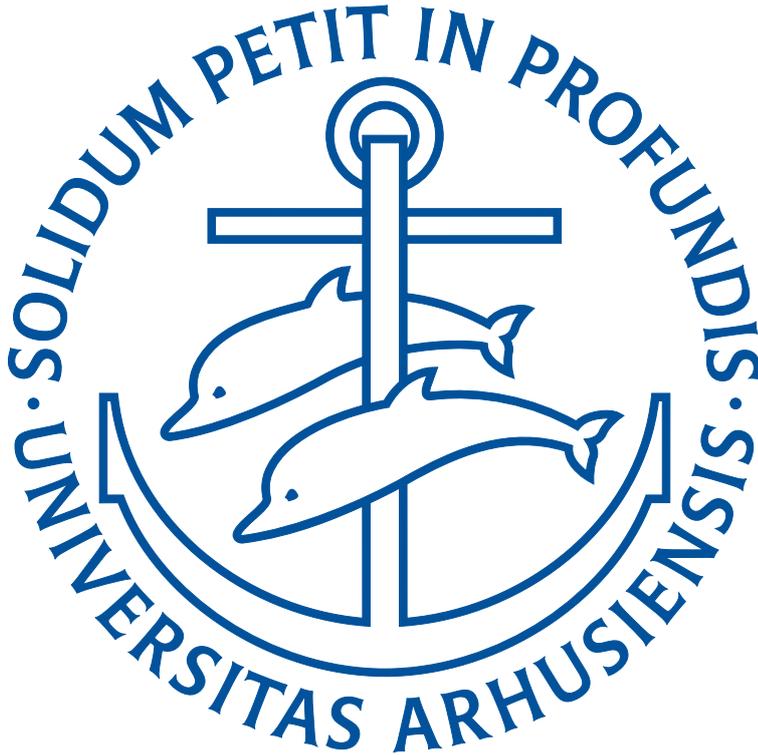


# Models of the Dark Universe Through Cosmological $N$ -body Simulations

---

CHRISTIAN SCHULTZ



SUPERVISION BY STEEN HANNESTAD

DEPARTMENT OF PHYSICS AND ASTRONOMY  
FACULTY OF SCIENCE AND TECHNOLOGY  
AARHUS UNIVERSITY  
AUGUST 2014

**PhD DISSERTATION**



# **Models of the Dark Universe Through Cosmological $N$ -body Simulations**

A DISSERTATION  
PRESENTED TO THE FACULTY OF SCIENCE AND TECHNOLOGY  
OF AARHUS UNIVERSITY  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE PhD DEGREE

BY  
CHRISTIAN SCHULTZ  
AUGUST 2014



## SUMMARY IN ENGLISH

Mankind has since ancient times looked towards the sky and wondered what it was that we saw out in the cosmos. Through time this has led to many interesting and colourful theories about *cosmology*. In this light, it may therefore be surprising to some that cosmology, as a scientific discipline, in fact, is very young. It was only with Einstein's theory of general relativity in 1916 that the mathematical foundation for modern cosmology was laid out. It is interesting now, only a hundred years later, to look back and recognise what an outstanding success modern cosmology has been.

Cosmology is by nature basic science, and broadly attempts to answer epistemological questions such as "what is our position in the cosmos" and "what *is* the universe really?". A unique trait of human nature is the constant wondering and fascination with nature to the extent that our senses permit us to experience it. In this regard cosmology plays a central role in mankind's self-perception. With scientific methods and processes we can, piece by piece, slowly map out the reality in which we exist.

It is with this in mind that the present PhD dissertation should be read. The dissertation is a small step towards deeper understanding of our physical reality, that is, what can observations tell us of the universe we inhabit. We have, e.g., realised that we live in a universe primarily made of components we cannot sense - so-called dark energy and matter. This dissertation is primarily an important step on the way to deeper insight into the nature of these mystical components. This has been done primarily through large computer simulations on the Aarhus University computer cluster, so-called  $N$ -body simulations. It is my clear intention that this PhD dissertation can serve as an introduction to my research for (graduate) students in the field of physics, also for readers without a background in astrophysics necessarily.

The dissertation starts out with a review of the scientific context in which my research has been conducted. Thereafter unfinished projects, that can potentially lead to publications in the short term, will be discussed. The last part of the dissertation is dedicated to the publications created during the PhD. Here it will be demonstrated how traditional models of warm dark matter do not work. A novel method to include neutrinos in  $N$ -body simulations will also be demonstrated. In this last part we will also compare with other published results.



## DANSK RESUMÉ

Menneskeheden har siden tidernes morgen kigget op mod himlen og undret sig over, hvad det var vi så derude i kosmos. Det har undertiden ført til mange interessante og kulørte teorier om *kosmologi*. I dette lys kan det derfor måske overraske, at kosmologi som videnskabelig disciplin faktisk er utrolig ung. Det var først med Einsteins generelle relativitetsteori i 1916, at det matematiske fundament for den moderne kosmologi blev lagt. Det er interessant nu, kun hundrede år senere, at konstatere hvilken enestående succes den moderne kosmologi har været.

Kosmologi er i sin natur fundamental grundforskning, og forsøger helt overordnet at besvare epistemologiske spørgsmål som: "Hvad er vores placering i kosmos?" og "Hvad *er* universet i det hele taget?". Et enestående træk ved menneskets væsen er den konstante undren og fascination over naturen, i den grad som vores sanser tillader os at opleve den. I den forstand spiller kosmologien en helt central rolle i menneskets selvopfattelse. Vi kan med videnskabelige metoder og processer, brik for brik, langsomt kortlægge den virkelighed i hvilken vi består.

Det er med dette in mente, at nærværende ph.d. afhandling skal læses. Afhandlingen er et lille skridt videre mod dybere forståelse af vores fysiske virkelighed, d.v.s. hvad kan observationer fortælle os om det univers vi bebor. Vi har f.eks. erkendt, at vi lever i et univers bestående primært af komponenter vi ikke kan sanse - såkaldt mørk energi og stof. Denne afhandling er først og fremmest et vigtigt skridt på vejen mod en større indsigt i disse mystiske komponents natur. Dette er gjort primært ved hjælp af store computersimuleringer kørt på Aarhus Universitets supercomputer, såkaldte *N*-body simuleringer. Det er min klare hensigt at denne ph.d. afhandling kan tjene som introduktion til min forskning for (kandidat-)studerende i fysik, også for læsere der ikke nødvendigvis har en baggrund i astrofysik.

Afhandlingen lægger ud med en redegørelse for den videnskabelige sammenhæng, hvori min forskning er skabt. Dernæst vil ikke-afsluttede projekter, der potentielt kan føre til publikationer på kort sigt, blive beskrevet. Den sidste del af afhandlingen er dedikeret til publikationerne skabt i løbet af ph.d.-studiet. Det vil her blive demonstreret, hvordan traditionelle modeller for varmt mørkt stof ikke virker. En ny metode til at inkludere neutrinoer i *N*-body simuleringer vil også blive demonstreret. I denne sidste del vil vi også sammenligne med andre publicerede resultater.



# CONTENTS

<b>Summary in English</b>	<b>i</b>
<b>Dansk resumé</b>	<b>iii</b>
<b>Introduction</b>	<b>1</b>
<b>I Theoretical background</b>	<b>5</b>
<b>1 Cosmological framework</b>	<b>7</b>
1.1 Friedmann-Lemaître-Robertson-Walker cosmology . . . . .	7
1.2 Cosmological perturbations . . . . .	10
1.3 The power spectrum . . . . .	14
1.4 Inflation and initial conditions for the universe . . . . .	16
1.5 Cosmological models . . . . .	18
1.5.1 Tensions in $\Lambda$ CDM . . . . .	18
1.5.2 $\Lambda$ WDM . . . . .	20
1.6 Gravitational lensing in cosmology . . . . .	23
1.6.1 Weak lensing . . . . .	24
1.6.2 Observing weak lensing . . . . .	27
<b>2 <i>N</i>-body simulations</b>	<b>29</b>
2.1 The TreePM algorithm . . . . .	29
2.2 Parallelisation . . . . .	31
2.3 Timestepping . . . . .	32
2.4 The equations of motion in GADGET-2 . . . . .	33
2.5 Halo finding in <i>N</i> -body simulations . . . . .	35
2.5.1 Amiga's Halo Finder . . . . .	36
2.5.2 Rockstar . . . . .	36
2.5.3 Halo finding in warm dark matter cosmologies . . . . .	37

<b>II</b>	<b>Research in progress</b>	<b>41</b>
<b>3</b>	<b>An algorithm for <math>N</math>-body ray-tracing</b>	<b>43</b>
3.1	The algorithm . . . . .	44
3.2	Current state and outlook . . . . .	47
<b>4</b>	<b>Dynamical dark energy in <math>N</math>-body simulations</b>	<b>49</b>
4.1	A relativistic fluid . . . . .	50
4.2	Implementation in GADGET-2 . . . . .	52
4.3	Halofit . . . . .	53
4.4	Current state . . . . .	54
<b>III</b>	<b>Scientific research and publications</b>	<b>55</b>
<b>5</b>	<b>Warm dark matter</b>	<b>57</b>
5.1	Current constraints on warm dark matter . . . . .	58
5.2	Introduction to the paper . . . . .	60
5.3	The future of warm dark matter . . . . .	75
<b>6</b>	<b>Neutrinos in <math>N</math>-body simulations - a novel approach</b>	<b>77</b>
6.1	Neutrinos in $N$ -body simulations . . . . .	78
6.2	Introduction to the paper . . . . .	81
6.3	The future and outlook . . . . .	95
<b>IV</b>	<b>Outlook</b>	<b>97</b>
<b>7</b>	<b>Conclusion and outlook</b>	<b>99</b>
<b>8</b>	<b>Acknowledgements</b>	<b>101</b>
	<b>Bibliography</b>	<b>103</b>

# INTRODUCTION

The discipline of physical cosmology (or cosmology for short) is the branch of physics dealing with the *cosmos*, that is, the ordered whole. Cosmology borrows much from many other fundamental physical fields such as particle physics or statistical mechanics depending on the scales of interest. One particular theory stands out, however, as the scientific discipline of cosmology was arguably born with Einstein's theory of general relativity. Although the correct equations governing our universe as a whole were quickly found as a consequence of Einstein's equations, it was only with the discovery that the universe is expanding that the true marvellous nature of our universe was unveiled.<sup>a</sup> For a long time there was no consensus on what type of universe we inhabited, e.g., Einstein postulated a static and compact variant dominated by matter whereas de Sitter hypothesised a flat universe dominated by the "cosmological constant". The issue of the origin of the universe also caused controversy, with supporters of a steady state model with continuous matter creation and an infinitely old universe on one side, and supporters of a Big Bang model with an initial singularity and a finite age on the other side. For a long time this debate divided the scientific community. It was only when new technology in the nineteen sixties allowed observation of the cosmic microwave background, that the Hot Big Bang model became consensus. This marked an important paradigm shift in the field of cosmology. Kragh (1999) gives an excellent account of the history of the field of physical cosmology.

Studying the zigzag history of cosmology teaches us that, among the plethora of mathematically allowed universes, we need precise observations to discern theoretical models governing radically different universes. Due to the obvious fact that we only have a single universe to observe, it is necessary to map how different theoretical models differ when measured with physical observables like, e.g, the galaxy density or matter content. We must constantly strive to falsify theories and constrain the allowed parameter space of competing models based on observations. Fortunately for the curious scientist, cosmology has experienced a renaissance in the last twenty-five years owing to high precision observations. Noticeably the COBE, WMAP and Planck satellites have given us a wealth of information on the fundamental parameters of the

---

<sup>a</sup>Credit for this discovery should go to Lemaître and to a smaller extent Friedmann, and not (only) to Hubble.

universe from observations of the cosmic microwave background. The Hubble Space Telescope, which could easily be called one of the wonders of the twentieth century, has, among many things, given us insight into the evolution of galaxies from galaxy observations. Supernova observations have established that, not only do we live in a universe that is expanding, the expansion is in fact accelerating. These observations of the cosmic microwave background, distant galaxies and supernovae have brought us much closer to answering the great questions of cosmology. Before these observations the questions were: "Does dark matter and dark energy exist? Was there an inflationary epoch?". Today the questions are rather: "What *is* dark matter and dark energy? What drove inflation and when did it end?".

While the observations of the last twenty-five years has contributed greatly to the development of the field of cosmology, the importance of fast, modern computers used in, e.g., data reduction and model prediction cannot be underestimated. Together with enhanced numerical procedures, these computers have allowed analysis of complicated systems with fewer simplifying assumptions, and has revolutionized not only the field of cosmology, but also theoretical physics as a whole. Once the weapons of choice of a theoretical physicist were the pen and the paper, now it would be fair to say they are on equal footing with the keyboard and the computer. It is on this numerical side of modern cosmology that this present dissertation is a contribution.

Modern science advances in small steps towards answering the great questions of our time, and in cosmology one of the most important questions is: What is dark matter and dark energy? While the definitive answer did not make it into this dissertation,<sup>b</sup> the research presented here should be viewed as an incremental step toward a better understanding of the physical nature of the dark universe. This has been done partly by examining the parameter space currently allowed by observations, and also by providing efficient and convenient numerical methods to estimate the impact of dark components. The research touches, among other things, on the subjects of beyond-the-Standard-Model warm dark matter and Standard Model hot dark matter in the form of neutrinos. By far, the main emphasis of the material presented is on the impact of these models on structure formation through  $N$ -body simulations and not of the phenomenological nature of the dark matter.

The bulk of the work behind the results presented here has been on the development and analysis of  $N$ -body simulations. We will also present some of the on-going work that were not finished by the deadline of this dissertation: An algorithm for ray-tracing through  $N$ -body simulations and a framework for presenting the impact of dynamical dark energy models on structure formation.<sup>c</sup> Throughout the disserta-

---

<sup>b</sup>Due to time constraints of course.

<sup>c</sup>A major factor in the unfinished state of this research is the termination of the PhD 3 months prior to normal completion.

tion, emphasis has been given to provide references expanding upon the details of the theoretical framework. This is done with the hope of guiding readers interested in building upon the conclusions of the research presented here in the right direction. Hopefully, the presentation here can serve as a convenient and relatively short introduction to the scientific context. Since the focus of the research behind this dissertation has been on  $N$ -body simulations, a considerable amount of code has been produced. This code has to an extent been ordered into structured units that should be relatively easy to pick up from. The code is not explicitly public yet, however, should anyone be interested in using it, access to the code base will be given.

The dissertation, presenting the work conducted during a four-year PhD position of the four-plus-four program of Aarhus University, is structured into three major parts and one smaller part summing up the conclusions. The first part provides an introduction to the theoretical framework necessary to understand the context of the other parts. It has been split into two chapters detailing the relevant mathematical details and the numerical details respectively. The second part provides an introduction to un-finished research into which a significant amount of time was invested. Some of this work will likely result in an additional publication shortly. The third part contains published research conducted during the PhD, along with the relevant publications. These publications are (see the bibliography section at the end for the full references)

- *S. Hannestad, T. Haugbolle, and C. Schultz.* Neutrinos in Non-linear Structure Formation - a Simple SPH Approach.
- *C. Schultz, J. Oñorbe, K. N. Abazajian, and J. S. Bullock.* The high- $z$  universe confronts warm dark matter: Galaxy counts, reionization and the nature of dark matter.

The publications are preceded by a discussion of the scientific context in which they should be viewed as contributions. The last part of the dissertation summarises the results. Throughout the text SI units will be employed unless otherwise specified.



## **Part I**

# **Theoretical background**



# COSMOLOGICAL FRAMEWORK

This chapter introduces some formalism and the mathematical constructs necessary to understand the context of later chapters. References have been liberally added in order to point the interested reader in the right direction. Throughout the chapter it has been attempted to conform to the standard notation and definitions in the works quoted, to the extent that this is possible. We note that the framework presented here is not done in the usual natural  $\hbar = c = 1$  unit system, but explicitly inserts factors of  $\hbar$  and  $c$  where appropriate. This unconventional choice is because  $N$ -body simulations generally do not assume natural units.

## 1.1 Friedmann-Lemaître-Robertson-Walker cosmology

The Einstein equation

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu}, \quad (1.1)$$

relates the geometry of the universe to its energy-content and dynamics. The units of the Einstein equation are

$$\begin{aligned} [G_{\mu\nu}] &= \frac{1}{\text{length}^2} \\ [T_{\mu\nu}] &= \frac{\text{energy}}{\text{length}^3} \\ [\Lambda] &= \frac{1}{\text{length}^2} \\ [g_{\mu\nu}] &= 1. \end{aligned}$$

The Friedmann-Lemaître-Robertson-Walker (FLRW) metric is the metric of a homogeneous and isotropic spacetime manifold. Isotropy, qualitatively, means that any observable property is independent of the direction of observation. Homogeneity means that any observable quantity is independent of the observer. Wald (1984) gives a rigorous definition of isotropy and homogeneity in a general relativistic sense. It can easily be shown that the requirement of homogeneity and isotropy requires the FLRW metric to describe a manifold of constant curvature. The FLRW metric can then take one of 3 forms

$$\begin{aligned} ds^2 &= -a^2(\eta)d\eta^2 + a^2(\eta) \begin{cases} d\chi^2 + \left(\frac{\sin(\sqrt{K}\chi)}{\sqrt{K}}\right)^2 (d\theta^2 + \sin^2\theta d\phi^2) \\ d\chi^2 + \chi^2(d\theta^2 + \sin^2\theta d\phi^2) \\ d\chi^2 + \left(\frac{\sinh(\sqrt{-K}\chi)}{\sqrt{-K}}\right)^2 (d\theta^2 + \sin^2\theta d\phi^2) \end{cases} \\ &\equiv a(\eta) (-d\eta^2 + d\chi^2 + f_K^2(\chi)d\Omega^2), \end{aligned} \quad (1.2)$$

corresponding respectively to a spacetime with constant positive curvature (a 3-sphere), zero curvature (a plane) or constant negative curvature (a hyperboloid). The coordinate  $\eta$  is called the conformal time<sup>a</sup>, the physical or cosmic time  $t$  experienced by a comoving observer is given by  $d\eta = c/ad t$ . The parameter  $a$  is called the scale factor since it stretches the metric by a time dependent factor, and it is normalised so that  $a = 1$  today. The curvature constant  $K$  is simply the inverse curvature radius of the manifold

$$K = \frac{\kappa}{R_0^2}, \quad (1.3)$$

where  $\kappa = \{-1, 0, 1\}$  is the sign of the curvature. This establishes the geometrical part of the Einstein equation. The other part of the equation - the dynamical part - is

<sup>a</sup>The name implies that in these coordinates, the flat FLRW metric is conformally equivalent to the non-expanding Minkowski spacetime, that is, the causal structure is the same. In particular, the FLRW metric is always conformally equivalent to the Minkowski spacetime. See Ibison (2007) for the explicit conformal transformations.

the stress-energy tensor. The most general form consistent with homogeneity and isotropy is

$$T_{\mu\nu} = \rho u_\mu u_\nu + P \left( g_{\mu\nu} + \frac{u_\mu u_\nu}{c^2} \right), \quad (1.4)$$

which incorporates the contents of the universe as a perfect (non-interacting) fluid.  $\rho$  is the matter-density,  $P$  is the pressure and  $u$  the spacetime tangent (four-velocity).

The time-time component of the Einstein equation of the FLRW universe yields the Friedmann equation which governs the evolution of the universe as a whole

$$H^2 \equiv \left( \frac{da/dt}{a} \right)^2 = \frac{8\pi G}{3c^2} \epsilon - \frac{Kc^2}{a^2} + \frac{\Lambda c^2}{3}.$$

This can be put into the form

$$\left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3c^4} a^2 \epsilon - K + a^2 \frac{\Lambda}{3}. \quad (1.5)$$

The  $\rho c^2 = \epsilon$  term is the energy density of the universe, the  $K$  term is the curvature term from eq. (1.3). Note that the dot in  $\dot{a}$  means derivative with respect to conformal time. The last constant term is the famous cosmological constant term of profound importance for cosmology, but often neglected in other general relativistic theories. It is sometimes absorbed into the  $\epsilon$  term and considered a vacuum energy density. The space-space component of the Einstein equation gives the acceleration equation

$$\frac{d}{d\eta} \left( \frac{\dot{a}}{a} \right) = -\frac{4\pi G}{3c^4} (\epsilon + 3P). \quad (1.6)$$

Alternative versions of the Friedmann equation are also commonly employed. Define

$$H_0 \equiv H|_{\eta=\eta_0}, \quad (1.7)$$

where  $\eta_0$  is the current conformal time. Further define the critical density

$$\epsilon_{cr} = \frac{3c^2 H^2}{8\pi G} \quad (1.8)$$

and the density parameters

$$\Omega \equiv \frac{\epsilon}{\epsilon_{cr}}, \quad \Omega_\Lambda \equiv \frac{\Lambda}{3H^2} \quad (1.9)$$

$$\Omega_{,0} \equiv \frac{\epsilon}{\epsilon_{cr}} \Big|_{\eta_0}, \quad \Omega_{\Lambda,0} \equiv \frac{\Lambda}{3H^2} \Big|_{\eta_0}. \quad (1.10)$$

The Friedmann equation can then be cast into the form

$$K = \frac{a^2 H^2}{c^2} (\Omega + \Omega_\Lambda - 1), \quad (1.11)$$

and thus the curvature term  $K$  can easily be expressed in terms of the present value of the density parameters:

$$K = \frac{H_0^2}{c^2} (\Omega_{,0} + \Omega_{\Lambda,0} - 1). \quad (1.12)$$

## 1.2 Cosmological perturbations

As is evident from everyday experience, the universe is not completely isotropic and homogeneous. Therefore the above framework provides the background in which one can do perturbation theory (which naturally will break down on scales that are relevant to everyday life). The natural assumption is the cosmological principle: The universe is, on average on very large scales  $\gtrsim 150 \text{ Mpc } h^{-1}$  (Marinoni et al., 2012), homogeneous and isotropic. The basic cosmological perturbation theory then considers, like any perturbation theory, small deviations from this average. Perturbation theory describes the largest scales, but will eventually break down once the deviations from homogeneity and isotropy become sufficiently large. The presentation in this section closely follows that in Bertschinger (2001) and Ma and Bertschinger (1995). We will in this section only consider a spatially flat metric, although it is easily generalizable to curved a curved metric by substituting  $\delta_{ij}$  with the metric on the corresponding three-space  $\gamma_{ij}$  (Bertschinger, 1993).

The starting point of the perturbation approach is the perturbed metric

$$ds^2 = a^2(\eta) \left[ -d\eta^2 + \delta_{ij} dx^i dx^j + h_{\mu\nu} dx^\mu dx^\nu \right], \quad (1.13)$$

in which the tensor  $h_{\mu\nu}$  (or more precisely  $h_{\mu\nu}/a^2$ ) represents the perturbation. In the perturbation approach we then regard  $h_{\mu\nu}$  as a tensor field residing on the background FLRW metric.

