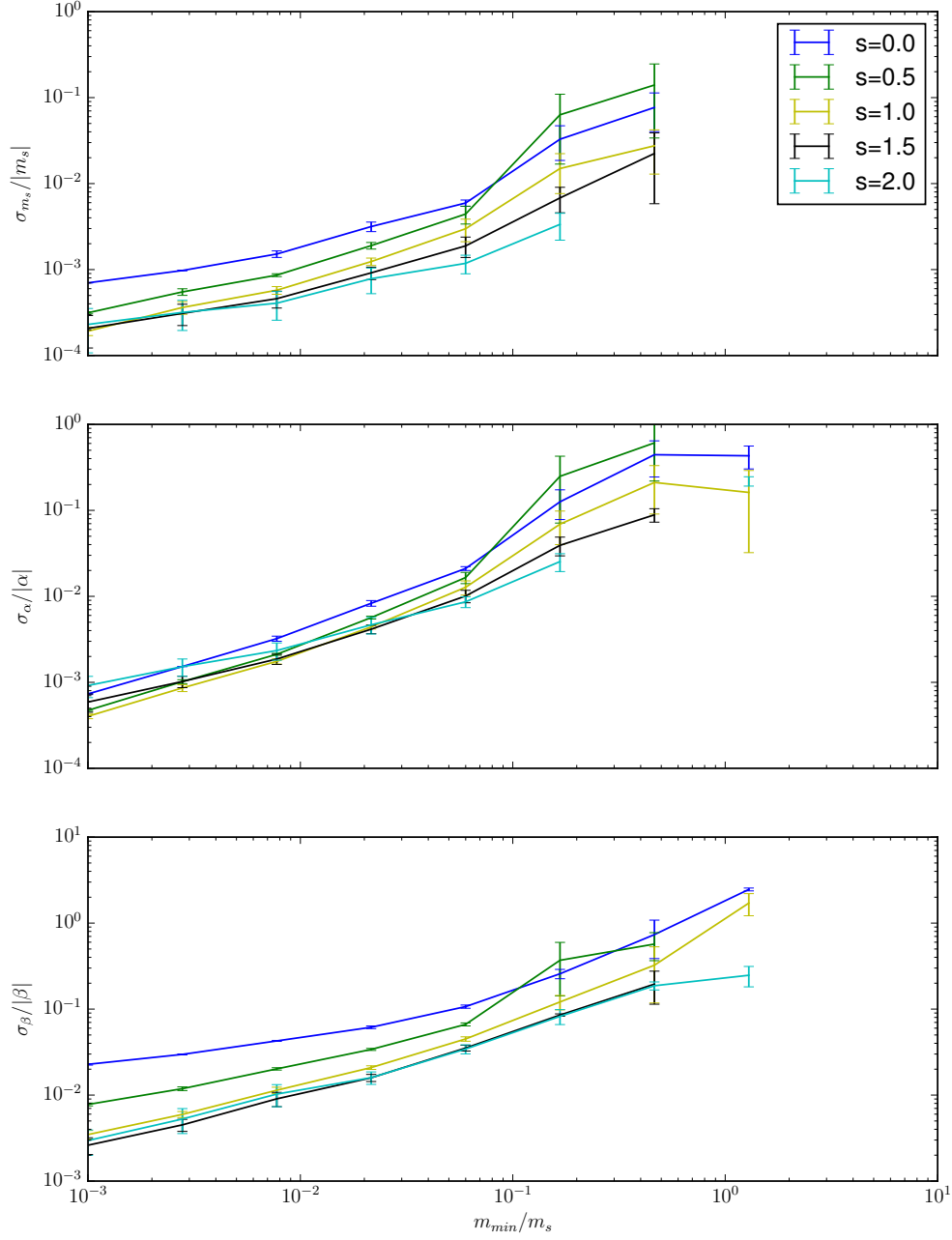


Figure A.5: Variance of parameters for sampled data. Figure corresponds precisely to Fig. A.3, except that the estimation was done from sampled masses. Some points were not able to be optimized, and therefore are left blank (towards higher truncation scale). Errorbars are derived from 20 realisations. The results are remarkably similar to the ideal analytical case, though the curves show signs of turning over at higher truncation scales.