The full metric tensor  $g_{\mu\nu}$  is a symmetric tensor describing four-dimensional space-time and has ten independent components. By extension the perturbation tensor  $h_{\mu\nu}$  can also have at maximum ten independent components. Lifshitz (1946) presented a decomposition of the perturbation tensor into irreducible representations of the rotation group. In general  $h_{\mu\nu}$  can be written

$$h_{00} = -2\psi, \quad h_{0i} = w_i^\parallel + w_i^\perp, \quad h_{ij} = 2(\phi\delta_{ij} + S_{ij}^\parallel + S_{ij}^\perp + S_{ij}^T). \quad (1.14)$$

Here  $h_{0i}$  is split into a curl-free ( $w_i^\parallel$ ) and divergence-free ( $w_i^\perp$ ) vectorfield, and  $h_{ij}$  is split into a trace part ( $\phi$ ) and a traceless tensor part ( $S_{ij}$ ). The tensor  $S_{ij}$  is further decomposed into a doubly longitudinal part ( $S_{ij}^\parallel$ ), a longitudinal-transverse part ( $S_{ij}^\perp$ )

and a doubly transverse part ( $S_{ij}^T$ ). Remembering that a curl-free vectorfield and the doubly longitudinal part of a traceless, symmetric tensor can be obtained from scalar fields<sup>b</sup> we can identify four scalars in eq. (1.14):  $(\psi, w_i^\parallel, \phi, S_{ij}^\parallel)$ . Likewise we can identify two vectors:  $w_i^\perp, S_{ij}^T$ . The decomposition into longitudinal and transverse components for both the vector mode and tensor mode are naturally coordinate dependent, and one can eliminate additional four degrees of freedom by choosing a gauge (since spacetime is four-dimensional choosing a gauge eliminates four degrees of freedom). Thus we are left with only six physically independent quantities which can be shown to be: Two scalars, one divergenceless vector and one traceless, symmetric and doubly transverse tensor (but naturally any linear combination of the above can be used). With an appropriate choice of gauge the metric can be written

$$ds^2 = a^2(\eta) \left[ -(1 + 2\psi) d\eta^2 + w_i^\perp d\eta dx^i + \left( (1 - 2\phi) \delta_{ij} - 2S_{ij}^T \right) dx^i dx^j \right]. \quad (1.15)$$

The two scalar modes  $\psi$  and  $\phi$  are, as the name implies, scalars under rotations and physically corresponds to perturbations of the (relativistic) Newtonian potential. As such, by virtue of the Poisson equation, they produce density perturbations and are the quantities relevant for structure formation.

The vector mode  $w_i^\perp$  corresponds to gravitomagnetism (also called frame-dragging or the Lense-Thirring effect), and has only decaying solutions so this component can be ignored.

The tensor mode  $S_{ij}^T$  corresponds to gravitational radiation, and the two transverse components are the two possible polarizations for the gravitational waves. These tensor modes are potentially important for constraining inflation models of the early universe. Since gravitational waves are physical observables they are gauge invariant.

By inspection of the Einstein equations it can be seen that the scalar, vector and tensor modes evolve independently, see Bertschinger (1993) for (some of) the gory details. We can thus freely restrict our attention to the scalar modes alone, and this is particularly handy since they source the structure formation (and since the vector mode decays it is not of cosmological interest). Furthermore, the tensor mode will propagate without feeling the changes in the scalar potentials (at the linear level. Non-linear effects like dust can contaminate the signal). Restricting our attention to only the scalar modes eq. (1.15) gives a diagonal metric known as the conformal Newtonian gauge (also known as longitudinal gauge)

$$ds^2 = a^2(\eta) \left[ -(1 + 2\psi) d\eta^2 + (1 - 2\phi) \delta_{ij} dx^i dx^j \right]. \quad (1.16)$$

If the energy-momentum tensor is isotropic (i.e. no anisotropic stress) then  $\psi = \phi$ , and in the weak-field limit  $\psi c^2 = \phi c^2 = \Phi$  plays the role of the Newtonian potential.

<sup>b</sup>  $w^\parallel = \nabla \alpha$  and  $S_{ij}^\parallel = (\nabla_i \nabla_j - \frac{1}{3} \delta_{ij} \nabla^2) \beta$  for suitable scalars  $\alpha, \beta$  respectively

The conformal Newtonian gauge is often used exactly because of its relation to the weak-field metric (hence its name)

$$ds^2 = -(1 + 2\Phi)c^2 dt^2 + (1 - 2\Phi)r^2 d\Omega^2. \quad (\text{Weak-field metric})$$

For convenience we also introduce some easy notation for the components of the stress-energy tensor. Define

$$\bar{\rho}\delta \equiv \delta\rho = \rho - \bar{\rho} = -\delta T^0_0/c^2, \quad (1.17)$$

$$(c^2\bar{\rho} + \bar{P})\theta \equiv ik^j\delta T^0_j, \quad (1.18)$$

$$\delta P \equiv \delta T^i_i, \quad (1.19)$$

$$(c^2\bar{\rho} + \bar{P})\sigma \equiv -(\hat{k}_i\hat{k}_j - \frac{1}{3}\delta_{ij})\Sigma^i_j \quad \text{where} \quad \Sigma^i_j \equiv \delta T^i_j - \delta^i_j\delta T^k_k/3, \quad (1.20)$$

where  $\delta T^i_j$  is perturbed stress-energy tensor that encapsulates the microphysics of the individual components. In the conformal Newtonian gauge the Einstein equations then takes the form

$$k^2\phi + 3\frac{\dot{a}}{a}\left(\dot{\phi} + \frac{\dot{a}}{a}\psi\right) = -\frac{4\pi G a^2}{c^2}\delta\rho, \quad (1.21)$$

$$k^2\left(\dot{\phi} + \frac{\dot{a}}{a}\psi\right) = \frac{4\pi G a^2}{c^2}(\bar{\rho} + \bar{P}/c^2)\theta, \quad (1.22)$$

$$\ddot{\phi} + \frac{\dot{a}}{a}(\dot{\psi} + 2\dot{\phi}) + \left(2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2}\right)\psi + \frac{k^2}{3}(\phi - \psi) = \frac{4\pi G a^2}{3c^2}\delta P/c^2, \quad (1.23)$$

$$k^2(\phi - \psi) = \frac{12\pi G a^2}{c^2}(\bar{\rho} + \bar{P}/c^2)\sigma, \quad (1.24)$$

Clearly  $\phi = \psi$  in the case of no anisotropic stress ( $\sigma = 0$ ) as noted above. No assumptions on the stress-energy tensor has been made up to this point, and the equations are therefore completely general within the gauge choice. The perturbed stress-energy tensor naturally depends on the choice of the unperturbed stress-energy tensor. For baryons and dark matter it is fitting to model it as that of a perfect fluid as in eq. (1.4). The continuity and Euler equations known from fluid dynamics can be obtained by invoking energy-momentum conservation on the stress-energy tensor  $T^{\mu\nu}_{;\mu} = 0$ . Rearranging terms and using the Einstein equations above then gives

$$\dot{\delta} = -(1+w)(\theta - 3\dot{\phi}) - 3\frac{\dot{a}}{a}\left(\frac{1}{c^2}\frac{\delta P}{\delta\rho} - w\right)\delta, \quad (1.25)$$

$$\dot{\theta} = -\frac{\dot{a}}{a}(1-3w)\theta - \frac{\dot{w}}{1+w}\theta + \frac{1}{c^2}\frac{\delta P/\delta\rho}{1+w}k^2\delta - k^2\sigma + k^2\psi. \quad (1.26)$$

$\delta, \theta$  and  $\sigma$  are the fluid density perturbation, fluid velocity perturbation and anisotropic stress respectively as defined previously, and  $w c^2 = P/\rho$  is the equation of state. Normally one can define a sound speed  $c_s^2 = dP/d\rho$  such that the equations simplify in

the case of adiabatic perturbations. This, of course, is only applicable to collisional components such as photons and baryons. Dark matter, being collisionless, has neither pressure nor sound speed.

An additional equation for the stress  $\sigma$  is needed to close the system. It should be noted that if the components interact (by other means than gravity, such as in the photo-baryon plasma in the early universe) the above equations should be modified by a full phase-space description as we will do shortly.

In the non-relativistic limit for a collisional gas ( $c_s \ll 1$ ,  $w \simeq 0$ ) the continuity and Euler equations may be written

$$\dot{\delta} = -\theta - 3\dot{\phi} \quad (1.27)$$

$$\dot{\theta} = -\frac{\dot{a}}{a}\theta + \frac{c_s^2}{c^2}k^2\delta + k^2\psi. \quad (1.28)$$

The acoustic term  $\frac{c_s^2}{c^2}k^2\delta$  is important for sufficiently large  $k$ . The shear term  $k^2\sigma$ , on the other hand, is far smaller and has as such been neglected. This would not be valid in the early universe with coupled baryons and photons, as there then would be a transfer of momentum and energy between these two terms.

The stress-energy tensor for components such as neutrinos and photons is given in terms of their full phase-space description

$$T_{\mu\nu} = \frac{c}{\hbar^3} \int dP_1 dP_2 dP_3 (-g)^{-1/2} \frac{P_\mu P_\nu}{P^0} f(x^i, P_j, \eta), \quad (1.29)$$

where  $g$  is the determinant of the metric tensor,  $P_\mu$  is the four-momentum, and  $f$  the phase-space distribution function. Perturbations in the stress-energy tensor are then given as perturbations of the distribution function. It is therefore convenient to expand it from its equilibrium form (Bose-Einstein, Fermi-Dirac etc.)

$$f = f_0 + \frac{\partial f_0}{\partial T} \delta T = f_0(1 + \Psi). \quad (1.30)$$

In the conformal Newtonian gauge the four-momentum is related to the proper momentum  $p_i$  by

$$P_i = a(1 - \phi)p_i. \quad (1.31)$$

For convenience the comoving proper momenta  $q$  is defined as  $q_j \equiv ap_j$ . Since  $p_i \propto a^{-1}$  this means that  $q$  is constant as the universe expands. We also define  $\epsilon = \sqrt{q^2 c^2 + a^2 c^4 m^2}$  as the energy a comoving observer would measure. With the definition in eq. (1.29) the stress-energy tensor can then be written in terms of  $f_0$  and

$\Psi$

$$T^0_0 = \frac{-1}{a^4 \hbar^3} \int [\epsilon f_0(q)(1 + \Psi)] q^2 dq d\Omega \quad (1.32)$$

$$T^0_i = \frac{1}{a^4 \hbar^3} \int [c q n_i f_0(q) \Psi] q^2 dq d\Omega \quad (1.33)$$

$$T^i_j = \frac{1}{a^4 \hbar^3} \int \left[ \frac{c^2 q^2 n_i n_j}{\epsilon} f_0(q)(1 + \Psi) \right] q^2 dq d\Omega \quad (1.34)$$

where  $n_i, n_j$  are unit-vectors defined by  $q_j = q n_j$ . The Boltzmann equation states that the Liouville operator equals the collision operator for the distribution function

$$\hat{\mathbf{L}}[f] = \mathbf{C}[f] \quad (1.35)$$

Utilising the covariant Liouville operator in the variables  $(x^i, q, n_j, \eta)$  the Boltzmann equation can then be written in the form

$$\frac{\partial \Psi}{\partial \eta} + ic \frac{q}{\epsilon} (\vec{k} \cdot \hat{n}) \Psi + \frac{d \ln f_0}{d \ln q} \left[ \phi - ic \frac{\epsilon}{q} (\vec{k} \cdot \hat{n}) \psi \right] = \frac{1}{f_0} \left( \frac{\partial f}{\partial \eta} \right)_C. \quad (1.36)$$

By solving this equation with some form of interaction term  $\left( \frac{\partial f}{\partial \tau} \right)_C$  the stress-energy tensor can be calculated.

### 1.3 The power spectrum

A homogeneous random field  $\mathcal{F}$  has the property that it is independent of translations  $\mathcal{F}(\vec{x}) = \mathcal{F}(\vec{x} + \vec{y})$ . An isotropic random field has the property that it is independent of any rotation  $\mathcal{R}$ ,  $\mathcal{F}(\vec{x}) = \mathcal{F}(\mathcal{R}\vec{x})$ . Naturally, the density contrast  $\delta$  obeys these properties in a FLRW universe, and thus the autocorrelation function of the density field  $\delta$  (also known as the two-point correlation function) depends only on the distance scale between points

$$\xi(r) \equiv \langle \delta(\vec{x}) \delta(\vec{x} + r) \rangle_x, \quad (1.37)$$

where  $\langle \rangle_x$  signifies the (volume) average over all  $x, y$  such that  $|x - y| = r$ . A Gaussian field is fully characterized by its autocorrelation function. The Fourier transform of  $\xi$  is

$$\xi(r) = \frac{V}{(2\pi)^3} \int |\delta_{\mathbf{k}}|^2 e^{-i\mathbf{k} \cdot \mathbf{r}} d^3 k. \quad (1.38)$$

With this expression for  $\xi$  we define the power spectrum of the density field as

$$P(k) \equiv \langle |\delta_{\mathbf{k}}|^2 \rangle = \langle |\delta_k|^2 \rangle. \quad (1.39)$$

where we have used that the density field is isotropic. From this definition it is clear that the power spectrum  $P$  is a measure of the root-mean-square of the density field on a scale  $k$ , and with the convention of the Fourier transform above, it is dimensionless. A convenient rescaling of the power spectrum gives the dimensionless power spectrum

$$\Delta^2 \equiv \frac{V}{(2\pi)^3} 4\pi k^3 P(k) \quad (1.40)$$

In some works the volume factor is absorbed into the power spectrum itself giving it the dimension of volume (and hence the name the dimensionless power spectrum for  $\Delta$ ). Physically  $\Delta$  is the contribution to the fractional density variance per unit interval in  $\ln k$ . Often the power spectrum is utilised to measure the variance in a certain geometry by convolving  $\xi$  with a filter function  $W_R$ . A common filter function is that of the spherical top-hat for some distance scale  $R$

$$W_R(r) = \frac{1}{V_W} \begin{cases} 1, & \text{if } r < R \\ 0, & \text{if } r > R \end{cases}$$

with  $V_W = 4\pi R^3/3$ . The Fourier transform of the convolution is then

$$\sigma_R^2 = \frac{V}{(2\pi)^3} \int |\delta_k|^2 \bar{W}_R^2 d^3 k = \frac{V}{(2\pi)^3} \int |\delta_k|^2 3^2 \left[ \frac{\sin kR}{(kR)^3} - \frac{\cos kR}{(kR)^2} \right]^2 d^3 k. \quad (1.41)$$

In terms of the dimensionless power spectrum this can be written

$$\sigma_R^2 = \int \Delta^2(k) \bar{W}_R d \ln k = 9 \int \Delta^2(k) \left[ \frac{\sin kR}{(kR)^3} - \frac{\cos kR}{(kR)^2} \right]^2 d \ln k \quad (1.42)$$

In the literature a smoothing scale of  $R = 8 \text{ Mpc } h^{-1}$  is often used, giving rise to the quantity  $\sigma_8$ .

A Gaussian field has the property that the Fourier phases are uncorrelated. In the linear regime each mode evolves independently and thus the Gaussianity is not broken until modes start to go non-linear. Since a Gaussian field is fully characterised by its variance, which is exactly the power spectrum, it is only when modes go non-linear, that the power spectrum is not a sufficient statistic. The universe started in an initially homogeneous and isotropic configuration and was thus fully quantified by the power spectrum. This is the motivation for the definition of the transfer function which evolves the different Fourier modes of the density field in the linear regime. The transfer function is defined such that

$$P_{a=0}(k) \equiv T(k)^2 P_{\text{primordial}}(k) \quad (1.43)$$

where  $P_{\text{primordial}}$  is the initial conditions for the universe set up by some mechanism (inflation for example). More specifically, the transfer function is conveniently defined such that the time evolution of the density field is encapsulated in the growth

function  $D(a)$

$$\delta = \delta_{\text{primordial}} T(k) \frac{D(a)}{a}. \quad (1.44)$$

The growth function can be calculated from the background cosmological parameters alone and is normalised to  $D(a = 1) = 1$ , whereas the transfer function takes all other parameters into consideration such as couplings to other components or horizon entry for a given mode. To fully calculate the transfer function in a realistic cosmology with interactions between the different sectors is a very complicated numerical task even when the universe is still linear. Fortunately, decades of research have culminated in sophisticated numerical tools that can calculate the transfer function given a background cosmology. See for example the CAMB suite (Lewis et al., 2000) or CLASS (Blas et al., 2011; Lesgourgues, 2011).

## 1.4 Inflation and initial conditions for the universe

The standard model of the universe (The Hot Big Bang Model) states that the initial configuration of the universe was in a singular point in spacetime known as the Big Bang.<sup>c</sup> Since the very early universe had an extraordinarily large energy density compared to any known physics, very little can be inferred about the early universe. However, at a scale of no more than  $10^{16}$  GeV the inflation epoch set in, rapidly expanding the universe by a factor of  $\sim e^{60}$ . The inflation epoch is very hard to circumvent when explaining the apparent flatness of the universe, the origin of the original perturbations or the thermal equilibrium of super-horizon scales. In the simplest models the initial inflation field is a scalar field in a slow-roll towards the global minimum. In these slow-roll models the inflaton field acts as a slowly decaying cosmological constant eventually decaying into the known particles of the Standard Model (or some intermediate particles that will later decay). With the decay of the inflaton field the universe is reheated with the production of highly relativistic particles, and the universe subsequently enters a radiation dominated phase. Specifically the inflation epoch stress-energy tensor is diagonal with energy density

$$\rho = \frac{1}{2} \left( \frac{d\phi}{dt} \right)^2 + V(\phi) \quad (1.45)$$

and pressure

$$P/c^2 = \frac{1}{2} \left( \frac{d\phi}{dt} \right)^2 - V(\phi) \quad (1.46)$$

---

<sup>c</sup>Essentially a white hole.

with  $\phi$  the inflaton field and  $V$  some potential. The Friedmann equation then takes the form

$$H^2 = \frac{c\hbar}{3M_{\text{planck}}^2} \left[ V(\phi) + 1/2 \left( \frac{d\phi}{dt} \right)^2 \right], \quad (1.47)$$

where  $M_{\text{planck}} = (c\hbar/8\pi G)^{1/2}$  is the reduced Planck mass, and the inflaton field obeying the Klein-Gordon equation

$$\frac{d^2\phi}{dt^2} + 3H \frac{d\phi}{dt} = -\frac{dV}{d\phi}. \quad (1.48)$$

The requirement for slow-roll is then for

$$\epsilon \equiv \frac{M_{\text{planck}}^2}{2c\hbar} \left[ \frac{dV/d\phi}{V} \right]^2 \quad \text{and} \quad \eta \equiv \frac{M_{\text{planck}}^2}{c\hbar} \frac{d^2V/d\phi^2}{V} \quad (1.49)$$

to be small, and this is *a priori* the only requirement on  $V$ . The ratio of the power in tensor perturbations to scalar perturbations,  $r$ , is proportional to  $\epsilon$  so observing this ratio gives valuable constraints on inflation. Under the assumption of a nearly de-Sitter spacetime and a small inflaton mass it can be shown that the power spectrum of the metric perturbations is scale invariant. This is encoded in the spectral index  $n_s$  defined such that  $P_\phi \propto k^{n_s+1}$ , with  $n_s = -1$  defining a scale invariant spectrum. Such scale invariance will in turn force the matter perturbations to have an initial spectrum proportional to  $k$ . However, since the inflaton field is not exactly a cosmological constant, and since it does have mass, a small deviation from this proportionality is expected. Current estimates of the primordial matter power spectrum is that  $P_\delta \propto k^{-0.96}$  (Ade et al., 2013). See also Lesgourgues (2006); Liddle and Lyth (2000) for some good overviews of inflation.

The exact physics of the inflation epoch are not very well constrained. In fact, inflation very effectively hides the physics in the pre-inflationary epoch from observation. Recently, however, the second generation of the South Pole based "Background Imaging of Cosmic Extragalactic Polarization" (BICEP2) telescope claimed to have observed primordial gravitational  $B$ -modes from the inflationary period (BICEP2 Collaboration et al., 2014). However, Mortonson and Seljak (2014) performed a joint analysis on Planck+BICEP2 data and found no preference for gravitational waves. Likewise, Flauger et al. (2014) found that while the BICEP2 data is consistent with a cosmology with a tensor-to-scalar ratio of  $r \sim 0.2$  and negligible dust foreground, it is also consistent with a cosmology with  $r \sim 0$  and a significant dust polarization signal. It is evidently premature to draw any firm conclusions from the BICEP2 data, but with a careful mapping of the dust foreground it may be possible to see the first real fingerprint of the inflationary epoch in a few years. To this extent, high frequency Planck data and Keck Array observations will likely play a crucial role.

## 1.5 Cosmological models

Any component relevant for the evolution of the universe as a whole can have its energy contribution parametrised by

$$\epsilon = \epsilon_0 a^{-3(1+w)}, \quad (1.50)$$

with  $w$  (assumed to be independent of time) the equation of state parameter relating a components pressure to its energy density  $P = w\epsilon$ . For dark matter, with the dominant contribution being the rest mass  $w = 0$ , whereas for a cosmological constant  $w = -1$ . In general a component needs to have  $w \geq -1$  in order to obey the weak energy condition of general relativity  $\epsilon + P/c^2 > 0$ . Dark energy models which have  $w < -1$  are named phantom dark energy, and are notoriously difficult to treat in a dynamical context.

The standard model of cosmology is the  $\Lambda$ CDM model<sup>d</sup> which is the cosmology of a flat universe primarily inhabited by a cosmological constant,  $\Lambda$ , and a cold dark matter (CDM) component. The cosmological constant is responsible for the late time acceleration of the expansion, whereas CDM acts as seeds for the growth of structure in the early universe. In the current universe the  $\Lambda$  component contributes roughly 70 % and the total matter component is roughly 30 % (with baryonic matter making up  $\approx 5$  %). The neutrino and radiation contributions are very small, but as we will see later the neutrino component affects the matter component on small scales at the percent level. In fact, neutrinos affect the power spectrum by an order of magnitude more than a simple  $\Omega_\nu/\Omega_m$  guess warrants. Some of the best estimates on these cosmological parameters come from CMB data. See, e.g., Ade et al. (2013) for the precise values from the Planck mission.

The observational results on the dark energy equation of state parameter  $w$  are not conclusive, and while  $w = -1$  gives the simplest model consistent with observations, one could also imagine having a value of  $w$  slightly above or below this limit. Observations are also consistent with a time varying equation of state parameter

$$w = w_0 + (1 - a)w_a. \quad (1.51)$$

We will return to this observation in chapter 4.

### 1.5.1 Tensions in $\Lambda$ CDM

The rotation velocity in galaxies gives an excellent way to probe the potential and thus the matter distribution. In the innermost regions of galaxies the baryonic component dominates, and thus a multicomponent mass-model is needed. Observations of the central part of disk galaxies point towards a linearly increasing rotation

<sup>d</sup>More accurately one could say that  $\Lambda$ CDM is the standard parametrization of the standard Hot Big Bang Model.

velocity with the distance from the center (Kormendy and Freeman, 2004). This can be interpreted as indicating a central dark matter core that is approximately constant  $\rho \propto r^0$  in the central part. Simulations, however, indicate a NFW profile with the central dark matter cores having a more cuspy distribution, scaling as  $\rho \propto r^\alpha$  with  $\alpha \simeq -1$ . This gives rise to a velocity profile of the central region that increases with the square root of the distance. This is especially problematic for dwarf spheroidals as they are believed to be completely dominated by dark matter, yet their stellar profiles show no sign of a cuspy centre. It is a general property of  $N$ -body simulations to have  $\alpha \lesssim -1$  at 1 kpc, and at present, this core-cusp problem is still an open question in cosmology. de Blok (2010) gives a decent review of the core-cusp problem and possible solutions dealing with interactions between the baryonic component and the dark matter component.

Another problem with the  $\Lambda$ CDM model is the dwarf problem (also called the missing satellites problem): Observations find far fewer dwarf galaxies in the Local Group than what is to be expected from simulations for a similar galaxy group (Klypin et al., 1999). It has been suggested that a possible solution to the dwarf problem may lie in a more accurate description of baryonic physics such as supernova feedback and increased tidal stripping due to the baryonic disk (Brooks et al., 2013). It is also a possibility that the missing satellites exist, but are too faint to be observed with existing telescopes. Katz and Ricotti (2012) and Bovill and Ricotti (2011) proposes a mechanism where the bulk of the star formation in the first galaxies happened in loosely bound globular clusters that were subsequently tidally stripped from the later-to-be dwarf galaxies due to the weak potential. This could potentially explain the apparent faintness of dwarf galaxies without invoking any baryonic feedback mechanisms.