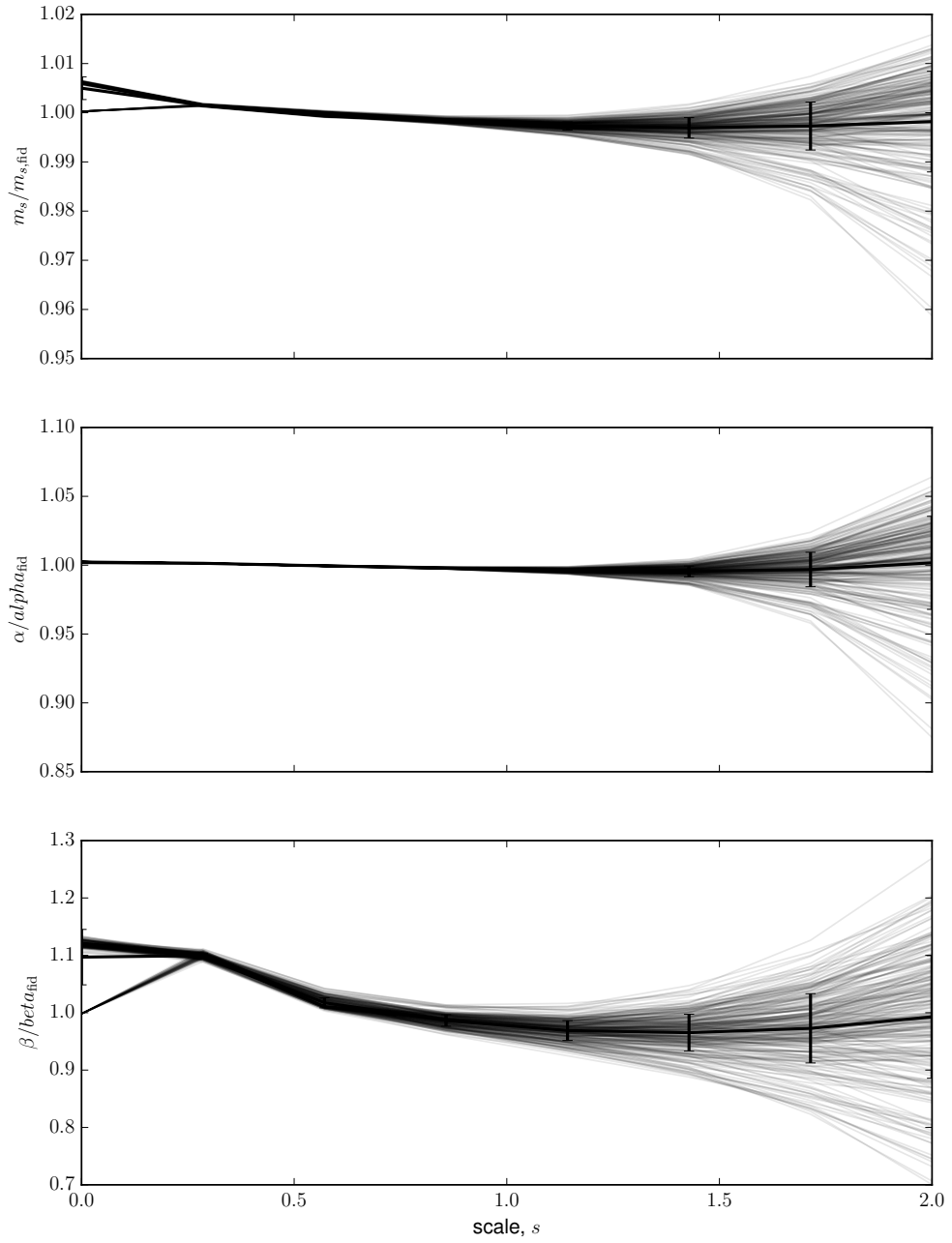


Figure A.6: Expected bias if halo masses are drawn precisely from an EPS mass function. Note the  $\sim 10\%$  bias in  $\beta$  for small scaling  $s$ . The bias for all three parameters reaches a minimum at  $s \approx 1$ , and Poisson scatter dramatically increases for  $s > 1$ .