Related to the dwarf problem is the apparent lack of faint galaxies in the voids defined by the bright galaxy population, the so called "Void Phenomenon", as discussed by Peebles (2001). Simulations show that the voids between the concentrations of large galaxy clusters is inhabited by dark matter haloes capable of hosting gas and stars in dwarf galaxies, contrary to what is observed. It is puzzling why the, albeit few, dwarf galaxies we observe in our local neighbourhood do not inhabit the voids where their formation seems more favourable. In other words the voids contain very few, if any, faint galaxies, and the few galaxies actually found are similar to the overall galaxy population. This is hard to understand within the standard  $\Lambda$ CDM scheme as smaller haloes formed earlier, and should not be sensitive to the later forming large scale filamentary structure. In the  $\Lambda$ CDM model the voids are never completely empty of matter and contain a significant amount of low mass haloes - which would mean dwarf galaxies should be the preferred population. Even in the case of baryonic feedback it is hard to justify why this would act more potently in the voids. However, see Tinker and Conroy (2009) for a possible explanation within the  $\Lambda$ CDM model. Here it was argued that galaxies are sensitive only to the mass of the halo it inhabits, and

not of the larger scale environment. Thus galaxies in voids should be a fair sample of the overall population. Furthermore, since the mass-to-light ratio increases for faint galaxies this in turn allows haloes of roughly comparable mass to host a wider range of galaxy luminosities. This means that we should not observe an increase of faint galaxies in voids, as the small mass haloes can also host brighter galaxies due to the decreased mass-to-light ratio.

Observations of the circular velocities in the Milky Way satellites show that most of the satellites have a  $V_{max} \lesssim 30$  km/s.<sup>e</sup> Since the maximal circular velocity of a galaxy tightly constrain the total mass of the halo within the luminous region, and since this mass correlates strongly with the total mass, this gives good estimates of the total masses of the satellites. By running a series of  $N$ -body simulations - the Aquarius simulations - Boylan-Kolchin et al. (2011) and Boylan-Kolchin et al. (2012) showed that the satellites for a typical Milky Way galaxy are too dense to host any of the Milky Way bright satellites. The question is then, if the simulations are to be believed, why are the most massive Milky Way subhaloes not also populated by the most luminous satellites? Are these massive subhaloes too big to fail in their galaxy formation, or is galaxy formation at these scales essentially stochastic? For this reason this issue has been named the too-big-to-fail problem, with regard to the failed star formation in the massive Milky Way subhaloes. Polisensky and Ricotti (2014) argued that the too-big-to-fail problem is an artifact of using pre-Planck values for  $\sigma_8$  and  $n_s$  in the Aquarius simulations.

The core-cups problem, the missing satellite problem and the too-big-to-fail problem are the most glaring problems with the standard model of cosmology. They all concern the small scale features of  $\Lambda$ CDM, in what is known as the "small scale crisis" of  $\Lambda$ CDM. See Weinberg et al. (2013) for a brief discussion of the small scale tension in the  $\Lambda$ CDM model and possible solutions. The smallest scales in  $N$ -body simulations are also the most sensitive to baryonic physics, and so the issues can possibly be solved with better understanding of baryonic physics such as feedback mechanisms. As such there is no reason to suspect that  $\Lambda$ CDM is inherently wrong. However, many alternative models have been proposed. For example, warm dark matter (WDM) has been proposed as an alternative, and Lovell et al. (2012) showed that this could alleviate the tension with the too-big-to-fail problem. Since  $\Lambda$ CDM is the simplest model consistent with most observations, any alternative model has to be tightly constrained by current observations.

### 1.5.2 $\Lambda$ WDM

The last decade has seen an increased interest in WDM models, much of it owing to the results of Bode et al. (2001). Here it was shown that the small scale crisis in  $\Lambda$ CDM

<sup>e</sup>Notable exceptions are the Magellanic Clouds and the Sagittarius Dwarf.

can be alleviated if one introduces a warm component instead. The main difference between CDM and WDM is the concept of free-streaming: WDM has a lower particle mass (thus being warmer) than a CDM particle.<sup>f</sup> This means WDM particles had a non-negligible mean free path in the early universe, hence erasing any initial perturbations on smaller scales. Thus a WDM model will have considerably less power on small scales relative to CDM. In CDM clustering happens in a bottom-up hierarchy where the smallest scales form first, whereas a WDM has a bottom-up hierarchy on scales larger than the free-streaming length, but a top-down hierarchy on the smallest scales. This effectively means that small structure formation is delayed with respect to CDM, and therefore, e.g., a lower abundance of dwarf galaxies is expected. Historically, also models with hot dark matter has been of interest, mainly because of the relation with the tiny neutrino mass. These models have a much larger free-streaming length and as such experience mainly a top-down formation hierarchy. In the framework of linear perturbation theory, a CDM component is characterised by decoupling when the particles were non-relativistic in the early and radiation dominated universe. Warm dark matter also decoupled in the radiation dominated phase, however, they decouple while still being relativistic. Hot dark matter decouples relativistically in the matter dominated universe, and since structure grows appreciably in the matter dominated phase, the delayed decoupling translates into a more dramatic suppression than in the case of a warm particle. It in this sense the temperatures of the different dark matter models should be understood: the temperature translates into different decoupling mechanisms.

Properly calculating the free-streaming scale requires integration of the Boltzmann equation. However, to estimate the relevant scale it is usually assumed that the particle species has decoupled and thus travels in free-fall in the expanding background. Matter perturbations only start to grow appreciably after radiation-matter equality at  $t_{rm}$  where they become Jeans unstable. Before this time any inhomogeneities can be smoothed out by particles streaming from overdense to underdense regions. This sets the free-streaming scale

$$\lambda_{FS} = \int_0^{t_{rm}} \frac{v(t)}{a(t)} dt \approx \int_0^{t_{nr}} \frac{c}{a(t)} dt + \int_{t_{nr}}^{t_{rm}} \frac{v(t)}{a(t)} dt \quad (1.52)$$

where the  $t_{nr}$  marks the transition from the particles being relativistic to non-relativistic. Assuming that the universe is radiation dominated when the particles become non-relativistic eq. (1.52) yields

$$\lambda_{FS} = \frac{2ct_{nr}}{a_{nr}} \left( 1 + \frac{1}{2} \ln \left[ \frac{t_{rm}}{t_{nr}} \right] \right) = r_H(t_{nr}) \left( 1 + \frac{1}{2} \ln \left[ \frac{t_{rm}}{t_{nr}} \right] \right) \quad (1.53)$$

<sup>f</sup>For the same production mechanism. Axions can be lighter in absolute terms than particles from other models and still act as cold dark matter.

with  $r_H(t_{nr})$  being the comoving horizon size at the time of the non-relativistic transition. This transition is set by  $kT_X \approx m_X c^2$ , and the free-streaming length can be written as

$$\lambda_{FS} \simeq 0.2 \text{ Mpc} \left( \frac{m_X c^2}{\text{keV}} \right)^{-1} \left( \frac{T_X}{T_\nu} \right) \left( 2 + \ln \left[ \frac{t_{rm}}{t_{nr}} \right] \right) \quad (1.54)$$

where  $T_\nu$  is the neutrino temperature. Any primordial perturbations on scales smaller than the free-streaming length will be wiped out leading to a radically different transfer function. The previous definition of the free-streaming scale follows that of Kolb and Turner (1990), however, it should be noted that several other definitions for the free-streaming scale are employed throughout the literature. Other definitions include the maximum Jeans scale in cosmic history or, as is usually the case for WDM, the scale at which the the power in WDM is half that of CDM

$$T^2(k_{FS}) = \frac{P_{\text{WDM}}(k_{FS})}{P_{\text{CDM}}(k_{FS})} = 1/2, \quad (1.55)$$

where  $k_{FS} = 2\pi/\lambda_{FS}$ . With this definition the free-streaming scale can be shown to be (Bode et al., 2001)

$$\lambda_{\text{FS,Bode}} = 1.2 \left( \frac{\Omega_{\text{WDM}}}{0.3} \right)^{0.15} \left( \frac{h}{0.7} \right)^{1.3} \left( \frac{m_X c^2}{\text{keV}} \right)^{-1.15} h^{-1} \text{Mpc}. \quad (1.56)$$

One could potentially imagine the thermal velocity of the WDM playing an appreciable role later when structures start to form. However, after matter radiation equality at  $z \sim 3200$  the thermal velocity cools in the Hubble flow  $\sigma_\nu \propto a$ . In fact, it is easy to show that the Jeans scale evolves as  $M_J \propto a^{-3/2}$ , and hence the Jeans scale will rapidly decay. Thus the primary suppression of the growth of structure in WDM models is not due to late-time velocities, but mainly an artefact of the initial conditions in the matter power spectrum. The effect of thermal velocities would only be relevant for a centrifugal barrier around the centre of haloes of order  $\sim 10$  pc and smaller. (Bode et al., 2001). We note here that the thermal velocities for ordinary neutrinos cannot be neglected due to their very small mass. In fact the thermal velocities of neutrinos have a significant impact in structure formation due to an on-going free-streaming effect. We will return to this later in section 6.

Several models have been proposed to account for a WDM particle, with the main difference being in whether the dark matter particle is produced in thermal equilibrium or not. The Neutrino Minimal Standard Model ( $\nu$ MSSM) is an extension of the Standard Model to also include three right handed sterile neutrinos. These sterile neutrinos can act as both CDM or WDM. See Boyarsky et al. (2009b) for an overview.  $\nu$ MSSM is also of considerable interest from a particle physics perspective as it can possibly explain many of the problems with the Standard Model such as the rate of neutrino oscillations and the baryon asymmetry. Since the sterile neutrinos are

mainly produced by neutrino oscillations they are not in thermal equilibrium if the production rate is lower than the expansion rate.<sup>8</sup> Another potential candidate for a WDM particle is the thermally produced supersymmetric gravitino. If the scale of supersymmetry breaking is sufficiently small the lightest supersymmetric particle is the gravitino which decouples while still being relativistic. If the scale of supersymmetry breaking is large, however, the lightest supersymmetric particle is the neutralino which decouples while being unrelativistic and hence is a candidate for CDM.

Due to the different production mechanisms of the WDM particles, the particle mass alone does not uniquely determine the free-streaming scale. This is obvious since axions, for example, can have a tiny mass and still act as CDM. Rather, as seen from eq. (1.54), for a fixed dark matter density  $\Omega_X$  it is the characteristic scale  $kT_X/m_X c^2$  at which the particles become non-relativistic that sets the free-streaming scale. In the case of a thermal model and a thermally suppressed sterile neutrino model this fact can be utilised to give a one-to-one correspondence between the particle masses such that the transfer functions are identical

$$m_{\text{sterile}} = 4.43\text{keV} \left( \frac{m_{\text{thermal}} c^2}{\text{keV}} \right)^{4/3} \left( \frac{0.1225}{\Omega_X h^2} \right)^{1/3}. \quad (1.57)$$

See also Viel et al. (2005). Since a sterile or thermal model have identical transfer functions with these masses, for the purpose of structure formation, these cosmologies are identical.

## 1.6 Gravitational lensing in cosmology

This section will cover the basics of weak gravitational lensing in cosmology. We will closely follow the excellent reviews Bartelmann and Schneider (2001) and Schneider (2005). We will in this section allow a non-vanishing curvature of the background.

It is a well known prediction from General Relativity that the curvature of spacetime translates into the bending of light in the presence of gravitational fields. Even Newtonian mechanics predicts this bending if one applies the equivalence principle. The most pronounced effects of this phenomenon are seen in the strong lensing of background galaxies by foreground galaxy clusters. However, weaker lensing effects can also be measured by statistical methods on a large sample of galaxies by the shearing of the observed ellipticities.

First we will examine the general case of a photon bundle propagating in an arbitrary spacetime. We consider an infinitesimally thin ray bundle around a fiducial ray. Let  $\gamma_0^\mu(\lambda)$  be the null geodesic of a fiducial ray arriving at the observer. Here

<sup>8</sup>This is true in the standard Dodelson-Widrow mechanism (Dodelson and Widrow, 1994). Other possible mechanisms include Higgs boson decay or inflation coupling. See, e.g., Kusenko (2009) for a review.

the ray is parametrised by a suitable choice of the affine parameter  $\lambda$ , namely  $\lambda = 0$  at the observer,  $\lambda$  increases with decreasing time and is locally the proper distance. We define  $\gamma^\mu(\lambda, \vec{\theta})$  as the null geodesic of a ray belonging to the bundle propagating at an angle  $\vec{\theta}$  relatively to the central, fiducial ray at the observer. We then define the 2-dimensional vector  $\vec{\xi}(\lambda, \vec{\theta})$  as the transverse separation of a segment of the ray bundle from the fiducial center ray at affine parameter  $\lambda$  (we will ignore any longitudinal deviation). Ignoring the detailed calculation, it follows from linearisation of the equations of geodesic deviation that

$$\vec{\xi}(\lambda, \vec{\theta}) = \mathcal{D}(\lambda)\vec{\theta}, \quad (1.58)$$

where the matrix  $\mathcal{D}(\lambda)$  satisfies the Jacobi equation

$$\frac{d^2 \mathcal{D}(\lambda)}{d\lambda^2} = \mathcal{T}(\lambda)\mathcal{D}(\lambda), \quad (1.59)$$

where  $\mathcal{D}(0) = 0$ ,  $d\mathcal{D}/d\lambda(0) = \mathcal{I}$  is the identity and  $\mathcal{T}$  is the optical tidal matrix. Inserting this into the above yields the equation for  $\vec{\xi}$ , the so-called transport equation

$$\frac{d^2 \vec{\xi}(\lambda, \vec{\theta})}{d\lambda^2} = \mathcal{T}(\lambda)\vec{\xi}(\lambda, \vec{\theta}). \quad (1.60)$$

This is the the fundamental equation for the transverse separation of the rays in the bundle. The optical tidal matrix is in general a very complicated function of the Ricci tensor and Weyl tensor and can be written in the form

$$\mathcal{T}(\lambda) = \begin{bmatrix} \mathcal{R}(\lambda) + \text{Re } \mathcal{F}(\lambda) & \text{Im } \mathcal{F}(\lambda) \\ \text{Im } \mathcal{F}(\lambda) & \mathcal{R}(\lambda) - \text{Re } \mathcal{F}(\lambda). \end{bmatrix} \quad (1.61)$$

$\mathcal{R}$  (not to be mistaken with the Ricci scalar) is called the source of convergence and leads to an isotropic focusing of circular ray bundles. It is obtained from the Ricci tensor and the photon wave vector.  $\mathcal{F}$  is called the source of shear and causes an anisotropic focusing and changing the shape of the photon bundle, and it is given in terms of the Weyl tensor.

### 1.6.1 Weak lensing

In this section we will apply the weak lensing approximation to eq. (1.60). In the general FLRW spacetime the shear must vanish due to isotropy. In that case it follows directly from eq. (1.61) that the tidal matrix  $\mathcal{T}$  is proportional to the identity matrix  $\mathcal{I}$ , and it can relatively easily be shown that  $\mathcal{D}$  is the angular-diameter distance  $af_K(\chi)$ . In fact one can show

$$\mathcal{T} = -\frac{3}{2} \left( \frac{H_0}{c} \right)^2 \Omega_m a^{-5} \mathcal{I}, \quad (1.62)$$

if the universe is assumed to be filled with a pressureless fluid such as CDM for example. Redefining  $\xi$  to the comoving separation  $\vec{x} = \vec{\xi}/a$ , and utilising  $d\lambda = a^2 d\chi$  from the definition of the chosen affine parameter, we can then put eq. (1.60) into the simple form

$$\frac{d^2 \vec{x}}{d\chi^2} + K \vec{x} = 0. \quad (1.63)$$

Here the second term is a curvature term describing the overall curvature of the background, and  $K$  can be found from eq. (1.12). This is the geodesic equation for a single null geodesic in a FLRW spacetime.

In the case of matter inhomogeneities we have the metric in eq. (1.16). In the weak field limit we assume isolated<sup>h</sup> matter inhomogeneities on scales much smaller than the Hubble length  $1/H_0$ , a weak gravitational field  $\psi = \phi = \Phi/c^2 \ll 1$ , and velocities much smaller than the speed of light  $v \ll c$ . Around each perturbation there then exists a neighbourhood which is sufficiently large to contain the perturbation entirely, and in which the Newtonian potential  $\Phi$  is well-defined and obeys the Poisson equation. We then obtain the metric

$$ds^2 = a^2(\eta) \left[ - \left( 1 + 2 \frac{\Phi}{c^2} \right) d\eta^2 + \left( 1 - 2 \frac{\Phi}{c^2} \right) (d\chi^2 + f_K^2(\chi) d\Omega^2) \right]. \quad (1.64)$$

This metric can then be used to determine the tidal matrix

$$\mathcal{T}_{ij} = -\frac{3}{2} \left( \frac{H_0}{c} \right)^2 \Omega_m a^{-5} \delta_{ij} - \frac{1}{a^2 c^2} (2\partial_i \partial_j \Phi + \delta_{ij} \partial_3 \partial_3 \Phi), \quad (1.65)$$

where  $x_3$  is the propagation direction. See Seitz et al. (1994) for a derivation. Here it should be noted that the optical tidal matrix has now acquired shear (off-diagonal) terms.

Now turning our attention to the contribution by the large scale structure, we need to quantify the transversal deviation of a ray observed at an angle  $\vec{\theta}$  with respect to a fiducial ray at  $\vec{\theta} = 0$ . Using eq. (1.65), the transport equation eq. (1.60) then yields

$$\frac{d^2 \vec{x}}{d\chi^2} + K \vec{x} = -\frac{2}{c^2} \left[ \nabla_{\perp} \Phi(\vec{x}(\vec{\theta}, \chi), \chi) - \nabla_{\perp} \Phi^0(\chi) \right], \quad (1.66)$$

where  $\Phi, \Phi^0$  is the Newtonian potential along the observed ray and the fiducial ray respectively. This is the transport equation in the weak field limit from which everything follows. Comparing to eq. (1.63) for the FLRW metric, we see that the only difference is the right hand side which is proportional to the off-diagonal terms of the tidal matrix.

---

<sup>h</sup>It might seem a bit odd to assume isolated matter inhomogeneities since we wish to quantify the weak lensing by large scale structures. However, the large scale structure consists of a superposition of many galaxies and corresponding halos which are localised.

The differential equation eq. (1.66) can conveniently be solved by Green's method. We then obtain the fundamental equation for  $\vec{x}$

$$x(\theta, \chi) = f_K(\chi)\theta - \frac{2}{c^2} \int_0^\chi d\chi' f_K(\chi - \chi') \left[ \nabla_\perp \Phi(\vec{x}(\vec{\theta}, \chi'), \chi') - \nabla_\perp \Phi^0(\chi') \right]. \quad (1.67)$$

With  $\vec{\beta} = \vec{x}/f_K(\chi)$  as the position of a source at comoving distance  $\chi$  in the absence of matter inhomogeneities, we define the distortion matrix  $A$  as the Jacobian of the lens mapping  $\vec{\theta} \mapsto \vec{\beta}(\theta, \chi)$

$$A \equiv \frac{\partial \vec{\beta}(\vec{\theta}, \chi)}{\partial \vec{\theta}} = \frac{1}{f_K(\chi)} \frac{\partial \vec{x}(\vec{\theta}, \chi)}{\partial \vec{\theta}} \quad (1.68)$$

$$A_{ij} = \delta_{ij} - \frac{2}{c^2} \int_0^\chi d\chi' \frac{f_K(\chi - \chi') f_K(\chi')}{f_K(\chi)} \frac{\partial^2 \Phi(\vec{x}(\theta, \chi'), \chi')}{\partial \theta_i \partial \theta_k} A_{kj}(\theta, \chi'). \quad (1.69)$$

The distortion matrix is related to the magnification  $\mu$  by  $\mu = 1/\det(A)$ , and it is the fundamental quantity to be determined in weak lensing surveys. The distortion matrix can be decomposed into a rotation and a symmetric matrix

$$A = R(\phi) \begin{bmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{bmatrix}. \quad (1.70)$$

The rotation matrix corresponds to a generally unobservable rotation of the source plane. The shear  $\gamma = \gamma_1 + i\gamma_2$  has 2 components and is due to the tidal gravitational potential along the light trajectory. It will cause a distortion of a circular image into an elliptical image. The convergence  $\kappa$  is a scalar, and causes an isotropic focusing, in effect changing the magnification only. The convergence is caused by the actual matter distribution along the trajectory. The shear field is often decomposed into a rotation-free part and a divergence-free part, the so-called E- and B-modes, by a linear combination of  $\gamma_1$  and  $\gamma_2$ . The B-mode is expected to be orders of magnitude smaller than the E-mode, in effect making it unmeasurable and limiting the shear to one degree of freedom. It should also be noted that since both the convergence and the shear are second order derivatives of the gravitational potential it is possible to write down an equation between them in Fourier space known as the Kaiser-Squires relation (Kaiser and Squires, 1993).

A weak lensing survey aims to quantify any of the parameters  $\mu, \kappa$  or  $\vec{\gamma}$ . To first order<sup>i</sup> the statistical properties are the same for the convergence and the shear. Since the intrinsic brightness of the faint source galaxies is very uncertain, weak lensing surveys often aim to measure the shear. Here it is a crucial, and natural in the light of the cosmological principle, assumption that the intrinsic source ellipticities are uncorrelated in a large sample of the sky. Any observation of correlation in the observed ellipticities is then accordingly a measure of the cosmic shear. An often used method to measure the shear is the aperture mass statistic which we will now define.

<sup>i</sup>Precisely what we mean by first order will become apparent later in chapter 3.

## 1.6.2 Observing weak lensing

The power spectrum is a fundamental prediction in cosmology, e.g. from galaxy surveys, and it is often used to intercompare  $N$ -body simulations, and to collate with observations. However, redshift measurements are time consuming, and due to the peculiar velocities of the galaxies, redshift measurements are not an exact measurement of distance. For weak lensing, this is even worse as the source galaxies are in general very distant and very faint. As a result precise photometric analysis is required to obtain redshift dependence. Consequently, galaxy surveys often record only the angular positions of galaxies and extract the angular (or two dimensional) power spectrum  $P_{12}(l)$  - the power spectrum of the angular correlation function  $\xi(\phi)$ . We will hence concern ourselves with the angular correlation of the shear or convergence, and not go into the theory of tomography for weak lensing surveys.

As can be seen from eqs. (1.69) and (1.70) the convergence and shear are quantities derivable from the local gravitational potential, and by use of the Poisson equation, of the total matter distribution. Weak lensing consequently constrains the total matter power spectrum. If, e.g., the convergence angular power spectrum is measured it can directly constrain the three dimensional matter power spectrum. Furthermore, since weak lensing is sensitive to the total matter distribution, it is very suited for predictions involving the parameters  $\Omega_{0,0}$ ,  $\Omega_{\Lambda,0}$  and the normalisation of the power spectrum  $\sigma_8$ . In the literature it has also been noted that a tomographic analysis, where galaxy samples have been collected into redshift bins, substantially improves the parameter estimation (Bartelmann and Schneider, 2001).

A convenient and an often used statistical tool to measure the shear given an observation is the aperture mass statistic. If we assume a circular aperture of radius  $\theta_0$  and choose a weight function  $Q(\phi)$  with compact support  $\mathcal{S}(Q) \subseteq [0; \theta_0]$ , we then define the aperture mass by

$$M_{ap}(\theta) = \int d^2\vec{\alpha} Q(|\vec{\alpha}|) \gamma_t(\vec{\alpha}), \quad (1.71)$$

where  $\gamma_t$  is the tangential part of the shear (the E-mode) as measured towards the center of the aperture and  $\alpha$  is an integration angle. The variance of the aperture mass can then, depending on the approximations and filter functions employed, be related to the shear power spectrum and consequently constrain the matter power spectrum.



## *N*-BODY SIMULATIONS

Perturbation theory successfully describes the early universe and large scales of the present universe. For small scales, below a few hundred megaparsecs in the present universe, the perturbation theory breaks down. Therefore more sophisticated methods are needed such as second order Lagrangian perturbation theory, Press-Schechter theory (and extensions) or numerical gravity solvers. A widely used numerical gravity solver is the *N*-body simulation GADGET-2 (Springel, 2005) which is the primary subject of this chapter. For a review of high-resolution, state of the art simulations as of 2012, see Kuhlen et al. (2012).

### 2.1 The TreePM algorithm

GADGET-2 is an *N*-body simulation suited for both hydrodynamical and cosmological simulations. Integrating gas physics into cosmological simulations is done by the smoothed particle hydrodynamics (SPH) method. Fundamentally, SPH assigns gas particles a density depending on an (adaptive) smoothing length, a smoothing kernel and a sum of the masses of the neighbour particles. From this density, a corresponding pressure can be assigned to the particles and the equations of motion can be determined, in the absence of viscosity, by these parameters. GADGET-2 can also account for viscosity by altering the equations of motion when two or more particles are approaching. Much of the progress in *N*-body simulations has been improving the hydrodynamics, and as such, GADGET-2 being from 2005 is slightly outdated and is slower and less accurate than more recent iterations (e.g. GADGET-3). See for example Springel (2009) for a more recent implementation using Voronoi tessellation.

However, for the results presented in this dissertation, state of the art baryonic treatment is not a primary concern, and the gravity solvers of  $N$ -body simulations have not changed much since GADGET-2.

Naively, two distinct classes of  $N$ -body solvers can be identified: Tree algorithms (notably the Barnes-Hut Tree by Barnes and Hut (1986)), and particle mesh (PM) algorithms. The Tree algorithm starts by a recursive division of the particle distribution with the full distribution being called the root node. This root node is divided into a number of daughter nodes. Each daughter node is then further subdivided into more sub-nodes by the same division scheme, and this division continues until each particle resides in its own unique sub-node called a leaf. In practice, GADGET-2 uses the Barnes-Hut subdivision where the divisioning strategy is an oct-tree: The cubical root node is divided into eight cubical sub-nodes, and each subnode is further divided into octants recursively as explained. This in effect creates a hierarchical "tree"-like structure, hence the name. Each node of the tree only needs to know its immediate parent, daughters and its center of mass. Starting from a particle, one then "walks" the tree from the root node. Each node of the root node is considered, and if a certain opening requirement is satisfied (an example is given below), it is opened and each of its sub-nodes are considered. Likewise, if they satisfy the opening criterion, they are opened, and each of their sub-nodes are considered. If at one point a node does not satisfy the opening criterion, the walk along this branch is discarded, and the center of mass is used for the calculation of the gravitational force. A typical opening criterion is if the angle subtended by the node is greater than  $\theta = d/S$  where  $d$  is the width of the node and  $S$  is the distance to the particle in question. This tree is then walked for each particle. In effect this breaks down to a direct sum for the nearest particles, and distant particles are grouped accordingly and only the monopole moment (or higher) is considered. For a node with mass  $M$  and length  $l$  GADGET-2 uses an opening criterion of

$$\frac{GM}{r^2} \left(\frac{l}{r}\right)^2 \leq \alpha |\mathbf{a}| \quad (2.1)$$

where  $r$  is the distance,  $\mathbf{a}$  the total acceleration of the last timestep and  $\alpha$  a tolerance parameter.

Particle mesh, as an alternative to Tree codes, simply assign particle masses to a grid by a suitable interpolation (Nearest grid-point, Cloud-in-Cell etc.). A Fourier transform is then applied, and the potential is calculated on each gridpoint and interpolated back onto the particles. Naturally, this algorithm can be refined by using an adaptive grid, or direct summation of the nearest particles. Strictly speaking, this is no longer a PM algorithm, but rather an adaptive particle mesh (APM) or a particle-particle-particle-mesh (P<sup>3</sup>M) algorithm.

GADGET-2, being a TreePM code, combines the Tree algorithm with the PM algorithm to calculate the gravitational force in a simulation box with periodic boundary

conditions. The gravitational potential is separated into two parts: A long range part and a short range part. Fundamentally the naive gravitational force calculation by a direct sum is a  $\mathcal{O}(N(N-1))$  calculation since for each particle it is necessary to calculate the gravitational potential from every other particle. In general, the potential at a particle  $i$  is given by the sum

$$\Phi(\vec{x}_i) = -G \sum_{i \neq j} \frac{m_j}{|\vec{x}_i - \vec{x}_j|} \quad (2.2)$$

Very distant particles, however, contribute very little to the sum. Changing to Fourier space, we therefore split the potential to a long-range part and short-range part

$$\Phi_{\mathbf{k}} = \Phi_{\mathbf{k}}^l + \Phi_{\mathbf{k}}^s. \quad (2.3)$$

with

$$\Phi_{\mathbf{k}}^s = \Phi_{\mathbf{k}} S_{\mathbf{k}}(r_s) = \Phi_{\mathbf{k}} \exp(-\mathbf{k}^2 r_s^2) \quad (2.4)$$

$$\Phi_{\mathbf{k}}^l = \Phi_{\mathbf{k}}(1 - S_{\mathbf{k}}(r_s)) = \Phi_{\mathbf{k}}(1 - \exp(-\mathbf{k}^2 r_s^2)). \quad (2.5)$$

$S_{\mathbf{k}}$  serves as a smoothing function with a smoothing scale  $r_s$ . GADGET-2 uses the above Gaussian form, but naturally other functions could be used. The short range force is heavily suppressed on scales larger than a few times the scaling length, and can be calculated with the Tree method in real space and be expressed in terms of the complementary error function. The long range force is calculated by first interpolating the particle masses onto a grid by a Cloud-in-Cell interpolation, and subsequently employing a Fourier transform to obtain the potential on the grid. A finite difference of the grid gives the force, and this is then interpolated back onto the particles.

The problem of calculating the potential is then divided into two parts: A direct sum calculated by the Tree method for the nearest particles, and a Fourier transform done on a particle mesh. It has per se gone from a  $\mathcal{O}(N(N-1))$  calculation to two separate  $\mathcal{O}(N \log N)$  calculations. The Tree algorithm is especially useful since it is also utilised in the domain decomposition used in the parallelisation strategy used in GADGET-2 which is the topic of the next section. In fact both the Tree and the PM algorithms are easy to parallelise.

## 2.2 Parallelisation

The parallelisation strategy in GADGET-2 is capable of running a full  $N$ -body simulation on anything from a single processor (with some overhead) up to thousands of processors on the most modern computer clusters. Furthermore, the force calculations are completely independent of the number of threads.<sup>a</sup> The basic principle

<sup>a</sup>Naturally, running the code on different architectures may change the results slightly.

is to utilise a space filling fractal curve that maps the three dimensional space to a one dimensional curve. GADGET-2 uses the 3 dimensional Peano-Hilbert curve that completely fills up the unit cube. One major advantage of the Peano-Hilbert curve is that it preserves locality: Points close on the curve are also close in space. Therefore, an equidistant partition of the curve tends to translate into compact spatial domains. Furthermore, due to this property, there is a useful correspondence between splitting the domain into octants and segmenting a Peano-Hilbert curve.

The parallelisation of the Fourier transform calculation in the PM part does not share the domain decomposition with the tree part unfortunately. Instead, the simulation volume is evenly divided into slabs along one dimension, and the number of slabs is distributed to (ideally) all of the processors. Each process then determines what slab its local particles overlaps with. The Cloud-in-Cell interpolation is performed locally for each process, and the local patch is communicated to the processes to which it belongs in the slab decomposition. There is no guarantee that the slab decomposition overlaps with the local patch of a process, however. Fortunately the tree part of the force calculation takes up roughly an order of magnitude more time than the PM part, so any communication bottlenecks in the slab decomposition are not severe.

## 2.3 Timestepping

GADGET-2 utilises leapfrog integration for the advancing the particle's position and velocity in successive timesteps. This simply means that particles experience two fundamentally distinct timesteps: A kick operation  $K$  where the velocity is changed and the position is kept constant, and a drift operation  $D$  where the velocity is kept constant and the position changes:

$$\begin{aligned} K(\Delta t): \quad & \mathbf{x} \rightarrow \mathbf{x}, & \mathbf{p} & \rightarrow \mathbf{p} + \mathbf{F}\Delta t \\ D(\Delta t): \quad & \mathbf{x} \rightarrow \mathbf{x} + \frac{\mathbf{p}}{m}\Delta t, & \mathbf{p} & \rightarrow \mathbf{p} \end{aligned}$$

For a uniform timestepping this integration scheme has the advantage of being symplectic.

Due to the large dynamic range of *N*-body simulations it is not feasible to keep a uniform timestepping where all particles experience the same timestep. In highly clustered regions the dynamical timescale is much shorter than in underdense regions. For this reason GADGET-2 utilises a hybrid timestepping scheme in the TreePM mode. On the coarsest level, particles share a uniform top level timestep determined by the cosmology. This timestep is then subdivided into several smaller steps that can be refined on the individual particle level. Due to the slowly changing nature of the large scale force, the long range kick operation is calculated only once at the

coarsest level. The short range timestepping is done on an individual level for each particle. Particles in highly clustered regions will have the short range force contributions evaluated several times for each long range timestep. Therefore, schematically the timestepping in GADGET-2 can be represented as

$$K_{\text{r}}\left(\frac{\Delta t}{2}\right)\left[K_{\text{sr}}\left(\frac{\Delta t}{2m}\right)D\left(\frac{\Delta t}{m}\right)K_{\text{sr}}\left(\frac{\Delta t}{2m}\right)\right]^m K_{\text{r}}\left(\frac{\Delta t}{2}\right) \quad (2.6)$$

The short range steps are always calculated on a power of 2 subdivided timescale in order to have the timesteps be synchronised.

## 2.4 The equations of motion in GADGET-2

This section is thought as a starting point for users interested in understanding the internal GADGET-2 equations of motion. Hopefully it will clear a few things up missing in the (otherwise excellent) documentation. Throughout we assume a periodic box with comoving coordinates  $\boldsymbol{x}$ .

A particle moving in a comoving frame has a proper velocity of  $a\dot{\boldsymbol{x}} + \boldsymbol{x}\dot{a}$ . Hence the Lagrangian is

$$\mathcal{L} = \frac{1}{2}(a\dot{\boldsymbol{x}} + \boldsymbol{x}\dot{a})^2 - \Phi, \quad (2.7)$$

where  $\Phi$  is the Newtonian potential. The generating function  $\psi = 1/2a\dot{a}\boldsymbol{x}^2$  corresponds to a canonical transformation into

$$\mathcal{L} \rightarrow \mathcal{L} - d\psi/dt = \frac{1}{2}a^2\dot{\boldsymbol{x}}^2 - \phi/a, \quad (2.8)$$

with

$$\phi = a\Phi + 1/2a^2\ddot{a}\boldsymbol{x}^2. \quad (2.9)$$

Introducing the canonical momentum  $\boldsymbol{p} = a^2\dot{\boldsymbol{x}}$  the Hamiltonian can then be written

$$\mathcal{H} = \frac{\boldsymbol{p}^2}{2a^2} + \frac{\phi}{a}. \quad (2.10)$$

This gives the equations of motion

$$\frac{d\boldsymbol{x}}{dt} = \frac{\boldsymbol{p}}{a^2} \quad (2.11)$$

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\nabla\phi}{a}. \quad (2.12)$$

See also Peebles (1980) and Quinn et al. (1997). Since the internal velocity in GADGET-2 is exactly the canonical momentum this gives the drift and kick operators

$$D(\Delta t) : \mathbf{x} \rightarrow \mathbf{x} + \mathbf{p} \int_t^{t+\Delta t} \frac{dt}{a^2} = \mathbf{x} + \mathbf{p} \int_a^{a+\Delta a} \frac{da}{a^3 H} \quad (2.13)$$

$$K(\Delta t) : \mathbf{p} \rightarrow \mathbf{p} - \nabla\phi \int_t^{t+\Delta t} \frac{dt}{a} = \mathbf{p} - \nabla\phi \int_a^{a+\Delta a} \frac{da}{a^2 H}. \quad (2.14)$$

It should be noted that the initial condition and snapshot files do not use  $\mathbf{p}$  as the velocity variable, instead they use  $\mathbf{p}/a^{3/2}$ .

It may not *a priori* be obvious what the Poisson equation for  $\phi$  is. To see this relation, note that when using physical quantities and a comoving gradient Poisson's equation takes the form

$$\nabla_{\mathbf{x}}^2 \Phi = 4\pi G a^2 \rho(\mathbf{x}, t). \quad (2.15)$$

Then, from eq. (2.9)

$$\nabla_{\mathbf{x}}^2 \phi = 4\pi G [a^3 \rho(\mathbf{x}, t) - a^3 \bar{\rho}(t)] = 4\pi G a^3 \delta\rho(\mathbf{x}, t), \quad (2.16)$$

where the acceleration equation for the background has been used. From this it is seen that  $\phi$  obeys a Poisson equation with a relative comoving density  $a^3 \delta\rho$  serving as the source term.<sup>b</sup> GADGET-2 uses eq. (2.16) when calculating the potential on the grid in the PM part.

The potential is calculated in Fourier space after the particle masses have been interpolated onto the grid. Let  $B$  be the box size,  $G_{PM}$  be the grid size and  $m(\mathbf{k})$  the mass in a gridpoint. In Fourier space the derivative operator  $\partial_i$  changes to  $-ik_i$ , so the Laplace operator  $\nabla^2$  changes to  $-k^2$ . Thus the potential can be calculated by dividing each point with a factor of

$$-k^2 = -\left(\frac{2\pi}{B}\right)^2 k_i^2, \quad (2.17)$$

where  $k_i$  is a 3-tuple of integers running from  $-G_{PM}/2$  to  $G_{PM}/2$ . The potential in a gridpoint is then

$$\phi = -\frac{G}{\pi B} \frac{G_{PM}^3 m(\mathbf{k})}{k_i^2}. \quad (2.18)$$

It should be noted here that the mass grid is automatically multiplied by  $G_{PM}^3$  when performing the Fourier transform with the normalization convention used in the FFTW libraries.

<sup>b</sup>Comoving since for matter  $\rho \propto a^{-3}$  so  $a^3 \delta\rho$  is constant.

## 2.5 Halo finding in $N$ -body simulations

The *raison d'être* for  $N$ -body simulations is to deduce physical observables and compare them with observations. Fundamentally the  $N$ -body simulation simultaneously solves the Boltzmann equation, Poisson's equation and the equations of baryonic interactions, and one has to post-process the phase space sample to yield observable parameters. Usually the connection between observations and simulations is the halo distribution. Galaxies are generally believed to be fair tracers of the underlying halo distribution in our universe and are easily observed. Therefore it is natural to compare the observed galaxy population with the halo distribution and mock galaxy catalogues from simulations.

The definition of a halo is fairly straightforward if vague - a halo is simply a gravitationally bound object. Defining a metric on haloes, however, is not so straightforward. The Press-Schechter approach of defining the virial radius at a relative overdensity of  $\delta \approx 200$  is ambiguous and completely ignores embedded substructure (such as the individual galaxy haloes in galaxy clusters). Despite of these shortcomings, the virial radius  $r_{vir}$  (and corresponding mass  $M_{vir}$ ) is often employed as a parameter. Another parameter often used is the peak in the rotation curve  $v_{max}$  and the corresponding radius  $r_{max}$ . This parameter is easier to compare to observations as there is no ambiguity in defining the halo edge. However, since the peak velocity is mostly set by the central parts of a halo, this parameter is not sensitive to mass-loss by tidal stripping. Furthermore, dynamical friction effects are usually tightly correlated with the total mass and not just the central region mass. The ambiguity in defining a "one-size-fits-all" halo metric accentuates the need to use a halo parameter relevant to the problem and scale in question.

Broadly speaking halo finding can be divided into two distinct categories:

- Density peak locators
- Particle grouping algorithms

Finding peaks in the density is, e.g., done by constructing an adaptive grid and identifying isolated regions as potential halo centers as explained below. Around these halo centers spherical overdensities are grown out to some density value and particles are assigned to these spheres. Particle grouping algorithms, on the other hand, identify potential halos by linking together particles "close" to each other with some suitable metric (either in real space or phase space). See Knebe et al. (2011, 2013) for comparisons between different halo finders.

Historically halo substructure has been a matter of some controversy over the years. The so-called "overmerging problem", where small halos orbiting more massive halos seemed to be destroyed, was only solved when Klypin et al. (1999) noted that it was due to the lack of sufficient spatial resolution within the host halo, and

not an inherent property of massive halos. The lesson to learn from this, is that halo finding calls for careful convergence checks and sufficient force resolution (at least kiloparsec scale for galaxy-like halos). As we shall see in section 2.5.3, this is especially important for WDM cosmologies. Fortunately, independent halo finding techniques seems to have converged fairly well when identifying substructure (Knebe et al., 2011).

### 2.5.1 Amiga's Halo Finder

The earliest example of a density peak locator is the Spherical Overdensity (SO) algorithm by Press and Schechter (1974) which grows spherical shells from identified peaks in the density field until a density threshold is reached. This first approach completely neglects any substructure. More recent iterations of this algorithm include Amiga's Halo Finder (AHF) (Knollmann and Knebe, 2009). AHF is an adaptive mesh code that recursively refines a coarse domain grid into increasingly refined grids based on the particle density. This creates a tree-like grid hierarchy that samples the density field. By travelling the tree from the finest grids to the domain grid potential halos are identified as isolated regions of a given refinement level. By construction the isolated region of a given refinement level is a subset of the coarser grid from which it was constructed. If two areas, which are isolated on one level, belongs to the same coarser grid the two branches are joined.

After AHF has constructed the tree, it is traversed from top and down. Once a branch splits into two sub-branches the branch containing the most particles is marked as the main branch, the others as substructures. Each leaf in a branch is then assumed to be a halo. When a sub-halo branch is joined with the main-branch all particles in a sphere of radius half the distance between the sub-halo and the main-halo is assigned to the sub-halo. Particles within the virial radius ( $\delta(r) > \Delta_{vir} \sim 200$ ) is assigned to main haloes (excluding those in sub-haloes). For every particle in a halo the total energy is calculated, and if it is negative it is toggled as belonging to the halo. If it is positive it is checked whether it belongs to a halo higher in the hierarchy (if any). After this unbinding the halo characteristics such as mass, velocity profile etc. is calculated.

### 2.5.2 Rockstar

The earliest example of a particle grouping algorithm is the naive Friends-of-Friends (FoF) algorithm by Davis et al.. Here groups of particles are linked if they are closer than the (constant) linking length, and in this way particles can be categorized as belonging to halos. A typical linking length is  $l_{link} = b l_{mean}$  where  $b$  is of order unity and  $l_{mean}$  is the mean particle separation. More recent implementations of the FoF algorithm groups particles "close" in 6-dimensional phase space together, and also

considers the temporal evolution from snapshot to snapshot. The increased amount of data compared to a 3-dimensional analysis yields more accurate halo characteristics. This is because a 3-dimensional position-based analysis cannot track different haloes during mergers as the assignment of a particle to either of the infalling haloes is essentially random in the overlapping region. The velocity distribution, however, is generally very different as the momenta of the particles in the two haloes have opposite sign along the merger axis. Defining a metric in phase space is, on the other hand, not so obvious and introduces a model dependence in the halo finding. The same considerations are true for substructure, and as such taking the full 6-dimensional dataset into account gives a better identification, especially if knowledge about the substructure at earlier times is available.

A prime example of a modern FoF algorithm is the Rockstar halo finder (Behroozi et al., 2013). Rockstar starts with a simple grouping of particles together in naive FoF groups based on a linking length of  $l_{link} = 0.28l_{mean}$ . The neighbour search is sped up by only considering the nearest 16 neighbours and joining overlapping groups so that heavily clustered regions will not take up a significant amount of time. The naive FoF grouping is done exclusively to divide the simulation volume into suitable workloads for each process.

After particles have been distributed to each process, Rockstar further subdivides the FoF groups into a tree structure by successively tighter linking constraints. For this purpose the phase space metric

$$d(p_1, p_2) = \sqrt{\left( \frac{|\vec{x}_1 - \vec{x}_2|}{\sigma_x} + \frac{|\vec{v}_1 - \vec{v}_2|}{\sigma_v} \right)^2} \quad (2.19)$$

is employed. Here  $\sigma_x$  and  $\sigma_v$  are the particle position and velocity dispersion within the given FoF group. The linking length is then determined by the requirement that at least 70 % of particles in a group are linked within at least one subgroup. The halo catalogue can then be calculated and masses and other halo properties assigned to haloes. If available, Rockstar uses temporal information to assign satellite membership before calculating halo masses.

### 2.5.3 Halo finding in warm dark matter cosmologies

Warm dark matter simulations present some unique challenges for halo finding when compared to CDM simulations. Cold dark matter, with its bottom-up formation history, has structure on all scales below a certain threshold. The free-streaming length in WDM cosmologies postpones the formation of the smallest structures and relatively fewer halos would be expected to exist below the free-streaming mass scale. Naively extracting the halo catalogues from WDM simulations, however, proves otherwise. Here it is found that for the smallest halos there is actually an increase in the

halo abundance, and that these halos are confined to the filaments of the cosmic web. This is true even for state of the art simulations (Angulo et al., 2013; Lovell et al., 2014; Schneider et al., 2012, 2013b). Much like the overmerging problem of the early CDM simulations, this was shown by Wang and White (2007) to be a numerical artefact rather than a physical property. Here it was found that self-bound non-linear structures with masses smaller than the free-streaming mass appears to originate through spurious fragmentation of the filaments, and that the fragmentation happens in regularly spaced clumps that reflect the mean interparticle distance in the initial conditions. Furthermore, the fragments appear even in the absence of clearly defined axes in the initial conditions such as is the case in glass-like initial conditions. The fragments appears at a scale  $m_p^{1/3} k_{peak}^{-2}$  where  $m_p$  is the particle mass and  $k_{peak}$  is the peak of the power spectrum. The interpretation of this result is that, if these artificial structures are not properly removed, the mass resolution of WDM  $N$ -body simulations grows only as  $N^{1/3}$  and not directly proportional to  $N$ .<sup>c</sup> The fragmentation is caused by having a much smaller force resolution (softening length) than mean interparticle separation. As noted in Angulo et al. (2013), this is because an excessive force resolution creates local potential minima around the simulation particles which accrete neighbouring particles over time. The fact that the effective resolution of WDM cosmologies grows only as  $N^{1/3}$  is a severe limitation unless these unphysical structures, or artificial haloes, are properly identified and removed. Fortunately, several authors have proposed algorithms to remove the artificial haloes.

Angulo et al. (2013) utilises a T4PM gravity algorithm, based on a phase space Delauney triangulation. While actually increasing the force resolution, this approach smooths the potential minima around particles and more accurately tracks the true continuous density field. Furthermore, potential haloes are divided into distinct classes by visual inspection. Proto-haloes that are not yet fully collapsed and virialized are considered an independent halo population. Failures of the halo finding algorithm to properly identify haloes were also addressed, and by properly removing these haloes they found that the (differential) halo mass function was heavily suppressed under a cut-off scale as expected.

Schneider et al. (2013b) noted that the artificial haloes in the low-mass end followed a power law distribution, and that the power law index did not depend strongly on the WDM particle mass. By fitting the low-mass end with a power law and subtracting this from the total mass function the artificial haloes were essentially removed. This approach yielded good results for the low- $z$  halo mass function, at higher redshift the power law behaviour was less pronounced (probably due to the relative number of artificial haloes being lower).

---

<sup>c</sup>However, this is also the case to some degree for cosmologies with no free-streaming (Melott, 2007).

Lovell et al. (2014) removes artificial haloes by tracing the particles back to their initial positions prior to collapse. Here it is found that the initial particle distribution in artificial haloes is highly elongated whereas genuine proto-haloes are more spheroidal. By further requiring physical haloes to be present in simulations with varying resolution it is possible to infer a mass cut where contamination by spurious haloes starts to dominate. From this an accurate halo mass function can be constructed, however this comes at an appreciable computational cost as several simulations of various resolution has to be run.

As the above discussion details, halo finding in cosmologies with free-streaming, such as WDM cosmologies, can be a formidable task. However, better understanding of the underlying numerical problems has catalysed the development of methods to properly account for the formation of spurious structures. As such, the dependence on the number of particles in these simulations is not as catastrophic as the  $N^{1/3}$  dependence in the worst case scenario.



## **Part II**

# **Research in progress**



## AN ALGORITHM FOR $N$ -BODY RAY-TRACING

Traditional algorithms utilise the discrete lens-plane algorithm (also called the multiple-lens-plane approximation) where the mass distribution from the simulation snapshots is projected onto a discrete set of planes, and the lensing is only calculated on these planes, see e.g. Hilbert et al. (2008). White and Hu (2000) developed a method where the path of the individual rays in a bundle were numerically integrated along the line of sight on the fly. This method was used by Li et al. (2010) to calculate the lensing signal by interpolating the potential from a grid onto the individual light rays and analytically calculating the lensing signal. In this section we will present our method for ray-tracing in  $N$ -body simulations which borrows some elements from this approach.