## A.6 stan CODE

Listing A.1: Stan code for fitting MRP to masses with arbitrary measurement uncertainties

```

functions {
  /**
   * gammainc() is the upper incomplete gamma function (not regularized)
   * @param real a, shape parameter
   * @param real x, position > 0
   * @returns the non-regularized incomplete gamma
   * NOTES: uses a recursion relation to calculate values for negative a.
   */
  real gammainc(real a, real x){
    int n;
    real ap1;
    real ssum;

    if(a>=0) return gamma_q(a,x) * tgamma(a);

    ap1 <- a+1;

    //Get floor(-a)
    n<-0;
    while(n<-a){
      n <- n+1;
    }

    //Get summed part
    {
      vector[n] sums;
      for(i in 0:n-1) sums[i+1] <- pow(x,i)/tgamma(ap1+i);
      ssum <- sum(sums);
    }
    return tgamma(a)*(gamma_q(a+n,x)-pow(x,a)*exp(-x)*ssum);
  }

  /**
   * truncated_logGGD_log gives the log PDF of a variate whose exponential has a
   * lower-truncated generalised gamma distribution.
   * @param vector y, variate (should be log10(m/Hs) in the usual TGGD)
   * @param real ymin, truncation in y-space.
   * @param real alpha, power-law slope
   * @param real beta, cut-off parameter
   */
  real truncated_logGGD_log(vector y, real ymin, real alpha, real beta){
    vector[num_elements(y)] x;
    real xmin;
    real z;

    z <- (alpha+1)/beta;
  }
}

```

```

        x <- exp(log10()*y*beta);
        xmin <- exp(log10()*ymin*beta);
        return sum(log(beta) + log(log10()) + log10()*y*(alpha+1) - x - log(
            gammainc(z,xmin)));
    }
}

data {
    int<lower=0> N;                // number of halos
    vector<lower=0>[N] log_m_meas; // measure halo masses
    vector<lower=0>[N] sd_dex;    // uncertainty in halo masses (dex)

    // CONTROLS FOR PARAMETER BOUNDS
    real<lower=0> hs_min;          // Lower bound of logHs
    real<lower=0,upper=20> hs_max; // Upper bound of logHs
    real<lower=-2,upper=0> alpha_min; // Lower bound of alpha
    real<lower=-2,upper=0> alpha_max; // Upper bound of alpha
    real<lower=0> beta_min;        // Lower bound of beta
    real<lower=0> beta_max;        // Upper bound of beta
    real<lower=0> mmin_min;        // Lower bound of log_mmin
    real<lower=0> mmin_max;        // Upper bound of log_mmin
    real<lower=0,upper=20> mtrue_max; // Upper bound of true masses
}

parameters {
    real<lower=hs_min,upper=hs_max> logHs;          // Characteristic halo
    mass
    real<lower=alpha_min,upper=alpha_max> alpha;    // Power-law slope
    real<lower=beta_min,upper=beta_max> beta;        // Cut-off parameter
    real<lower=mmin_min,upper=mmin_max> log_mmin;    // Truncation mass
    vector<lower=log_mmin,upper=mtrue_max>[N] log_mtrue; // True mass estimates
}

model {
    vector[N] y;
    real ymin;
    y <- log_mtrue-logHs;
    ymin <- log_mmin-logHs;

    y ~ truncated_logGGD(ymin, alpha, beta);
    log_mtrue ~ normal(log_m_meas,sd_dex);
}

```



---

HALOMOD SUPPLEMENTS

---

## B.1 HANKEL TRANSFORM

The transformation from 3D power spectrum to real-space correlation function is given by Eq. 5.2, and is a zeroth-order spherical Hankel transformation. It is necessary to perform this transformation to calculate the two-halo correlations, and also for the one-halo term if an analytic form for the self-convolution of the density profile does not exist.

The transformation poses a challenge, as it is highly oscillatory in log-space, in which it is necessary to perform the integration. The bounds of the integration must be chosen carefully so as to terminate at a midpoint of the oscillations, while being high enough to fully capture the convergence, and low enough to enable the steps to be smaller than the oscillation period. While this can be achieved, performance drops significantly for large  $r$ , in which many steps are required to converge sufficiently.

halomod largely avoids this issue by using a technique outlined in Szapudi et al., 2005, and developed in Ogata, 2005. The technique uses a quadrature rule based on the roots of the Bessel function and a double-exponential transformation.

The general rule is defined as

$$\int_0^\infty f(x) J_\nu(x) dx \approx \pi \sum_{k=1}^\infty w_{\nu k} f(y_{\nu k}) J_\nu(y_{\nu k}) \psi'(hr_{\nu k}), \quad (\text{B.1})$$

where  $h$  is an integration step-size (similar to the trapezoidal rule) and

$$y_{\nu k} = \pi \psi(hr_{\nu k})/h \quad (\text{B.2})$$

$$\psi(t) = t \tanh(\pi \sinh(t)/2) \quad (\text{B.3})$$

$$\psi'(t) = \frac{\pi t \cosh(t) + \sinh(\pi \sinh(t))}{1 + \cosh(\pi \sinh(t))} \quad (\text{B.4})$$

$$w_{\nu k} = \frac{Y_\nu(\pi r_{\nu k})}{J_{\nu+1}(\pi r_{\nu k})}. \quad (\text{B.5})$$

Here  $J_\nu(x)$  represents a Bessel function of the first kind or order  $\nu$ , and  $Y_\nu(x)$  a Bessel function of the second kind. Lastly,  $r_{\nu k}$  are the roots of the Bessel function  $J_\nu$ , divided by  $\pi$  (alternatively, the roots of  $J_\nu(\pi x)$ ).