Other recent ray-tracing simulations are also worthy of mention. Teyssier et al. (2008) performed a full sky survey with 70 billion  $N$ -body particles in a 2 Gpc/h box in the Horizon simulation via the grid-based RAMSES  $N$ -body code. Fosalba et al. (2007) performed a full sky survey with the MICE simulation (using GADGET-2) with  $2048^3$  particles in a 3 Gpc/h box by storing a HEALPIX tessellated grid in a successive series of spherical shells, in effect creating an onion-like dataset in analogy to real data.

### 3.1 The algorithm

The fundamental quantity to be determined is the components of the distortion matrix eq. (1.70) - the shear  $\gamma$  and the convergence  $\kappa$ . The fundamental equation to be solved is therefore eq. (1.69)

$$A_{ij} = \delta_{ij} - \frac{2}{c^2} \int_0^\chi d\chi' \frac{f_K(\chi - \chi') f_K(\chi')}{f_K(\chi)} \frac{\partial^2 \Phi(\vec{x}(\theta, \chi'), \chi')}{\partial \theta_i \partial \theta_k} A_{kj}(\theta, \chi'). \quad (1.69)$$

Starting from an observer position in the  $N$ -body simulation, this equation has to be integrated along the actual ray paths in the *past-directed* lightcone. However, the  $N$ -body simulation is run *forwards*, i.e., from high redshift to low redshift. With the actual path of the light rays being slightly bend in the presence of matter there is in general no way of ensuring *a priori* that all of the actual light rays converge on the observer position. Several options are available to address this issue:

1. One could neglect any deviations from the straight path from source to observer, in effect replacing  $\Phi(\vec{x}(\theta, \chi'), \chi')$  with  $\Phi(f_K(\chi)\theta, \chi')$ . This assures convergence of the light rays at a later, predetermined time, and the rays can be set up at a suitable timestep in the  $N$ -body simulation and propagated forward in time on-the-fly. This straight-path approximation is known as the Born approximation and is one of the most commonly employed ray-tracing approximations.
2. It is also a possibility to run the  $N$ -body backwards in time. Starting with a normal  $N$ -body simulation it is in principle possible to reverse it after it has reached a certain time by reversing the gravity and the expansion of the box. This leaves out the possibility of having irreversible interactions occur such as dissipation in gas physics. Furthermore, unless the exact same timestepping for each particle is used in the backward run as in the original, the end configuration will be different than the initial configuration. This is a consequence of the integration scheme in GADGET-2, where particles can have different timesteps, not being strictly symplectic (although close to). There is also the possibility that the reverse simulation will be dominated by numerical noise if it is run too far into the past - in effect creating a glass-like configuration.
3. If at each timestep the particles in the observer's backward lightcone are dumped to the disk the ray-tracing can be done by post-processing. This is in practise not viable due to I/O operations being a bottleneck in simulations and, of course, also due to the needed disk space. These limitations allow only a section of the backwards lightcone to be stored on the disk.

While the Born approximation is commonly employed and allows easy on-the-fly ray-tracing we wish to circumvent this approximation and test its validity. Due to the

intricacies in a backward  $N$ -body simulation this approach is rather uncertain. Furthermore, no quantitative studies of reversing  $N$ -body simulations have been published. This leaves only the option of storing a section of the observer's backward lightcone on the disk, and doing the ray-tracing as post-processing.

In addition to the Born approximation, many previous ray-tracing techniques furthermore employ the Limber approximation and neglect lens-lens coupling. The Limber approximation is simply replacing the two dimensional Laplacian in the diagonal of the distortion matrix eq. (1.69) by the three dimensional Laplacian and utilising the Poisson equation for the density field. For our present purpose there is no need to employ this approximation (other than to test its validity). Lens-lens coupling is the appearance of the distortion tensor in the integrand in eq. (1.69). By ignoring lens-lens coupling one effectively sets  $A_{kj} = \mathcal{I}$  in the ray-tracing. Employing the Born approximation and neglecting the lens-lens coupling corresponds to a first order limit of eq. (1.69). Physically, this means that it is implicitly ignored that light from a distant source sees a distorted image of the lower redshift matter due to inhomogeneities along the line of sight. However, as shown in Li et al. (2010) the lens-lens coupling can actually be calculated semi-analytically and efficiently by interpolating the potential from a grid onto the rays. Consequently we will also include lens-lens coupling in our ray-tracing algorithm.

An important consideration when doing ray-tracing is resolution. A source plane at redshift  $\approx 2$  corresponds to a comoving distance of  $\approx 3.6$  Gpc/h in a  $\Lambda$ CDM universe. Thus the box size of the  $N$ -body simulation should be larger than this value to avoid repetition along the line of sight. This will severely compromise the formation and evolution of small scale structure unless a very large particle number is used, and it is consequently not a viable option. One way to circumvent this complication is to "tile" a series of independent  $N$ -body simulations and decreasing the box size as the rays converge on the observer. See White and Hu (2000) for a discussion of this tiling technique. The approach we use is due to Hilbert et al. (2008), and utilises the periodicity of the simulation box instead. Suppose we choose a line of sight parallel to the simulation box edges,  $\vec{L}_{los} = (1, 0, 0)$  for example. In that case the light rays along line of sight will experience the same structure, albeit at different times, after each box traversal (which is normally of order 512 Mpc). If we instead choose the line of sight at a skewed angle, light rays will travel further before encountering the same structure. In fact, with a box size of  $B$  and a suitable choice of  $\vec{L}_{los} = (n_1, n_2, n_3)$  where  $n_1, n_2$  and  $n_3$  are co-prime, the periodicity along the line of sight is  $|\vec{L}_{los}|B$ . The periodicity is ultimately limited by the intersection of the chosen field of view with itself. With this method the light rays propagate in the same simulation volume, and it is only necessary to run a single  $N$ -body simulation.

The algorithm we employ can, due to the above considerations, be condensed to the following steps in the modification of the GADGET-2 code:

1. At a user defined timestep a plane with a specified field of view is created in the simulation volume. This plane is oriented along a line of sight as previously described. In this plane a number of new particles are created on a grid, all massless and all moving with the speed of light in the direction of the line of sight. Note that the particles thus move parallel to each other and do not converge at a point. The separation of the particles will be of the order of the PM grid normally.
2. These light particles only need to experience the gravitational potential calculated on the PM mesh (although including the potential contribution from the Tree part is simple enough), and thus their timestep can be larger than that of normal particles. Each time the PM part of the gravitational potential is calculated, it is interpolated onto the ray particles and saved to disk. The ray particles are then advanced along the line of sight with the speed of light.
3. This process continues until the last step in the simulation is reached, with the potential grid being stored at each timestep where the long-range part of the potential is being calculated.

With this procedure a sequence of potential grids is created for a range of timesteps, and each grid contains the long-range part of the potential. The actual ray-tracing is then done after the simulation itself by solving eq. (1.69) to find the distortion matrix starting from an observer and moving back in time using an interpolation scheme as in Li et al. (2010). This interpolation technique calculates the potential for each point on the ray's path and as such is basically exact within the weak lensing approximation (at least to the accuracy provided by the PM grid and timestepping in the simulation). The interpolation is possible since each ray at all times is confined within successive cubes defined by the grids on two potential planes, and the ray is assumed to move on a straight line within these cubes. The tracer rays, originating from the observer, each have a slightly different initial direction corresponding to the chosen field of view. Note that with this method it is easy at a later stage to make a tomographic analysis simply by following the ray tracers to different timesteps. This method is also called a pencil beam ray-tracing technique due to the geometry of the ray tracers, and it is naturally not suited for a full sky survey.

Note that this does not guarantee that all of the tracer rays stay within the rectangular section of the simulation box defined by the grids. However, one can hope that, with weak lensing only slightly bending the rays in random directions, this "leakage" of rays outside the section is negligible. Furthermore, up to a certain redshift, no tracer rays will have leaked since they all start at the observer and only slowly diverge. In any case, a smaller part of the full field of view can be utilised by slightly altering the initial direction of the tracer rays. If it is found that this still does not prevent a leakage of rays at a satisfactory level, the Born approximation can be employed, or

the  $N$ -body simulation can be rerun with a larger field of view. The first option is by far the best of these since the Born approximation is believed to be quite good.

During the actual ray-tracing it is assumed that the individual tracer rays actually move from one grid onto the next, in effect making most rays move slightly faster than the speed of light. This is done to prevent any rays from lagging behind. In addition, the  $N$ -body potential planes should in fact not be planes but spherical caps. Neglecting these two effects corresponds to the commonly employed flat-sky approximation. Naturally this only works for sufficiently small field of views. In a future version the algorithm may be changed to create the potential grids by pixelation of a sphere, for example by the HEALPIX algorithm.

Assuming  $1024^2$  particles in the  $N$ -body potential plane this corresponds to about 4 MB of data stored at each timestep, since only the potential itself (assumed to be a 4-byte float) needs to be stored<sup>a</sup>. The actual grid position can easily be inferred from the starting redshift and the background cosmology, conserving disk space. With a normal simulation having of the order  $10^3 - 10^4$  evaluations of the long-range potential this should amount to a manageable portion of disk space, especially since the ray particles will only be created after a certain redshift so it is not necessary to dump the potential for all timesteps. Usually any ray-tracing beyond  $z \simeq 2$  is redundant based on observational limitations.

## 3.2 Current state and outlook

At the time of writing the modification of GADGET-2 is complete, that is, GADGET-2 has been modified to allow ray-tracer particles in a plane as previously described, and neutrinos are fully incorporated. The interpolation and actual ray-tracing is still to be done. This will be done in a separate application that reads in the planes and does the interpolation and integration, and it is expected to be fully parallel and MPI compliant. A major goal is to design the ray-tracing post-processing application so that it can be made public and used by other groups (especially within the EUCLID consortium). It should in this regard be noted that a considerable amount of time has been spent on the overall design of the application and producing efficient, yet easily modifiable by third parties, code and coding practises.

One of the most precise ray-tracing simulations (see Hilbert et al. (2008)) utilises snapshots from the Millennium simulation, and despite using a different ray-tracing algorithm than the one described above, it is expected that our implementation will yield somewhat similar results - if neutrinos are disregarded. Since neutrino effects have never been included in ray-tracing simulations, and since the relevant scales for weak lensing are in the non-linear regime, it is difficult to make any predictions on

---

<sup>a</sup>And a negligible amount of header information such as the current timestep etc.

what to expect in the full analysis. Hilbert et al. (2008) shows that the convergence power spectrum very accurately models that of the first order prediction (employing the Born approximation and neglecting lens-lens effects), and the shear power spectrum likewise. However, there is a considerable discrepancy with popular fitting formulas. They furthermore confirm the assertion that the B-mode of the shear field is unimportant for any weak lensing survey, limiting the shear to one degree of freedom. These results indicate that it might be reasonable to apply the Born approximation and ignore lens-lens effects, and thus making ray-tracing substantially easier from an algorithmic point of view. We hope that our results with neutrinos included will confirm these results.

In the near future, with a somewhat small amount of work (and sanity checks), the ray-tracing could be finished so the weak lensing statistics in a ray-tracing neutrino- $N$ -body simulation can be quantified. Since one of the design goals of the project is to have a public and accessible high quality code some work should go into code-restructuring and maintenance. In a slightly more distant future the ray-tracing algorithm will be modified to examine the dependence of the weak lensing statistic on the exact ray-tracing technique applied. For example, instead of choosing a specific line of sight, one could apply a random rotation, translation and parity operation on the ray particles after they have traversed a box length. This would make it less probable for the ray tracers to encounter the same structure twice at the cost of destroying continuity in the mass distribution. Experiments will also likely be conducted to examine the effect of including the short range part of the potential in the potential planes. Another future goal is to change the underlying  $N$ -body simulation to model non- $\Lambda$ CDM cosmologies such as dynamical dark energy models, which will be the subject of chapter 4. Weak-lensing of  $\Lambda$ WDM cosmologies would also be interesting because of the damping of small scale structure (however this effect should be most pronounced for  $z \gtrsim 4$ ).

In conclusion: This project has the aim of performing ray-tracing in an  $N$ -body simulation with neutrinos. A major goal is to provide a robust and user-friendly public code. This aim has been partially fulfilled. So far a new, simple and accurate method for including neutrinos in  $N$ -body simulations has been developed. The overall algorithmic structure of the ray-tracing has been decided, and modification of the GADGET-2 part is complete. With a few months work the actual stand-alone ray-tracer code capable of handling ray-tracing from any  $N$ -body simulation providing a suitable dataset could be completed. It should be noted that a public ray-tracing algorithm by the name of Raytrix do exist, however it is not documented and comes with no user guide, see Hamana (2010). Raytrix utilises the multiple-lens algorithm and takes  $N$ -body input, however, it is not likely to be compatible with the approach described above.

## DYNAMICAL DARK ENERGY IN $N$ -BODY SIMULATIONS

Several authors have implemented dynamical dark energy models in  $N$ -body simulations, see Baldi (2012) for a good overview. In this section we will present our model to implement a dark energy component into an  $N$ -body simulation. Although the original aim of the project was to do a full non-linear treatment of a dark energy fluid, we realised that this is ambiguous. Instead we opted for a linear treatment<sup>a</sup> within an  $N$ -body simulation. Our approach, however, is still valid for a wide class of dark energy models and has the advantage that it is very easy to use. Furthermore, it can in principle be used to do a linear treatment of any other fluid in an  $N$ -body simulation, for example the neutrino fluid approximation that is the topic of chapter 6. It is our goal to make the code public, and a paper demonstrating the code, by improving the Halofit results of Takahashi et al. (2012) to a wider range of dark energy models, is in the works and should be published shortly.

The standard  $\Lambda$ CDM model of cosmology introduces a cosmological constant  $\Lambda$  to drive the acceleration of the expansion at late times  $a \gtrsim a_{m\Lambda} \sim 0.75$ . This cosmological constant has an equation of state  $w = -1$  that is constant throughout space-time. However, any component with  $w < -1/3$  can provide an acceleration in the expansion as can easily be seen from eq. (1.6). A component that acts as a late time accelerator is called dark energy, and there exists a myriad of different models both with time and space dependent equation of state parameter  $w$ , so called dynamical dark energy (DDE) models. The equation of state parameter can also be viewed as the "pa-

---

<sup>a</sup>Partly due to time constraints.

parameter of relevance" since this determines when the dark energy component starts to contribute appreciably to the background energy (and pressure) density. Furthermore, it also determines the magnitude of the acceleration. Many models implement the dark energy component as a fluid with a sound speed  $c_s$ . This sound speed serves as a scale that inhibits dark energy self clustering, and below this scale the dark energy component acts as a constant (in space) background contribution. However, the dark energy is not in any sense a classical fluid. As we will see shortly, non-linear treatment of a fluid DDE model can be ambiguous.

## 4.1 A relativistic fluid

Since our goal is to include dark energy in  $N$ -body simulations we work in the Newtonian limit where velocities  $|\mathbf{u}| \ll c$  and the Newtonian potential  $\Phi \ll c^2$ . A full general relativistic treatment of an ideal fluid in this limit gives the following form of the continuity and Euler equation

$$\dot{\rho} + 3\mathcal{H}(\rho + P/c^2) + \nabla \cdot [(\rho + P/c^2)\mathbf{u}] = 0, \quad (4.1)$$

$$\dot{\mathbf{u}} + \mathcal{H}\mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \frac{\nabla P + \mathbf{u}\dot{P}/c^2}{\rho + P/c^2} + \nabla\Phi = 0. \quad (4.2)$$

Here  $\mathcal{H} = aH$ ,  $\rho$  is the physical dark energy density and  $\mathbf{u}$  is the dark energy fluid velocity. The gradients are co-moving and the overdots means derivative with respect to conformal time as before. See for example Peebles (1980) or Kodama and Sasaki (1984) for an extensive treatment or Anselmi et al. (2011) for a concise introduction. The pressure  $P_w$  is related to the density  $\rho_w$  through an equation of state  $P(\rho_w)$ . In the limit we are working in, the Poisson equation takes the form

$$\nabla^2\Phi = -4\pi G a^2 \sum_{\alpha} (\delta\rho_{\alpha} - 3(\bar{\rho}_{\alpha} + \bar{P}_{\alpha}/c^2)\mathcal{H}\chi_{\alpha}/c^2), \quad (4.3)$$

where  $\mathbf{u} = \nabla\chi$ . Going to Fourier-space this takes the form

$$k^2\Phi = -4\pi G a^2 \sum_{\alpha} \left( \delta\rho_{\alpha} + 3(\bar{\rho}_{\alpha} + \bar{P}_{\alpha}/c^2) \frac{\mathcal{H}^2}{c^2 k^2} \frac{\theta_{\alpha}}{\mathcal{H}} \right), \quad (4.4)$$

where  $\theta = \nabla \cdot \mathbf{u}$ .<sup>b</sup> In the Newtonian  $N$ -body limit the last term can be neglected so that the final Poisson equation takes the form

$$k^2\Phi = -4\pi G a^2 (\delta\rho_m + \delta\rho_w). \quad (4.5)$$

This equation is still valid for non-linear densities and is a modified form of the normal GADGET-2 Poisson equation (2.16).

<sup>b</sup>Note the slightly different definition of  $\theta$  here and in section 1.2 (where  $\theta$  has units 1/length).

The continuity and Euler equations in eqs. (4.1) and (4.2) are not in a closed form until an equation of state has been specified. At the background level the equation of state parameter  $w$  determines the relation between pressure and density

$$\bar{P}/c^2 = w\bar{\rho}. \quad (4.6)$$

The ratio of the perturbed quantities is set by the sound speed

$$\delta P = c_s^2 \delta \rho. \quad (4.7)$$

In general the sound speed  $c_s^2$  is a gauge dependant quantity, since it depends on the fluid rest frame. A gauge invariant sound speed  $\hat{c}_s$  can be defined by writing the sound speed in the fluid rest frame to linear order

$$\hat{c}_s^2 \delta = c_s^2 \delta + 3(1+w)(c_a^2 - \hat{c}_s^2) \frac{\mathcal{H}^2}{c^2 k^2} \frac{\theta}{\mathcal{H}}. \quad (4.8)$$

Here the adiabatic sound speed  $c_a^2 = \dot{P}/\dot{\rho}$  has been defined with  $c_a^2 = wc^2$  for a constant equation of state parameter  $w$ . The gauge effects of eq. (4.8) cannot be neglected for early times as we shall see shortly.

Unfortunately choosing  $w$  and  $c_s^2$  independently is not consistent with a full non-linear treatment of a dark energy fluid. If  $w$  and  $c_s^2$  are chosen independently there is no guarantee that  $\rho + P/c^2$  will stay non-negative, and this makes the fluid rest frame undefined as we will now see. The rest frame is defined to be the unique frame in which the stress-energy tensor  $T_{\mu\nu}$  is diagonal. However, the off-diagonal terms of  $T_{\mu\nu}$  are all proportional to  $\rho + P/c^2$ , so in the limit  $\rho + P/c^2 = 0$  all vectors are eigenvectors of the  $T_{\mu\nu}$ , and the rest frame is consequently no longer uniquely defined. Since the fluid sound speed is inherently a quantity dependent on the rest frame of the fluid, this in turn means that the sound speed is ambiguous. Crossing the "phantom divide" defined by  $\rho + P/c^2 = 0$  therefore correspondingly loses physical meaning. This argument, due to Kunz and Sapone (2006), shows how the fluid description is inconsistent in the limit  $\rho + P/c^2 = 0$ . This is, in fact, already apparent from the Euler equation which diverges at the phantom divide. Furthermore, crossing into  $\rho + P/c^2 < 0$  breaks the weak energy condition of general relativity and physically corresponds to having the momentum of a fluid being opposite the flow (which is always the case for fluid descriptions of phantom dark energy models). The Euler equation diverges if

$$\delta = -\frac{1+w}{1+c_s^2/c^2}, \quad (4.9)$$

and this can happen even for modest densities if  $w \sim -1$ . This will lead to a severe numerical instability when the phantom divide is approached, and is a cause of concern for any fluid model. For this reason, in order to consistently do a full non-linear

treatment of fluid dark energy models, one has to resort to alternative models where a self-consistent equation of state is specified. For example, a barotropic model with an equation of state only dependent on the density  $P(\rho)$  could be tailored to ensure that no crossing takes place. See Linder and Scherrer (2009) for a convenient parametrization of barotropic models.

Since a full non-linear treatment of a dark energy fluid is inconsistent in the absence of a full equation of state, one has to resort to the linearised continuity and Euler equation.

$$\dot{\delta} = -(1+w)\theta - 3\left(\frac{\hat{c}_s^2}{c^2} - w\right)\mathcal{H}\delta - 9(1+w)(\hat{c}_s^2 - w)\frac{\mathcal{H}^2}{c^2 k^2}\theta \quad (4.10)$$

$$\dot{\theta} = -\left(1 + 3\frac{\hat{c}_s^2}{c^2}\right)\mathcal{H}\theta + \frac{\hat{c}_s^2 k^2}{1+w}\delta + k^2\Phi. \quad (4.11)$$

The last term proportional to  $\frac{\mathcal{H}^2}{c^2 k^2}$  in the continuity equation is coming from the gauge dependence of the sound speed, and cannot always be neglected. This is because of the acoustic term in the Euler equation being of the same order at early times (Anselmi et al., 2011). For the scales of  $N$ -body simulations, however, it is small and could be neglected for most realistic scenarios. Embedding a linear component in a non-linear environment is similar to how neutrinos are some times implemented in  $N$ -body simulations as we shall see in chapter 6. Here we will see that a full non-linear treatment is not necessary to account for the change in the total matter power spectrum. With this in mind, we don't consider a linear treatment of a dynamical dark energy component a serious drawback.

## 4.2 Implementation in GADGET-2

Since GADGET-2 is inherently an Eulerian code, the continuity and Euler equations are best solved on the background PM mesh. For this reason the accuracy of the DDE simulation are at the level of the grid size. Since many DDE models usually modify the power spectrum the most at scales around  $k \sim 1 \text{ Mpc}^{-1}/h$  a suitable PM mesh has to be chosen for the simulation (Alimi et al., 2010).

We implement the linearised continuity and Euler equations (4.10) and (4.11) on the PM mesh and integrate them forward in time for every PM timestep. The dark energy perturbations exist in Fourier space, and interacts with matter only through the gravitational potential. The dark energy contribution to the potential is not Gaussian smoothed into a long range and short a range contribution since it is applicable at the grid scale. The initial dark energy perturbations are read from a file onto a grid and the initial velocity divergence is set as  $\theta = \mathcal{H}\delta$ . The equations to be solved in Fourier space can easily be modified from the linear fluid equations to another set of equa-

tions if necessary, for example including a time varying equation of state parameter is straightforward.

We have also implemented a framework to solve the full non-linear continuity and Euler equation on the PM mesh directly in the simulation. Due to the non-linearity these are kept in real space, and it is necessary to track each individual component of velocity and not just the divergence. All the MPI structures and communication steps are properly set up, and it should be relatively easy to get the code to run for a given equation of state, although a few minor modifications remain to be done. As a starting point we recommend employing an equation of state

$$\delta p = c_s^2(\mathbf{k}, t)\delta\rho \quad (4.12)$$

tailored by some physically motivated theory, and subject to the additional constraint that the phantom divide is not crossed. This could be done, for example, by requiring that  $c_s^2$  tends to  $w$  as the phantom divide is approached as is the case for baryons (the physical meaning of an imaginary sound speed should probably not be pushed too far). Kunz and Sapone (2006) provides a good discussion of the intricacies of crossing the phantom divide. Here it is also shown how the phantom divide, within linear perturbation theory, can be crossed in a well-defined manner, if only a proper frame (i.e. not rest frame) is chosen for the pressure perturbations.