In the case of the zeroth-order spherical Hankel transform, we may make some simplifications. We firstly take Eq. 5.2 and make the substitution  $z = kr$ :

$$\xi(r) = \frac{1}{2\pi^2 r^3} \int_0^\infty P(z/r) z^2 j_0(z) dz, \quad (\text{B.6})$$

where  $j_0$  is the spherical Bessel function of order zero. We also note that the spherical Bessel functions are related to the standard functions by

$$j_\nu(x) = \sqrt{\frac{\pi}{2x}} J_{\nu+1/2}(x), \quad (\text{B.7})$$

so we may update Eq. B.6 to

$$\xi(r) = \frac{1}{2\pi^2 r^3} \int_0^\infty P(z/r) z^2 \sqrt{\frac{\pi}{2z}} J_{1/2}(z) dz. \quad (\text{B.8})$$

Eq. B.8 is in the form required to implement the method of Ogata, 2005, with  $\nu = 1/2$ . In this case, the zeros of the function,  $r_{\nu k} = 1, 2, 3, \dots$

We may make one further optimization by noticing that

$$w_{1/2k} = \frac{y_0(x_k)}{j_1(x_k)} = \frac{-\cos(x_k)/x_k}{\sin(x_k)/x_k^2 - \cos(x_k)/x_k} = 1, \quad (\text{B.9})$$

where the last equality holds because  $x_k = \pi r_{1/2,k}$  are defined as the roots of  $\sin(x)$ . Thus we finally arrive at the equation

$$\xi(r) \approx \pi \sum_{k=1}^{\infty} f(y_{1/2,k}) J_{1/2}(y_{1/2,k}) \psi'(hr_{1/2,k}), \quad (\text{B.10})$$

with

$$f(x) = P(x/r) x^2 \sqrt{\frac{\pi}{2x}}. \quad (\text{B.11})$$

In practice, the infinite sum is truncated at some finite value  $N$ . The benefit of this method, due to the double-exponential transformation, is that this  $N$  can be quite small for a very accurate result. Since the values,  $y_{\nu k}$ , at which the sum is evaluated approach the zeros of the bessel function as  $k \rightarrow 0$ , we find a sharp cut in the values at some  $k$ , dependent on  $h$ . Beyond this, there is no value in further iterations of the sum and so we set this  $k$  as the limit of the sum. We have found that this cut-off can be very accurately modelled as  $N = 3.2/h$ . This value of  $N$  captures all possible information while reducing random error as much as possible.

The value of  $h$  is mostly dictated by the existence of information in the integrand at low  $x$ . As expected, a smaller value of  $h$  will in general give more accurate results (modulo accumulated numerical addition error). Using our relation for  $N(h)$ , we finely increment  $h$  in the Hankel transform of the non-linear matter power at a range of scales  $r$ . Figure B.1 shows the value of  $h$  which deviates from the “correct” result (very small  $h$ ) by a percent. Though there is no real pattern here, it is clear that over all scales a value