### 4.3 Halofit

Halofit is a fitting form for the total non-linear power spectrum proposed by Smith et al. (2003) built on earlier fitting forms, and further develops the abstractions in the 'halo model' (see for example Ma and Fry (2000)). The fit was further improved to include neutrinos by Bird et al. (2012a), and a background dynamical dark energy component with equation of state  $w \neq -1$  by Takahashi et al. (2012). The basic idea of the halo model is to decompose the density field into a distribution of isolated halos. Correlations in the total field can be then be broken down into correlations between distinct halos in a quasilinear, large scale term, and correlations in the matter distribution within single haloes in a non-linear small scale term. For this purpose Halofit introduces the transition scale  $R_G = k_\sigma^{-1}$  that separates the power spectrum into the two terms: A two halo quasi-linear term  $\Delta_Q^2$ , and a non-linear one-halo term  $\Delta_H^2$ . The full power spectrum is thus

$$\Delta_{\text{NL}}^2 = \Delta_Q^2 + \Delta_H^2 \quad (4.13)$$

with the convention  $\Delta^2 = k^3 P(k)/2\pi^2$ . The scale  $k_\sigma$  is usually identified by smoothing the linear power spectrum  $\Delta_L$  with a Gaussian filter (unlike the spherical top-hat of eq. (1.41)) so that

$$\sigma^2(R_G) = \int \Delta_L^2 \exp(-k^2 R_G^2) \frac{dk}{k} = 1, \quad (4.14)$$

since this is a characteristic scale for where linear theory is no longer applicable. With  $y = k/k_\sigma$  the two halo terms have the form

$$\Delta_Q(k)^2 = \Delta_L^2(k) \left\{ \frac{[1 + \Delta_L^2(k)]^{\beta_n}}{1 + \alpha_n \Delta_L^2(k)} \right\} \exp(-y/4 + y^2/8) \quad (4.15)$$

$$\Delta_H(k)^2 = \frac{a_n y^{3f_1(\Omega_m)}}{(1 + \mu_n y^{-1} + \nu_n y^{-2}) \left( 1 + b_n y^{f_2(\Omega_m)} + [c_n f_3(\Omega_m) y]^{3-\gamma_n} \right)}. \quad (4.16)$$

The parameters  $(\alpha_n, a_n, \dots, \gamma_n)$  in these terms are then polynomials (or log-polynomials) in the quantities

$$\begin{aligned} n_{\text{eff}} + 3 &\equiv \left. \frac{d \ln \sigma^2(R)}{d \ln R} \right|_{R_G} \\ &= \frac{2}{\sigma^2(R)} \int d \ln k \Delta_L^2(k) y^2 \exp(-y^2) \Big|_{R_G} \end{aligned} \quad (4.17)$$

$$\begin{aligned} C &\equiv - \left. \frac{d^2 \ln \sigma^2(R)}{d \ln R^2} \right|_{R_G} \\ &= (3 + n_{\text{eff}})^2 + \frac{4}{\sigma^2(R)} + \int d \ln k \Delta_L^2(k) (y^2 - y^4) \exp(-y^2) \Big|_{R_G}, \end{aligned} \quad (4.18)$$

and the relevant cosmological parameters (such as  $\Omega_m$  for the non-linear growth-factors  $f$ ).

## 4.4 Current state

As stated above the linear equations are fully implemented in GADGET-2, and a series of runs has been produced in order to explore additional dependencies of the Halofit parameters on the dark energy sound speed. We expect this will result in a publication shortly. In the longer term we would like to address the effects of fully non-linear models on the power spectrum. Furthermore, we also hope to combine the dynamical dark energy models with a full ray-tracing method like the one described previously.

## **Part III**

# **Scientific research and publications**



## WARM DARK MATTER

Warm dark matter has for some time been a contender to CDM to fill the spot as the missing matter component in the universe. As mentioned in section 1.5.1 the standard  $\Lambda$ CDM paradigm has difficulty in accurately modelling the small scale features of the dark matter clustering. Warm dark matter provides an attractive alternative as it inherits the CDM large scale properties where  $\Lambda$ CDM has proven itself extremely successful. Furthermore, WDM models only introduce a single new parameter: the free-streaming scale, and is therefore a minimal extension.<sup>a</sup> Only on scales comparable to the free-streaming scale and below do the two models differ. On these small scales, however, baryonic processes also play an unknown but potentially important role in the clustering properties especially at late times. These processes are difficult to simulate since very small scale phenomena like supernovae suddenly have significant impact on much larger scales. In this light, careful investigation of competing models is important and they should be tightly constrained in order to determine their viability. If WDM models fail the constraints imposed by high- $z$  observations, the next place to look may very well be more accurate hydrodynamical models or a careful re-analysis of established parameters. For example, the too-big-to-fail problem could be solved if the Milky Way halo is found to be slightly less massive than currently thought (Wang et al., 2012), although observations of the proper motion of the Leo I dwarf spheroidal located in the Milky Way's vicinity seem to indicate otherwise (Boylan-Kolchin et al., 2013).

---

<sup>a</sup>One could make the case that it is not an extension *per se*, but simply choosing a different dark matter particle mass as there is a one-to-one correspondence between the particle mass and free-streaming length for a thermal relic.

## 5.1 Current constraints on warm dark matter

Several independent probes have been used to constrain the WDM particle mass (that can be translated into a free-streaming length). Most probes utilise high- $z$  data since the impact of WDM would be most pronounced at these early times simply because gravity will have had less time to act, and correspondingly, the matter distribution will clearly show the characteristic break that marks the bottom-up up formation regime from the top-down regime. The WDM particle cannot be too light and act as hot dark matter. Hot dark matter models have been excluded for some time due to the very limited structure formation. For neutrino hot dark matter, the amplitude of the CMB fluctuations imply that hot dark matter would only start to go non-linear today. Kang et al. (2013) showed that WDM models with  $m_{\text{thermal}} \lesssim 0.75$  keV could not simultaneously reproduce the stellar mass function and the Tully-Fisher relation. Barkana et al. (2001) gave the same estimate based on the Gunn-Peterson limit: the fact that the universe is reionized no later than at  $z \sim 6$ . A simple analysis by Polisenky and Ricotti (2011) found that  $m_{\text{thermal}} \gtrsim 2.2$  keV by assuming that the number of satellites in WDM simulations equal or exceed that of the Milky Way.

de Souza et al. (2013) utilises the gamma-ray burst database Swift to give a lower bound on the particle mass. Gamma-ray bursts (GRB), being some of the the most luminous events observed in the universe, can be observed very far into the early universe. Since GRBs are believed to be the result of core collapse supernovae of very heavy stars they are a useful tool to probe the smallest scales. Assuming that the GRB population is a fair tracer of the SFR, observing GRBs origination from the very early universe can give a measure of the early SFR. de Souza et al. (2013) makes the simple assumption that the GRB rate is a product of the SFR and a power law in  $z$ , however, this relation naturally has a large degree uncertainty given the observational difficulties in obtaining the high- $z$  SFR. Since high metallicity tends to prevent the formation of very heavy stars, the GRB rate could be rapidly suppressed once a significant build-up of metals have taken place. In any case, in the lack of any hard evidence of a clear GRB/SFR-bias, a simple and conservative relationship is justified. Using a maximum likelihood analysis the constraint on the (thermal) WDM particle mass from this approach is  $m_{\text{thermal}} \gtrsim 1.6 - 1.8$  keV.

The most stringent constraints on the WDM particle mass comes from the Lyman- $\alpha$  forest (Boyarsky et al., 2009a; Narayanan et al., 2000; Seljak et al., 2005; Viel et al., 2005, 2013). A decent review of the constraints from Lyman- $\alpha$  is given in Markovič and Viel (2013) which also presents a short introduction to future constraints from weak lensing surveys like Euclid. Observations of Lyman- $\alpha$  forest are useful in constraining WDM models since they probe the high- $z$  universe in a large range of redshifts ( $z \sim 2 - 6$ ) and the matter power spectrum when it is starting to go non-linear - optimal conditions to differentiate a WDM model from a CDM model. The basic idea is to relate the flux fluctuations  $\delta_{\text{flux}}$  to the the matter fluctuations  $\delta_{\text{matter}}$ . This

is usually done by the effective bias method

$$P_{\text{flux}}(k) = b^2(k, z)P_{\text{matter}}(k), \quad (5.1)$$

pioneered by Croft et al. (2002), and then calibrating the bias  $b$  by numerical simulations that incorporate baryonic effects to properly account for the thermal history of the intergalactic medium (IGM). Viel et al. (2013) derive constraints based on high resolution quasar data from the Keck HIRES spectrometer and the Magellan spectrograph MIKE in the redshift range  $4.48 < z < 6.42$ . Utilising a mock quasar sample generated by a grid of 54 hydrodynamical  $N$ -body simulations with a simple star formation criterion a conservative range of thermal histories of the IGM is examined. Since the data is in the high- $z$  universe feedback processes are believed to have much smaller impact, and this removes a significant cause of uncertainty on the thermal properties of the IGM. It should be noted, however, that the uncertainty (and the uncertainty on the uncertainties) is not very well established for the high- $z$  universe, so a conservative parameter range is indeed necessary. They find that  $m_{\text{thermal}} > 3.3$  keV and, importantly, also find that the effect of WDM free-streaming is not significantly degenerate with other physical effects.

Tremaine and Gunn (1979) gave a constraint for any fermionic dark matter candidate known as the Tremaine-Gunn bound. Essentially the argument is that for a particle following a Fermi-Dirac distribution the average phase-space density cannot exceed the degenerate, microscopic phase-space density. For a non-interacting species only experiencing free-fall, the Vlasov equation (the Boltzmann equation (1.36) with zero interaction term) states that the phase-space density is conserved, and hence the present phase space density cannot exceed the maximum of the infinitesimal (fine-grain) equilibrium distribution. This maximum depends sensitively on the fermion mass. In other words, in a finite (coarse-grained) phase-space volume the averaging over volume can only decrease the phase-space density. Since phase-space information can be inferred from observations of the stellar orbits in galaxies this can be used to constrain the fermion mass.<sup>b</sup> The Tremaine-Gunn bound is most relevant to objects with high density and low velocity dispersion, which make dark matter dominated dwarf spheroidals an obvious case of study. Usually an isothermal density profile is assumed, as this is a decent fit to observations, and the velocity profile can be assumed to be Maxwellian (as it has to be for an isothermal sphere). A warm relic with an appreciable primordial thermal velocity has a tighter limit on the achievable phase-space density, regardless of the current thermal velocity, and this leads

<sup>b</sup>Since only the fermionic equilibrium phase-space density has a maximum value, this method does not work for bosonic dark matter. The approach may work for a non-thermally produced fermionic particle like a Dodelson-Widrow neutrino, however only if a suitable maximum in the distribution function can be defined. In the  $\nu$ MSM model the sterile neutrino spectrum is a rescaled Fermi-Dirac for example.

to a more smooth halo core with lower density and correspondingly larger core radius. Generally the halo core radius scales as  $\sim m_X^{-2}$ , see for example Peebles (1993). Thus studying the cores of dwarf spheroidals can probe the particle mass of a WDM candidate. Shao et al. (2013) performed a suite of  $N$ -body simulations in order to examine the cores of simulated dark matter only dwarf spheroidals. They subsequently applied their results to the dwarf spheroidals of the Milky Way system and found a particle mass of  $m_X \sim 0.5$  keV. It should be emphasised that this is not a *limit* on the particle mass, rather it is an observationally motivated *best-guess* of the particle mass based on the phase-space considerations induced by the WDM initial thermal velocities. A crucial assumption in this approach (besides the obvious assumption that these haloes do have a core) is that the formation of the core is primarily influenced by the dark matter component, which is a seemingly a reasonable assumption given that dwarf spheroidals mainly consist of dark matter. If baryonic processes play an appreciable role, the phase-space density may increase, and the Tremaine-Gunn bound is not applicable. The fact that most constraints on the WDM particle mass puts it in the  $m_X \gtrsim 1.5$  keV range, however, is a source of concern for WDM models relying on a thermally produced fermion unless some energy exchange between the baryonic and dark component takes place in the halo, or dwarf spheroidals prove themselves core-less.

## 5.2 Introduction to the paper

Cosmology has a tradition of using as many independent probes as possible to falsify or tighten the constraints on models, WDM being no different. It is in this light that our work Schultz et al. (2014) should be viewed: as a further examination of the WDM parameter space using a complementary approach to the methods discussed above and victim of different systematics. We here establish that simply counting high- $z$  galaxies, and assuming that galaxies are a fair tracer of dark matter haloes, gives a non-trivial bound on the WDM particle mass. When confronted with a crude reionization model the tension between observations and WDM models is further increased. The method we use is robust and does not necessitate running a large amount of simulations. The overall goal of the work was to show that, even with an extremely conservative approach, the WDM models usually employed ( $m_{\text{thermal}} \sim 0.5 - 1.5$  keV) to match low- $z$  observations are excluded by high- $z$  considerations. Furthermore, we wanted to assess how much the constraints would improve by upcoming James Webb Space Telescope (JWST) observations. The conservative approach serves as a proof-of-concept, and more exhaustive analysis would very likely tighten the constraints considerably. We note that our method is also applicable to other models of warm dark matter than those studied here as long as the star formation in the model is not more efficient than in  $\Lambda$ CDM. Thus our method is not in any

way sensitive to the initial transfer function.

# The high- $z$ universe confronts warm dark matter: Galaxy counts, reionization and the nature of dark matter

Christian Schultz,<sup>1,2★</sup> Jose Oñorbe,<sup>1</sup> Kevork N. Abazajian<sup>1</sup> and James S. Bullock<sup>1</sup>

<sup>1</sup>Center for Cosmology, Department of Physics and Astronomy, 4129 Frederick Reines Hall, University of California, Irvine, CA 92697, USA

<sup>2</sup>Department of Physics and Astronomy, Aarhus University, Ny Munkegade, DK-8000 Aarhus C, Denmark

Accepted 2014 May 13. Received 2014 May 13; in original form 2014 January 15

## ABSTRACT

We use  $N$ -body simulations to show that high-redshift galaxy counts provide an interesting constraint on the nature of dark matter, specifically warm dark matter (WDM), owing to the lack of early structure formation these models. Our simulations include three WDM models with thermal-production masses of 0.8, 1.3 and 2.6 keV, as well as CDM. Assuming a relationship between dark halo mass and galaxy luminosity that is set by the observed luminosity function at bright magnitudes, we find that 0.8 keV WDM is disfavoured by direct galaxy counts in the *Hubble Ultra Deep Field* at  $>10\sigma$ . Similarly, 1.3 keV WDM is statistically inconsistent at  $2.2\sigma$ . Future observations with *JWST* (and possibly *HST* via the Frontier Fields) could rule out 1.3 keV WDM at high significance, and may be sensitive to WDM masses greater than 2.6 keV. We also examine the ability of galaxies in these WDM models to reionize the universe, and find that 0.8 and 1.3 keV WDM produce optical depths to the cosmic microwave background (CMB) that are inconsistent at 68 per cent C.L. with current *Planck* results, even with extremely high ionizing radiation escape fractions, and 2.6 keV WDM requires an optimistic escape fraction to yield an optical depth consistent with *Planck* data. Although CMB optical depth calculations are model dependent, we find a strong challenge for stellar processes alone to reionize the universe in a 0.8 and 1.3 keV WDM cosmology.

**Key words:** galaxies: abundances – cosmology: theory – dark ages, reionization, first stars – dark energy.

## 1 INTRODUCTION

Dark matter dominates the evolution of gravitational perturbations, leading to the formation of haloes and galaxies. In the prevalent paradigm of cold dark matter (CDM), the primordial perturbation spectrum extends to very small scales; galaxy formation proceeds from the bottom up, commencing in the smallest dark matter haloes where gas cooling can occur. If instead there exists a non-negligible minimal scale for primordial perturbations as in the case of warm dark matter (WDM), halo formation is delayed, and early galaxy formation is suppressed considerably.

Early galaxy formation has been understood to be a challenge for WDM models for some time (Barkana, Haiman & Ostriker 2001; Somerville, Bullock & Livio 2003). Today, the tension is only heightened by mounting evidence that structure formation is proceeding in earnest at very early cosmic times. There are now direct detections of galaxies at redshifts as high as  $\sim 10$  (Ellis et al. 2013; Oesch et al. 2013), clearly indicating that there are collapsed structures at this time. More indirectly, studies of quasar spectra show that the intergalactic medium (IGM) was almost fully ionized

by redshift  $z \sim 6$  (Fan et al. 2006) and the measured electron scattering optical depth from the cosmic microwave background (CMB) may imply reionization as early as  $z \sim 10$  (Ade et al. 2013). The maintenance of reionization back to these early times seems to require contributions from numerous, low-mass galaxies (Kistler et al. 2009; Kuhlen & Faucher-Giguere 2012; Robertson et al. 2013). In this paper, we examine how current and future observations of high- $z$  galaxies, together with observational probes of reionization, can constrain the dark matter power spectrum on small scales, and by extension the particle nature of dark matter.

There has been considerable interest in the WDM paradigm for galaxy formation, owing to potential problems with the  $\Lambda$ CDM model on sub-galactic scales. Most recently, it has been recognized that the observed central densities of low-luminosity Milky Way dwarf satellite are significantly lower than expected in dissipationless CDM simulations (Boylan-Kolchin, Bullock & Kaplinghat 2012). This issue can be alleviated if the dark matter is warm (Lovell et al. 2012, 2014; Polisensky & Ricotti 2014). Here, we specifically study a WDM model (1.3 keV thermal particle mass) that corresponds to the cutoff scale that alleviates the central-density problem. It should be noted here that models including ultralight axions alongside a CDM component can also alleviate the sub-galactic problems in the pure CDM models (Marsh & Silk 2014).

\*E-mail: cs06@phys.au.dk

The two most popular classes of WDM particle candidates are ‘thermal’ particles and sterile neutrinos. Thermal WDM is coupled to the primordial plasma in the early Universe, and is diluted to the proper (observed) dark matter density by an unspecified process. Sterile neutrinos, on the other hand, can be produced at the proper dark matter abundance through scattering processes due to their mixing with active neutrinos with the Dodelson–Widrow mechanism (Dodelson & Widrow 1994, sometimes referred to as non-resonant production), through resonant production in the case of a large cosmological lepton asymmetry (Shi & Fuller 1999), or through coupling with other fields (Kusenko 2006; Shaposhnikov & Tkachev 2006). An important characteristic of the different models is the free-streaming length they introduce, with a given particle mass having a different free-streaming length for the different WDM particles and for the different sterile neutrino production mechanisms. In this paper we primarily state particle masses in terms of thermal WDM particles, i.e. the ‘thermal mass’, but we also provide conversions to the Dodelson–Widrow sterile neutrino mass in summary statements and tables. We also quote the wave numbers where the associated power spectra fall to half the value of a standard CDM model, which allows our results to be interpreted generally for any model that results in truncated small-scale power, as can arise for standard CDM particles in the case of non-standard inflation (e.g. Kamionkowski & Liddle 2000; Zentner & Bullock 2002).

Recent work has constrained the warm dark particle mass by a number of methods. Some of the currently tightest constraints come from observations of the Lyman  $\alpha$  ( $\text{Ly}\alpha$ ) forest produced by neutral gas along the line of sight to distant quasars. The neutral gas follows the gravitationally dominant dark matter clustering in the mildly non-linear regime probed by the  $\text{Ly}\alpha$  forest, and therefore it can be a powerful probe of the dark matter perturbation spectrum at small scales. However, the  $\text{Ly}\alpha$  forest is a challenging tool, requiring disentangling the effects of pressure support and thermal broadening of the  $\text{Ly}\alpha$  forest features from the effects of dark matter perturbation suppression from WDM. In addition, modelling the dependence on the physics of the neutral gas requires assumptions of the thermal history of the IGM and its ionizing background, which are done as parametrized fitting functions. Many of the limitations of the  $\text{Ly}\alpha$  forest on constraints of the primordial power spectrum are discussed in Abazajian et al. (2011). Setting aside the limitations of the method, the  $\text{Ly}\alpha$  forest provides stringent constraints, with recent quoted limits at  $m_{\text{WDM}} > 3.3 \text{ keV}$  ( $2\sigma$ ; Viel et al. 2013).

The lack of early structure formation in WDM has motivated limits from the rate of high- $z$  gamma-ray bursts (de Souza et al. 2013). Similarly, Pacucci, Mesinger & Haiman (2013) utilize strongly lensed ultrafaint, high-redshift galaxies to constrain the particle mass by halo mass function considerations. At low redshift, WDM models can be constrained by studying the abundance of small galaxies. Work by Polisensky & Ricotti (2011) and Lovell et al. (2014) uses  $N$ -body simulations of Milky Way-sized dark matter haloes and constrains the particle mass by assuming that the number of simulated dark matter satellites equals or exceeds the number of observed Milky Way satellites, and report limits on thermal WDM particle masses of  $m_{\text{WDM}} > 2.3 \text{ keV}$  and  $> 1.5 \text{ keV}$ , respectively. It should be recognized that constraints from satellite counts are sensitive to halo-to-halo variation in sub-structure counts as well as assumed completeness corrections to the observed Milky Way satellite luminosity function. More recently, Horiuchi et al. (2014) have tried to adequately account for the halo-to-halo scatter and focused on counts around M31 (which are higher than around the MW at fixed luminosity) and find  $m_{\text{WDM}} > 1.8 \text{ keV}$ .

Given the potential systematic problems with known WDM constraints, it is useful to explore alternative probes. In the rapidly evolving field of high- $z$  galaxy surveys, the Lyman-break technique has proven useful for discovering galaxies and estimating the UV luminosity function out to redshifts  $z \sim 9$ , although there are candidates in the literature at redshifts as high as  $z \sim 12$  (Bouwens et al. 2007, 2011; McLure et al. 2013; Oesch et al. 2013; Schenker et al. 2013b). New Fourier techniques seem promising in finding fainter candidates below the normally required detection threshold  $S/N \sim 4.5$  (Calvi et al. 2013). Furthermore, Lyman-break galaxies seem to be fair tracers of the overall halo population (Conroy, Wechsler & Kravtsov 2006). Thus, the UV luminosity function interconnects with the halo mass function of dark matter, a quantity which is readily constructed from simulations and from which different dark matter models can be distinguished.

Additional physical mechanisms may ease the tension between simulations and observations presented by, e.g. the missing dwarf galaxies or the too-big-to-fail problem (Boylan-Kolchin et al. 2012). One example is the work Bovill & Ricotti (2011). Here, it is noted that the Milky Way halo may be populated by fossils of early dwarf galaxies that formed before reionization, and that these fossils today have very low surface brightness rendering them outside of current observational bounds. Therefore, there may indeed be a population of low-luminosity dwarf galaxies near the Milky Way. Katz & Ricotti (2013) expand on this result, and argues that the bulk of the old globular clusters in the Milky Way formed in these first (now fossil) dwarf galaxies. Furthermore, the protoglobular clusters were an important mode of star formation in these galaxies, and could in fact be the main driver for reionization.

In this work, we study in detail the effects of WDM models on high- $z$  dark matter halo counts using high-resolution cosmological simulations, and extend these results empirically to infer the observable effects on galaxies and reionization. We compare the high- $z$  luminosity functions of galaxies recently measured from the *Hubble Ultra Deep Field* (HUDF) with the inferred luminosity function derived using  $N$ -body simulations in WDM and CDM cosmologies. Our predicted luminosity functions are normalized to match observed bright galaxy counts using abundance matching. All magnitudes quoted below are in the AB system. The same models allow us to study cosmological reionization in WDM models and compare them to CDM.