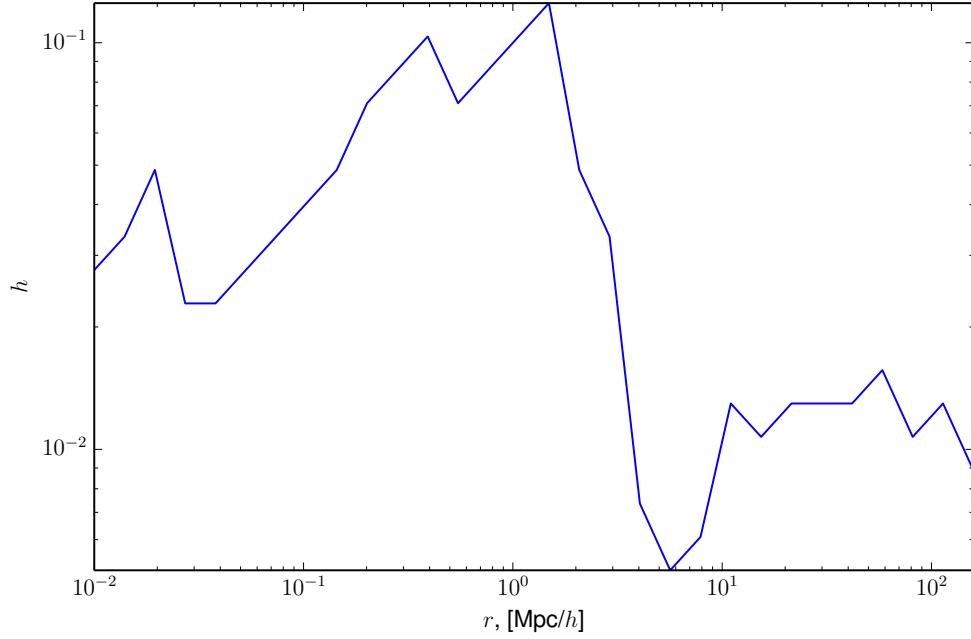


Figure B.1: Maximum value of  $h$  which yields sub-percent accuracy in the Hankel transform of the non-linear matter power, over a broad range of scales  $r$

of  $h = 0.005$  is sufficient to transform with percent accuracy. Our relation between  $N$  and  $h$  then yields  $N = 640$ .

In `halomod` we set  $N$  and  $h$  to these values to ensure better-than-percent accuracy over all reasonable scales.

## B.2 PROJECTED CORRELATION FUNCTION LIMITS

The finite upper limit can be handled empirically, keeping in mind that the correlation function may become negative at large scales (to a small degree). To perform this convergence test, we cumulatively integrate the real-space correlation function at several scales  $r_p$ , determining where the contribution to the integral becomes less than a percent. While there is no continuous relation which describes this location as a function of  $r_p$ , we deduce the following stepwise prescription:

$$r_{\max} = \max(80.5, 5r_p). \quad (\text{B.12})$$

This prescription captures the idea that for small  $r_p$ , we do not want to integrate out to negative correlation, while for larger  $r_p$ , we must integrate well past the point at which the correlation becomes negative. In practice, measurements of the projected correlation function usually set a finite upper limit of  $\sim 60h^{-1}\text{Mpc}$ , and we allow the user to set this arbitrarily if desired.

The finite lower limit has more difficult convergence properties. We make the substitution  $y = r - r_p$  in Eq. 5.69 and multiply by  $y$  to estimate the contribution of each (logarithmic) scale to the integral. We then proceed to perform an approximate analytic analysis of the convergence by assuming that as  $y \rightarrow 0$ , the correlation can be expressed

as a power law  $\xi(r) \approx r^{-a}$ . This yields the following expression for the contribution of each scale  $r$ :

$$C(y, a, r_p) = \frac{y(y + r_p)^{(1-a)}}{\sqrt{y(y + 2r_p)}}. \quad (\text{B.13})$$

This equation shows a characteristic shape, with a peak around the scale  $r_p$ . This indicates that the integral limit must be less than the scale of this peak. In fact, we deduce the scale at which the contribution is 1% of the value at the peak, to ensure sub-percent accuracy.

To do this, we solve  $\frac{dC(y, a, r_p)}{dy} = 0$  for  $y = r_{\text{peak}}$ , and calculate  $C(r_{\text{peak}}, a, r_p) = C_{\text{max}}$ . The next step is to solve  $C(y, a, r_p) = f_{\text{peak}} C_{\text{max}}$ , where  $f_{\text{peak}}$  is the fraction of the peak value we wish to locate (we set this to 0.01 in our implementation). However this is not analytically solvable. We make the valid assumption that at the solution,  $y \ll r_p$  (since the peak is around  $r_p$ ). This enables us to solve the simple equation

$$r_p^{1-a} \sqrt{y_{\text{cut}} / (2r_p)} = f_{\text{peak}} C_{\text{max}}, \quad (\text{B.14})$$

which has the solution

$$y_{\text{cut}} = \theta(a) r_p f_{\text{peak}}^2, \quad (\text{B.15})$$

where

$$\theta(a) = \frac{2^{2a+1} (-2a^3 + (p+9)a^2 - (3p+13)a + 3p+7)}{(a-1)^2 (p+2a-1) \left(\frac{p+1}{a-1}\right)^{2a}} \quad (\text{B.16})$$

$$p = \sqrt{4a^2 - 8a + 5} \quad (\text{B.17})$$

$\theta(a)$  is a weakly but monotonically decreasing function, of the order  $\approx 0.1$  in the range of  $a$  we are interested in. In `halomod`,  $y_{\text{cut}}$  is calculated for each  $r_p$ , ensuring accuracy at any scale.