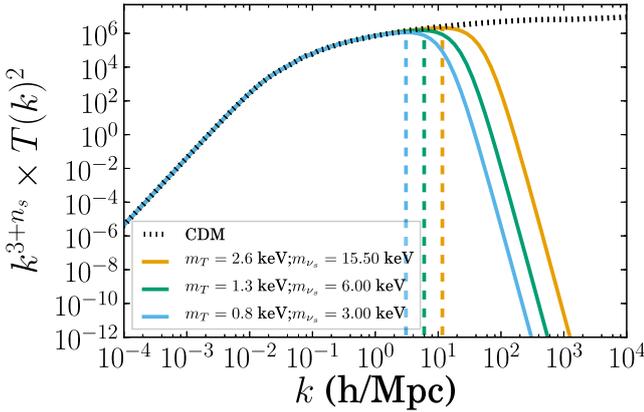
## 2 THEORY AND SIMULATIONS

### 2.1 Power spectrum

WDM has a non-negligible thermal velocity which imprints a free-streaming scale in the matter perturbation distribution arising from the early Universe. Below this free-streaming scale, structure formation is suppressed. This scale is conveniently parametrized by the Jeans mass

$$M_J(z) = \frac{4}{3\pi} \rho_{\text{dm}}(z) \left( \frac{\pi \sigma_v^2(z)}{4G\bar{\rho}(z)} \right)^{3/2}, \quad (1)$$

where  $\rho_{\text{dm}}$  is the dark matter density,  $\bar{\rho}$  the mean density and  $\sigma_v$  the velocity dispersion. The Jeans mass is constant approximately until matter–radiation equivalence and thus erases the initial conditions (ICs) below this mass scale. After matter–radiation equality the Jeans mass drops rapidly and decays with the cooling of the dark matter in the Hubble flow as  $\sim a^{-3/2}$  (Schneider, Smith & Reed 2013a). The transfer function relates the primordial matter power spectrum to the linear power spectrum at a later redshift. For WDM,



**Figure 1.** Shown are the matter power spectra for the different WDM models considered in this paper. The more massive the WDM particle, the more CDM-like the matter transfer function becomes due to the lower thermal velocities. For reference, each model is labelled by its equivalent thermal and sterile neutrino mass. The dashed lines mark the maximum of the transfer function.

the matter power spectrum can be seen as a suppression of power above a certain wavenumber  $k$ . In fact, the transfer function for WDM relative to the CDM case can be approximated by the fitting function (Abazajian 2006):

$$T(k) = [1 + (\alpha k)^\nu]^{-\mu}, \quad (2)$$

where the smoothing scale is set by

$$\alpha = a \left( \frac{m_s}{\text{keV}} \right)^b \left( \frac{\Omega_{\text{dm}}}{0.26} \right)^c \left( \frac{h}{0.7} \right)^d h^{-1} \text{ Mpc}. \quad (3)$$

Here,  $m_s$  is the (non-thermally produced) sterile neutrino WDM particle mass (Abazajian 2006). The relationship between the mass of a thermal particle ( $m_{\text{WDM}}$ ) and the mass of the sterile neutrino ( $m_s$ ) for which the transfer functions are nearly identical (Viel et al. 2005) is

$$m_{\text{WDM}} = 0.335 \text{ keV} \left( \frac{m_s}{\text{keV}} \right)^{3/4} \left( \frac{\Omega_{\text{m}}}{0.266} \right)^{1/4} \left( \frac{h}{0.71} \right)^{1/2}. \quad (4)$$

Fig. 1 shows the matter power spectrum for the three WDM models we explore in this paper (labelled by thermal mass and equivalent sterile mass) along with a CDM model for comparison. For comparison purposes, the scales where these WDM transfer functions equal half the CDM transfer function,  $k_{1/2}$ , are listed in Table 1, along with wavenumber of maximum power  $k_{\text{max}}$ .

**Table 1.** WDM models simulated in this paper. The first column gives the thermal WDM mass. This is the default model label we use throughout the work. The second column gives the equivalent sterile neutrino mass. The last two columns list the wave numbers where the transfer functions reach half the value of CDM and the wave numbers where the power spectrum is maximized, respectively.

$m_T$	$m_{\nu_s}$	$k_{1/2} [h^{-1} \text{ Mpc}]$	$k_{\text{max}} [h^{-1} \text{ Mpc}]$
2.6 keV	15.5 keV	21.4	2.62
1.3 keV	6.0 keV	9.47	1.28
0.8 keV	3.0 keV	5.24	0.764

## 2.2 Numerical simulations

Our simulations were performed with the GADGET-2 code, in TreePM mode (Springel 2005). In order to generate the ICs, we have used the MUSIC code (Hahn & Abel 2011). The method uses an adaptive convolution of Gaussian white noise with a real-space transfer function kernel together with an adaptive multigrid Poisson solver to generate displacements and velocities following second-order Lagrangian perturbation theory (2LPT). For more specific details on the MUSIC code, we refer the reader to Hahn & Abel (2011). The CAMB (Lewis, Challinor & Lasenby 2000; Howlett et al. 2012) package was used to generate the CDM transfer functions used to generate the ICs for this cosmology. The WDM transfer functions were obtained from CDM using equations (2) and (3). Only the ICs were modified, and the thermal velocities of the WDM particles are not included in the simulations, since they have not been found to be significant in affecting WDM structure formation (Bode, Ostriker & Turok 2001; Villaescusa-Navarro & Dalal 2011).

The cosmological parameters used were  $h = 0.71$ ,  $\Omega_{\text{m}} = 0.266$ ,  $\Omega_{\Lambda} = 0.734$ ,  $n_s = 0.963$  and  $\sigma_8 = 0.801$ . All simulations are  $1024^3$  particles in  $(50 \text{ Mpc } h^{-1})^3$  boxes started at  $z_{\text{ini}} = 125$ . The implied particle mass is  $m_p = 8.6 \times 10^6 h^{-1} M_{\odot}$ . We employ three WDM models with thermal particle masses of 0.8, 1.3 and 2.6 keV, which are equivalent to oscillation-produced Dodelson–Widrow sterile neutrino particle masses of 3, 6 and 15.5 keV. Note that the 1.3 keV (thermal; 6 keV sterile) case is equivalent in the structure formation cutoff scale of the M2L25 model of Boyarsky et al. (2009) and Lovell et al. (2012).

Three separate issues require special attention when running these simulations: (1) the dependence of the halo mass function at high redshift on the chosen starting redshift,  $z_{\text{ini}}$ ; (2) systematic errors induced by the finite volume of the simulation; and (3) artificial haloes that emerge in WDM as a result of shot noise in regimes where the underlying power spectrum is suppressed. We discuss each of these issues in turn.

Concerning the initial redshift  $z_{\text{ini}}$ , recent advances in the techniques used for numerical calculation of perturbations, such as the 2LPT (see, for example Jenkins 2010), have improved the convergence of simulations using different  $z_{\text{ini}}$ . Several groups (e.g. Lukić et al. 2007; Prunet et al. 2008; Knebe et al. 2009; Jenkins 2010; Reed et al. 2013) have worked to quantify the effect of  $z_{\text{ini}}$  on the final results of cosmological simulations. These works stress the point that not using 2LPT algorithm leads to simulations that converge very slowly as the start redshift is increased. In order to reduce as much as possible any  $z_{\text{ini}}$  effect we have used the 2LPT algorithm incorporated in MUSIC (Hahn & Abel 2011) to generate all the ICs of our simulations. Additionally, all the simulations presented in this work use the same initial redshift ( $z_{\text{ini}} = 125$ ). Therefore, any systematic effects associated with starting redshift will be present in all cases and cancel when considering the ratios between the WDM and CDM halo mass functions.

Systematic errors from the finite volume of the simulation box can be divided into three categories: shot noise, sampling variance and lack of power from modes larger than the simulation box. Shot noise is especially important for the most massive haloes since only a few exists in the simulation volume, and it generally decreases as  $1/\bar{n}V$  where  $V$  is the simulation volume and  $\bar{n}$  the number count. However, for smaller halo masses, shot noise is dwarfed by sample variance (Hu & Kravtsov 2003). The average density in the simulation volume may happen to be an over- or underdense part of the universe, and since haloes are biased tracers of the density field, this will lead to differences in the halo mass function. The best way to

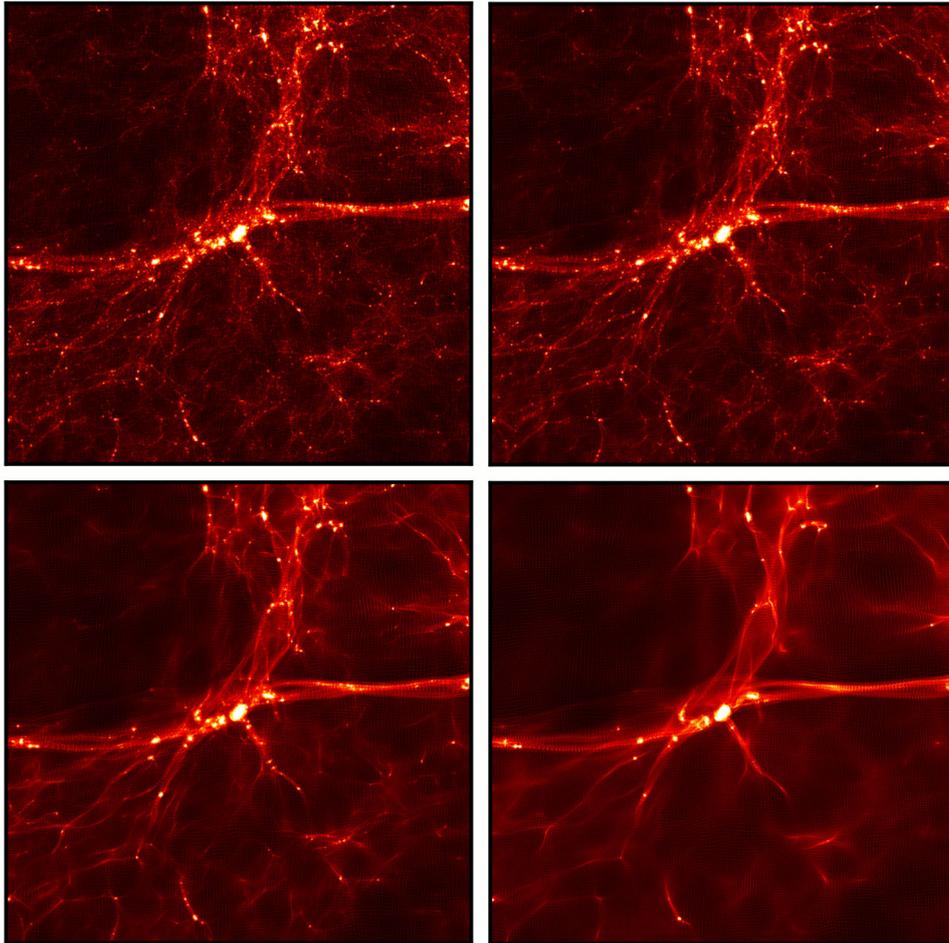
correct this is to run independent samples of the underlying density field (different seeds for the ICs), but this comes at a considerable cost in terms of CPU hours. Alternatively, the sampling variance can be estimated by analytic methods as given by equation 4 in Hu & Kravtsov (2003), with a Sheth–Tormen bias for example. However, such a bias is based on fits to  $\Lambda$ CDM simulations. It seems plausible that such a bias would not change significantly if used in a WDM cosmology as it is primarily determined by nature of halo collapse, but to avoid any complications with the error estimate, we directly calculate the sample variance in the halo mass function by the jackknife technique. We do not consider any contributions to the halo mass function from scales larger than the simulation box, since we are mostly interested in the low-mass end, and the simulated volume is significantly larger than the scale of clusters at the redshifts of interest.

Finally, below a specific mass scale dependent on numerical resolution, it has been well established that WDM produces artificial haloes in simulations (Wang & White 2007; Angulo, Hahn & Abel 2013), an effect of the shot noise due to the finite particle count. These haloes are usually visible as regularly spaced clumps in the filaments of the cosmic web, and they form below a mass scale proportional to  $m_p^{-1/3}$ , where  $m_p$  is the simulation particle mass. However, force resolution also plays a role, and an excessive force resolution, as compared to the mass resolution, can increase the

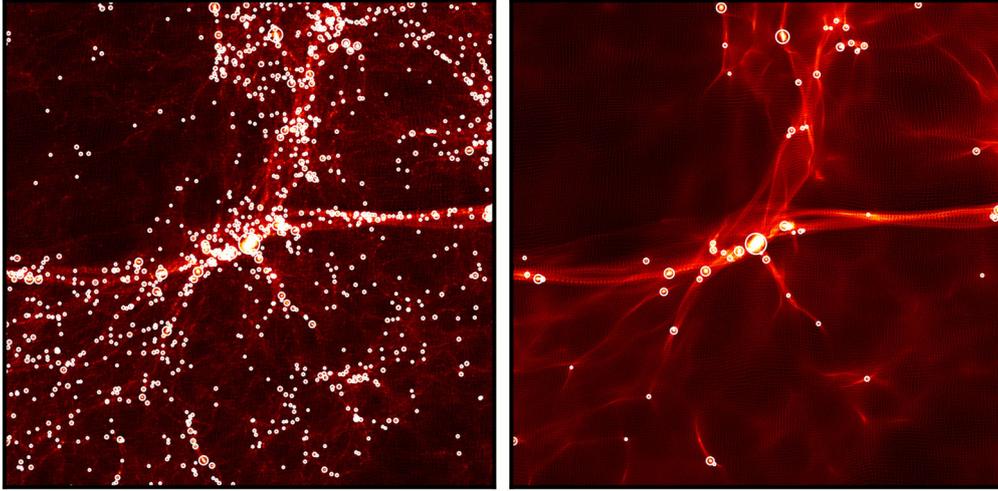
number of artificial haloes (Angulo et al. 2013). Schneider et al. (2013a) showed that the artificial haloes can be modelled by a power law increase in the WDM halo mass function below the mass scale. Most attention has been given to correcting the halo mass function for low redshift, since contamination of the halo population is the largest here, due to the fact that the artificial haloes have had more time to form and accrete.

There have been no focused studies on artificial halo contamination is at redshifts  $z \gtrsim 5$ . In the results presented below, down to the halo mass scale adopted for our completeness limit, we see little if any indication of a low-mass upturn in our WDM halo mass functions; such an upturn would be indicative of significant artificial halo contamination. Moreover, since any artificial haloes present would provide an increase in the halo mass function (thus making WDM more like CDM) ignoring them only makes our WDM constraints more conservative. In what follows, we have conservatively chosen to ignore any corrections for artificial haloes in our catalogues.

Fig. 2 provides a qualitative depiction of the differences inherent in WDM compared to CDM simulations. Shown are  $10 \times 10 \times 6$  (comoving  $h^{-1}$  Mpc) slices of each of our simulation volumes, centred on the most massive halo at a redshift of  $z = 6$ . On large scales the slices look similar, but on smaller scales there is a clear lack of structure in the WDM models.



**Figure 2.** Simulation images for CDM and WDM at  $z = 6$ , each initialized with the same random seed. The panels are  $10 h^{-1}$  Mpc square and  $6 h^{-1}$  Mpc deep; they are centred on the most massive halo in the box. The upper-left panel is CDM, with 2.6 keV WDM in the upper right. The bottom panels correspond to WDM: 1.3 keV (left) and 0.8 keV (right). The lack of structure for the lightest WDM models is striking compared to CDM.



**Figure 3.** Simulation snapshots from CDM (left) and 0.8 keV WDM (right) overlaid with circles to indicate identified dark matter haloes that are more massive than  $3.4 \times 10^8 h^{-1} M_{\odot}$ . The size of the circle is proportional to the virial radius of each halo. The CDM slice is filled with collapsed structure at  $z = 6$ , while the WDM slice is largely devoid of collapsed haloes that are massive enough for hydrogen cooling. Note that artificial haloes would show up as regularly separated haloes in the filaments, suggesting that contamination by artificial haloes is likely negligible here.

### 2.3 Halo catalogues

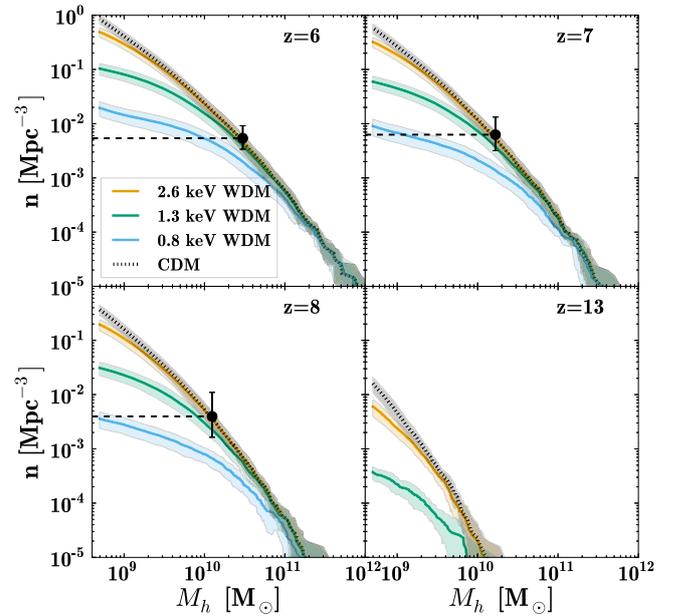
We used the Amiga Halo Finder (AHF; Knollmann & Knebe 2009) to identify haloes in our simulations. The halo mass  $M_h$  used in this work is calculated using the overdensity ( $\Delta_{\text{vir}}$ ) formula from Bryan & Norman (1998) for our cosmology at each specific redshift. Note that our conclusions do not change when using different overdensity definitions, e.g.  $\Delta_{200} = 200\rho_{\text{crit}}$ . As explained above, to build our mass–luminosity relation using the abundance-matching technique we took into account the merger history of each halo and used its maximal mass obtained over its lifetime  $M_{\text{peak}}$  instead of  $M_h$ . In any case, this correction turned out to be small due to the lack of sub-structure at high redshifts. We used a requirement of at least 40 simulation particles to constitute a halo, setting a halo mass completeness limit of  $M_h = 3.4 \times 10^8 h^{-1} M_{\odot}$ .

Compared to the density maps shown in Fig. 2, the differences between WDM and CDM become even more apparent when we compare halo counts. Fig. 3 shows two of the same density slices overlaid with white circles to indicate identified dark matter haloes more massive than our  $M_h = 3.4 \times 10^8 h^{-1} M_{\odot}$  completeness limit. Circle sizes are proportional to the virial radius of each identified halo. The difference in collapsed structures is striking between these two simulations. For example, the void in the upper-left corner is completely empty of any haloes in the 0.8 keV WDM run.

Fig. 4 provides a more quantitative demonstration of the differences in halo abundances from model to model, where each panel shows the cumulative dark halo mass function at redshifts  $z = 6, 7, 8$  and  $13$ . The CDM result (dotted line with shading) is in all cases above the WDM models (solid lines with shading, as labelled). Angulo et al. (2013) found a suppression of the halo mass function of the form<sup>1</sup>

$$\frac{n_{\text{WDM}}(M)}{n_{\text{CDM}}(M)} = \frac{1}{2} \left( 1 + \frac{M_1}{M} \right)^{-\alpha} \left[ 1 + \text{erf} \left( \log \frac{M}{M_2} \right) \right]. \quad (5)$$

<sup>1</sup> Strictly speaking Angulo et al. (2013) has  $\alpha = 1$  fixed; however, they also correct for artificial haloes. We find that keeping  $\alpha$  as a free fitting parameter is necessary to provide reasonable fits, probably owing to a strong evolution with redshift.



**Figure 4.** Shown are cumulative halo mass functions at selected redshifts for our CDM and WDM simulations. Note how the suppression of WDM increases with  $z$ . At  $z > 12$  no dark matter haloes are identified at all in the 0.8 keV WDM model. The central lines denote the simulated halo mass function, while the shaded areas indicate jackknife uncertainties. The horizontal dashed lines correspond to the known lower limit on the cumulative galaxy abundance at each redshift based on the faintest HUDF observations; assuming that galaxies reside in haloes, any viable mode must produce a total abundance of haloes above this line. The point with  $2\sigma$  error bar is placed at the halo mass corresponding to the HUDF luminosity limit as inferred from abundance matching, discussed in Section 3. For constraints, we use the values of the galaxy luminosity functions observed and inferred from the halo abundance-matching method, shown in Fig. 7.

We have verified this expression provides a good fit to the WDM/CDM abundance ratio for  $z \lesssim 10$ , with decreasing accuracy with increasing redshift. In our simulations, at  $10^9 M_{\odot}$ , the 0.8 keV model is suppressed by more

than an order of magnitude at all redshifts relative to  $\Lambda$ CDM.

As can be seen in the  $z = 13$  panel of Fig. 4, no haloes at all exist in 0.8 keV WDM model. Indeed, we find that no haloes have formed before  $z = 12$  for 0.8 keV WDM and none before  $z = 15$  in the 1.3 keV model. Detections at these epochs should be robust in the future with *JWST*. However, even current detections offer an interesting test: the point with error bar ( $2\sigma$ ) corresponds to the *lower limit* on the cumulative abundance of galaxies at those redshifts, as set by the faintest galaxies observed in the HUDF (Bouwens et al. 2007; McLure et al. 2013; Oesch et al. 2013). Its horizontal position (corresponding halo mass) is based on the luminosity limit and our adopted  $M_h-L$  relation presented in the next section. Importantly, the total abundance of galaxies at each redshift must be above the data point shown (regardless of its horizontal positioning on the plot). One can see without any further analysis that the 0.8 keV WDM model will have trouble producing enough galaxies to match *current* observations at  $z > 8$ ; there are simply not enough collapsed objects of any mass to account for the known galaxies at this epoch. The viability of the other WDM models is not immediately apparent from Fig. 4 in itself, since the halo mass function is not directly observed.

In order to provide a more precise connection with observations we will need a mapping between halo mass and galaxy luminosity. This is a primary subject of the next section.

### 3 PREDICTING OBSERVABLES

#### 3.1 Observed luminosity functions

We will normalize our predictions using observed high- $z$  galaxy counts. In doing so, we follow the literature and assume that high- $z$  luminosity function is well characterized by a Schechter function

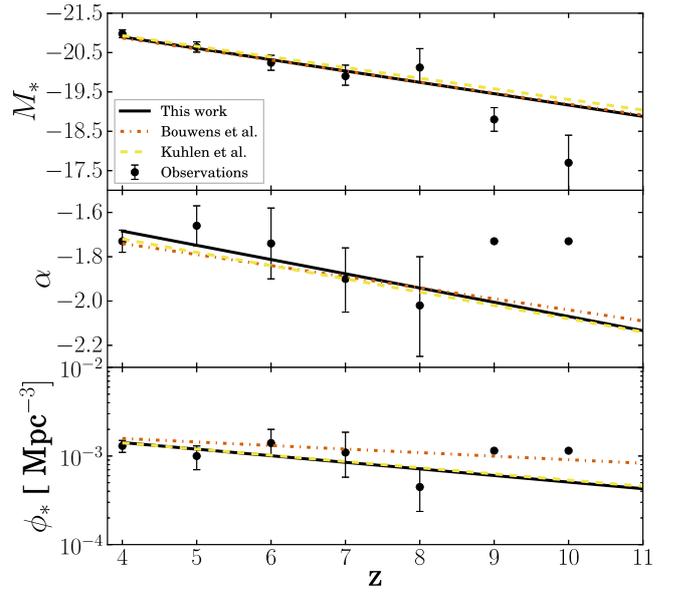
$$\phi(L) dL = \phi_* \left( \frac{L}{L_*} \right)^\alpha \exp \left( -\frac{L}{L_*} \right) \frac{dL}{L_*}. \quad (6)$$

Robust observations of luminosity functions with measures of  $\phi_*$ ,  $L_*$  and  $\alpha$  exist out to  $z \sim 8$  (Bouwens et al. 2011; McLure et al. 2013; Schenker et al. 2013b) and current observations can provide constraints on the normalization (with other parameters fixed) out to  $z \sim 10$  (Oesch et al. 2013).

We parametrize the evolution of the luminosity function with redshift by fitting quoted observational results for  $\log \phi_*$ ,  $L_*$  and  $\alpha$  and fitting them linearly as a function of  $z$  from  $z = 4-8$ . Fig. 5 shows the fit used in this work in comparison with fits from other authors. The data points used for this fit (plotted) are taken from Bouwens et al. (2007) for  $z = 4-6$  and from McLure et al. (2013) for  $z = 7-8$ . Points at higher  $z$  (which assumed fixed values for  $\alpha$  and  $\phi_*$ ) are shown for reference from Oesch et al. (2013). Note that formally the luminosity density becomes divergent if  $\alpha < -2$ ; however, due to the introduction of a minimum cutoff scale in halo masses in equation (12), this is not a cause for concern. It is important to stress here that even small changes in the fit parameters provides drastic changes in the reionization history. This is especially true for changes to the faint end slope  $\alpha$  of the luminosity function (Bouwens et al. 2011). Future observations from the *JWST* can hopefully much better constrain  $\alpha$  at high redshift.

#### 3.2 Connecting haloes to galaxies

In assigning luminosities to dark matter haloes we assume that brighter galaxies reside in more massive haloes and that the rela-



**Figure 5.** Shown are the fits for the Schechter function parameters  $M_*$ ,  $\alpha$  and  $\phi_*$ . The black symbols show observed values with quoted uncertainties (Bouwens et al. 2007; McLure et al. 2013; Oesch et al. 2013). The lack of uncertainties for  $\phi_*$  and  $\alpha$  at  $z = 9$  and  $10$  is due to the fact that they were fixed in the estimation of  $M_*$ . Since there is a large correlation between  $\alpha$  and  $M_*$ , these observations were not used for the fit. See Oesch et al. (2013) for further discussion. The solid black line is the fit used in this work, and we show other fits from Bouwens et al. (2011) and Kuhlen & Faucher-Giguere (2012) for comparison. Note that our fit uses newer observations from the 2012 HUDF results than the comparison fits. For reference, our fits are  $\log \phi_*(z) = -2.6 - 0.074z$ ,  $M_*(z) = -22.1 + 0.29z$  and  $\alpha(z) = -1.4 - 0.064z$ .

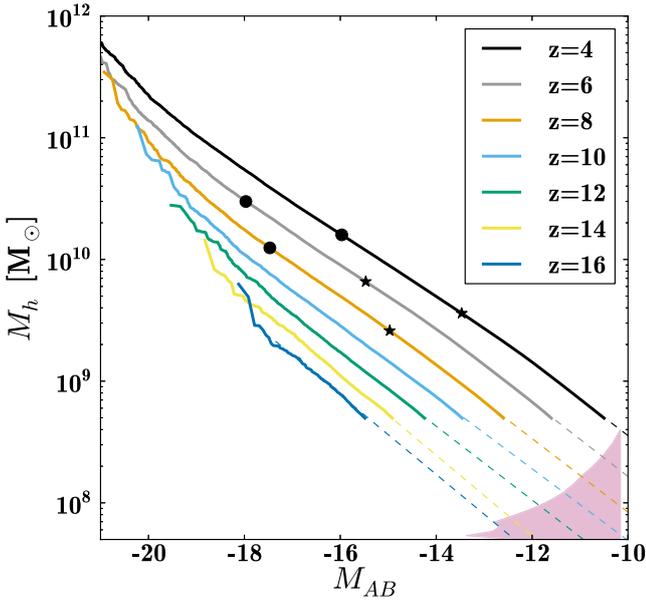
tionship between halo mass and galaxy luminosity is monotonic, following the same relation for all dark matter models.

Fundamentally, we rely on the abundance-matching technique (Kraivtsov et al. 2004; Vale & Ostriker 2004), which defines the relationship between halo mass and galaxy luminosity (or alternatively stellar mass) by equating the cumulative number density of haloes to the cumulative number density of galaxies observed. The power of this approach is that the observed luminosity function is fully reproduced (at least down to luminosities where the observations are complete or to where the matching is performed) while sweeping all uncertainties galaxy formation physics under the rug. In principle, other halo parameters could be used as the rank order of choice (e.g. maximal circular velocity; Trujillo-Gomez et al. 2011). Recently, Behroozi and collaborators have argued that halo mass is the most robust variable to use for these purposes (Behroozi et al. 2014).

Specifically, we set the relationship between halo mass  $M_h$  and UV luminosity  $L$  at different redshifts  $z$  via

$$n_{\text{CDM}}(M_h, z) = \Phi_{\text{gal}}(L, z), \quad (7)$$

where  $n_{\text{CDM}}$  is the cumulative dark halo count in CDM and  $\Phi_{\text{gal}}$  is the cumulative luminosity function as given by the Schechter function fits discussed in Section 3.1. The resultant relationships at various redshifts are plotted in Fig. 6. As can be seen, we are fundamentally assuming that the relationship between halo mass and galaxy luminosity obeys a power law at faint magnitudes (normalized at the bright end by observations) with  $M_h \propto L_{AB}^a$ ,  $a \approx 0.75$  (or  $\log M_h \sim -0.3M_{AB}$ ).



**Figure 6.** Shown is the relationship between halo mass and galaxy magnitude adopted in this work. The solid lines are derived from our simulations using abundance matching as described in the text. They are truncated at the point where halo catalogues become incomplete. The dashed lines are power-law extrapolations, fit to the solid lines at the faint end as  $\log M_h = C + b M_{AB}$ , with resulting slopes of  $b = -0.35$  to  $-0.30$ . The shaded area is an estimate of the cooling limit for haloes, below which galaxies cannot form efficiently, via equation (12). The circles indicate the current HUDF magnitude limit, the asterisks are the expected *JWST* limits.

Our fundamental assumption is that the halo mass–luminosity mapping is a power law at faint magnitudes. Furthermore we are conservatively assuming that the relationship between halo mass and galaxy luminosity is the same in CDM and WDM (obeying near power-law behaviour at faint luminosities). This approach demands that all models match the observations at the bright end (where halo counts overlap), and makes the assumption that there is no special break (towards more efficient galaxy formation) in the luminosity–halo mass relation in WDM for small haloes. This is conservative because there is no reason to expect that WDM haloes will be more efficient at making galaxies than CDM haloes. Indeed, WDM haloes collapse later and have had less time to form stars, so we might expect them to be less efficient at forming stars than their CDM counterparts.

Herpich et al. (2014) performed hydrodynamical simulations of several different WDM cosmologies. They found only slight differences between the stellar masses of the different dark matter models they considered, and the difference they did see was towards less efficient formation in WDM as discussed above. For the most extreme case, comparing a 1 keV WDM model to CDM, they found a ratio of  $M_{*,\text{CDM}}/M_{*,\text{WDM}} \sim 2$  in stellar masses at  $z = 0$ . Therefore, the main differences in the star formation histories are produced at late times, and therefore, this relatively small effect is more reduced at the high redshifts of interest for this work. Small variations in the stellar feedback implementation have a much greater impact on the final stellar mass of a galaxy than WDM particle mass. While the results of Herpich et al. (2014) are based on low- $z$  simulations it seems reasonable to expect the same to be true at higher redshift. Appendix A further elaborates on using the CDM halo catalogue as the fiducial model.

The assumption that the power-law behaviour for the faint end of the halo mass–luminosity mapping translates into a suppression of the luminosity function in the corresponding WDM model. This suppression is then exactly equal to the ratio of the halo catalogues at the given mass scale, which is an aggressively conservative assumption of a cancellation of two disparate physical mechanisms. Furthermore, this requires that the WDM luminosity functions diverge from a pure Schechter fit. We stress that several other processes could also cause the luminosity function to diverge from a power law; however, since no such divergence is yet observed, this translates into constraints on the nature of dark matter.

Before moving on, we note that while the general abundance-matching approach has proven successful and robust at reproducing galaxy properties in the low-redshift universe, it is less well tested at higher redshifts. For example, the scatter at fixed halo mass appears to be fairly minimal at low- $z$  (Behroozi et al. 2014) and the relationship between halo mass and luminosity is well described by a power law for faint systems (Moster, Naab & White 2013; Garrison-Kimmel et al. 2014). At high redshift however, the relationship between halo mass and UV luminosity could in principle exhibit significant scatter, though a power-law relationship for the smallest galaxies appears consistent with the data (Behroozi et al. 2014). We adopt a strict one-to-one relationship between halo mass and UV luminosity as a starting point in investigating the expected differences between CDM and WDM on galaxy counts in the high- $z$  universe. Because we are looking at differential effects between the two models, driven by the declining number of low-mass haloes in WDM, we anticipate that this approach provides a fair starting point, though it would be useful to extend this approach to more complicated mappings in the future.<sup>2</sup>

### 3.3 Reionization

With predictions for luminosity functions in hand, we can directly connect those to expectations on reionization. Star forming galaxies at  $z \gtrsim 6$  are the primary candidate for the main process driving the reionization of the IGM. Changes in the abundance of early galaxies therefore translates into different reionization histories.

The reionization process is a tug-of-war between ionizing radiation from short-lived massive stars and atomic recombination in the IGM. In terms of the volume filling fraction of ionized hydrogen  $Q_{\text{HII}}$  this is captured in the differential equation (Kuhlen & Faucher-Giguere 2012)

$$\frac{dQ_{\text{HII}}}{dt} = \frac{\dot{n}_{\text{ion}}}{\langle n_{\text{H}} \rangle} - \frac{Q_{\text{HII}}}{\bar{t}_{\text{rec}}}, \quad (8)$$

where  $\dot{n}_{\text{ion}}$  is the creation rate of ionizing photons, and  $\langle n_{\text{H}} \rangle$  is the comoving density of baryons

$$\langle n_{\text{H}} \rangle = X_{\text{p}} \Omega_{\text{b}} \rho_{\text{crit}}, \quad (9)$$

and  $\bar{t}_{\text{rec}}$  the mean time of H II recombination

$$\bar{t}_{\text{rec}} = \frac{1}{C_{\text{HII}} \alpha_{\text{B}}(T_0) \langle n_{\text{H}} \rangle (1 + Y/4X)(1 + z^3)}. \quad (10)$$

Here,  $\alpha_{\text{B}}$  is the case B recombination coefficient,<sup>3</sup>  $T_0$  is the IGM temperature and  $X$  and  $Y = 1 - X$  are the primordial

<sup>2</sup> Notably, we take the inherent scatter in the halo mass–luminosity mapping into consideration.

<sup>3</sup> Commonly used cases A and B definitions differentiate mediums that allow the Lyman photons to escape or that are opaque to these lines (except Ly $\alpha$ ) respectively. Case B is most appropriate for this reionization calculation.

hydrogen and helium abundances, respectively. Since recombination is not isothermal and uniformly distributed, the gas clumping factor  $C_{\text{HII}} = \langle n_{\text{H}}^2 \rangle / \langle n_{\text{H}} \rangle^2$  is also introduced to quantify the effects these approximations have. Allowing a fraction of  $f_{\text{esc}}$  of the produced ionizing photons to escape the gas clouds where the massive stars are born, the injection of UV photons into the IGM is given by the differential luminosity function  $\phi$  down to a limiting luminosity  $L_{\text{lim}}$

$$\dot{n}_{\text{ion}} = \zeta_{\text{ion}} f_{\text{esc}} \int_{L_{\text{lim}}}^{\infty} L \phi(L) dL = f_{\text{esc}} \zeta_{\text{ion}} \rho_{\text{UV}} \Big|_{L_{\text{lim}}}. \quad (11)$$

Here,  $\zeta_{\text{ion}}$  is a parameter converting the galactic UV luminosity to ionizing photon luminosity, or more precisely the amount of Lyman continuum photon emission per 1500 Å unit UV luminosity density. Note that  $f_{\text{esc}}$  and  $\zeta_{\text{ion}}$  are completely degenerate parameters, as reionization is only sensitive to the product. Any change in one of these parameters could be attributed to the same relative change in the other.

Critical in this analysis is what value to assign to the limiting luminosity in equation (11), that is, the minimal UV luminosity expected possible from early galaxies. Naïvely, one might expect no such lower limit. However, to capture the hot primordial gas needed for star formation, a sufficiently deep potential well is required. This effectively puts a lower bound on the possible UV luminosities (and therefore halo masses) due to photoevaporation. Furthermore, star formation can only take place once the hot gas has cooled sufficiently, and this introduces a limiting mass threshold below which stars cannot form. In this work we only consider the cooling limit for halo masses, and we use the parametrization adopted by Sobacchi & Mesinger (2013)

$$M_{\text{cool}} = 10^8 \left( \frac{1+z}{10} \right)^{-3/2} M_{\odot}. \quad (12)$$

The shaded red area in Fig. 6 shows an estimate of the region where galaxy formation is suppressed owing to this cooling limit, effectively mapping the halo mass limit to a luminosity cutoff. Based on this, we will explore cutoff magnitudes between  $M_{AB} = -10$  and  $M_{AB} = -13$  in what follows.

An important constraint on reionization comes from the Thomson optical depth to the CMB,

$$\tau_e = \int_0^{z_{\text{R}}} \frac{c(1+z)^2}{H(z)} Q_{\text{HII}} \sigma_{\text{T}} (n_{\text{H}}) (1 + \eta(z) Y / 4X). \quad (13)$$

Here,  $z_{\text{R}}$  is the redshift of recombination,  $\sigma_{\text{T}}$  is the Thomson cross-section and  $\eta = 1$  when helium is singly ionized and  $\eta = 2$  when helium is doubly ionized after  $z \lesssim 4$ .

In this work, we do not include an evolution in the reionization parameters. Kuhlen & Faucher-Giguere (2012) found that for their best-fitting scenario, evolution in the limiting luminosity alone is not enough to match both Ly $\alpha$  constraints and reionization constraints, and the data provide no conclusive evidence for an evolution in any case. Evolution in  $f_{\text{esc}} \zeta_{\text{ion}}$  (resulting perhaps from evolution in the stellar initial mass function) may be more plausible. We will present our results with different values of  $f_{\text{esc}} \zeta_{\text{ion}}$  and limiting luminosity, but they will remain fixed with redshift. Evolution in the clumping factor  $C_{\text{HII}}$  may be expected, but no definitive determination of its evolution exists. For example, Finlator et al. (2012) presents a detailed analysis of the evolution, suggesting that the clumping factor rises from  $C_{\text{HII}} < 1$  for  $z > 10$  to  $C_{\text{HII}} \sim 3.3$  at  $z \sim 6$ . It should be noted that our model equation (8) does not include the detailed distribution of hydrogen where some dense clumps reionize later than less dense clumps. This simplification is likely inaccurate in

the final phase of reionization around  $z \sim 6$ ; however, we expect it to be an appropriate approximation on average for higher- $z$  and in the large cosmological volumes of interest here.

In what follows, we adopt the reionization parameters  $C_{\text{HII}} = 3$ ,  $\zeta_{\text{ion}} = 10^{25.3} \text{ erg s}^{-1} \text{ Hz}$ ,  $X_{\text{p}} = 0.75$ ,  $T_0 = 2 \times 10^4 \text{ K}$  and  $\alpha_B = 1.6 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1}$ . The escape fraction and limiting luminosities vary, and will be indicated in the relevant figures.

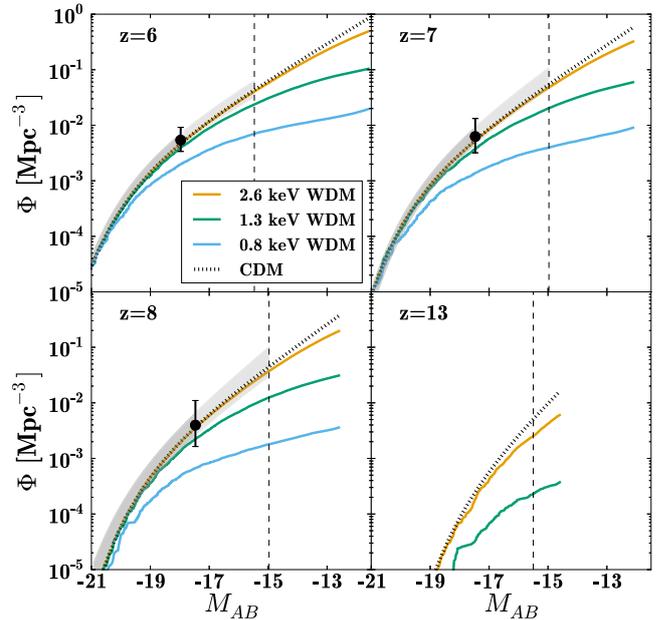
## 4 RESULTS

Here, we explore the constraints on WDM models by direct number counting and the inferred reionization history, and illustrate how future galaxy count surveys can improve these constraints on WDM models.

### 4.1 WDM constraints from galaxy counts

As explained in Section 3.2, given our (conservative) assumption of the star formation efficiency being the same in the different dark matter models, deviations in the implied WDM luminosity function from the observed Schechter fit translates into constraints on the WDM model.

Fig. 7 shows the implied luminosity functions for CDM (dotted black) and each of our WDM models (solid, colours indicated). The symbol with error bar is the known (observed) cumulative count of galaxies at the faint HUDF limit, with errors indicative of the  $2\sigma$  uncertainty calculated as in Fig. 4. The shaded band corresponds to a  $1\sigma$  uncertainty of the observed best-fitting Schechter function



**Figure 7.** Shown are cumulative luminosity functions for our CDM and WDM models at various redshifts (comoving volumes). In each panel, the symbol with error bar marks the observed cumulative count at the limit of published HUDF luminosity functions with  $2\sigma$  uncertainties shown. The shaded bands correspond to  $1\sigma$  uncertainties, and the vertical line marks the approximate reach of a hypothetical deep field observation with *JWST*. For redshifts where the luminosity function has been observed by HUDF, the *JWST* limit has been assumed to be 2.5 mag fainter. For  $z = 13$ , a limit of  $-15.5$  has been assumed (see Windhorst et al. 2006). The 0.8 keV WDM model is heavily disfavoured by current observations, and the 1.3 keV model is marginal. At redshift  $z = 13$ , *JWST* observations will likely be able to rule out 1.3 keV WDM and perhaps be sensitive to 2.6 keV.

(seen in Fig. 5), extrapolated down to an approximate *JWST* deep field limit (indicated by the vertical line). We see clearly here that the 0.8 keV WDM model (solid cyan) is strongly disfavoured by current observations of the galaxy luminosity functions. The 1.3 keV WDM model, while currently consistent with observations, demonstrates significant deviations from CDM at  $z = 13$  at magnitudes observable with *JWST*. Deep galaxy counts at this and earlier epochs may be even sensitive to 2.6 keV WDM.

We quantify how much the different models are disfavoured with a  $\chi^2$  test,

$$\chi^2 = \sum_i \left( \frac{\Phi - \Phi_{\text{obs}}}{\sigma} \right)^2. \quad (14)$$

Here,  $\Phi_{\text{obs}}$  is the abundance at the faint-end limit from observed luminosity functions (Bouwens et al. 2007; McLure et al. 2013; Oesch et al. 2013) and  $\sigma$  is the error on the simulated luminosity function, which is given by the jackknife error on the halo mass function at the corresponding abundance.<sup>4</sup> The CDM luminosity function is a fit to the redshift evolution of the Schechter function parameters (shown in Fig. 5) based on current observations, and not the actual quoted fits at each redshift, and this produces a small but non-zero  $\chi^2 = 2$  for the CDM case with 5 degrees of freedom (corresponding to the observations from  $z = 4$  to  $z = 8$ ) from the luminosity functions (85 per cent consistency). The  $\chi^2$  for the 2.6, 1.3 and 0.8 keV models are 2.27, 14.4 and 372, respectively, with probabilities for these models at getting the observed luminosity functions of 81 per cent, 1.3 per cent and  $\ll 10^{-10}$ . Therefore, the 1.3 keV WDM model is disfavoured at approximately 98.6 per cent C.L. ( $2.2\sigma$ ), and the 0.8 keV WDM model is disfavoured at very high significance,  $>10\sigma$ . Note that the statistical methods in the modelling and constraints here do not reflect systematic uncertainties in the halo abundance-matching method. Because we arrive at a consistent and smooth power-law relation for the abundance matching, the systematic effects are likely small, though difficult to quantify.

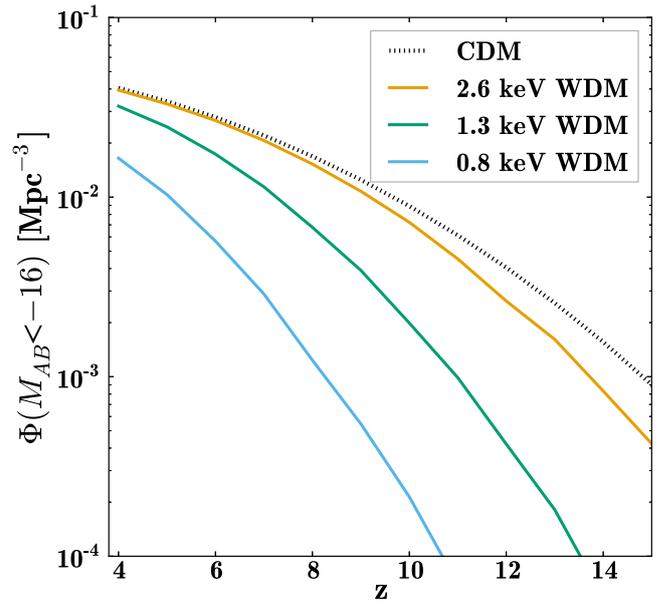
Had we considered WDM artificial haloes and the fact that the star formation efficiency is slightly lower in WDM than in CDM, the constraints would improve slightly, since there would then be fewer and less luminous galaxies. However, these effects are not likely to be very important at redshifts  $z \geq 4$ , and are a priori unknown.

Faint galaxy counts at even higher redshift will be particularly sensitive to WDM models. We demonstrate this in Fig. 8. Here, we show the cumulative number density of galaxies brighter than  $M_{AB} = -16$  as a function of redshift for each of our models. The differences between CDM and WDM are significant, especially for the lower mass WDM cases. Deep *JWST* observations should be sensitive to galaxy detections at least this faint out to  $z = 15$ , and therefore will provide a direct probe of the small-scale power spectrum by counting galaxies.

## 4.2 WDM constraints from reionization

Fundamental to the ability of galaxies to reionize the universe is the production rate of ionizing photons,  $\dot{n}_{\text{ion}}$ , which is proportional to the total UV luminosity density,  $\rho_{\text{UV}}$ , coming from these sources (see equation 11). For an underlying galaxy luminosity function with a steep faint-end slope  $\alpha$ , the total luminosity density implied will

<sup>4</sup> We use the jackknife errors instead of the errors on the observed luminosity function, since they are larger (shown in Fig. 7). The  $\chi^2$  test here is approximate, but our conclusions are robust regarding the models' consistency.



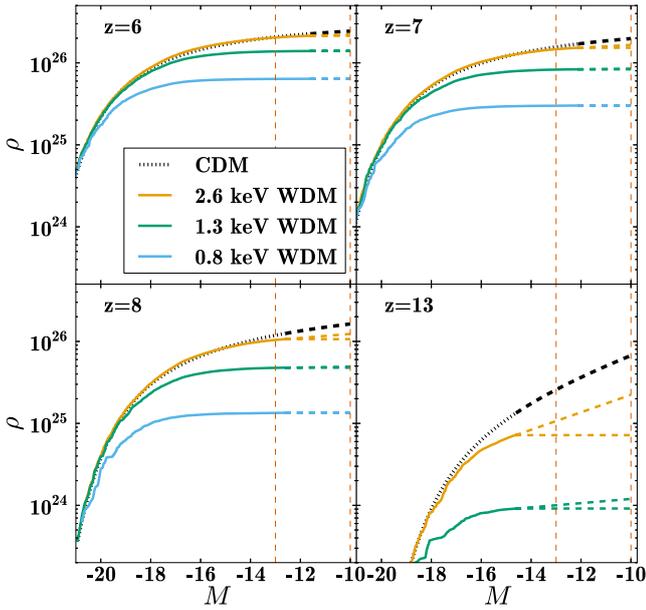
**Figure 8.** Predicted number density of galaxies brighter than  $M_{AB} = -16$  as a function of redshift for our CDM and WDM models. *JWST* should be capable of detecting galaxies of this brightness across the redshift range plotted, and perhaps be sensitive to differences between 2.6 keV WDM and CDM at  $z > 12$ .

be sensitive to the assumed faint-end cutoff used to calculate  $\rho_{\text{UV}}$ . Fig. 9 shows the luminosity density in our models as a function of faint-end cutoff at selected redshifts. Because WDM models have flatter faint-end luminosity function slopes, the total  $\rho_{\text{UV}}$  is less sensitive to the faint-end cutoff, i.e. the implied cumulative  $\rho_{\text{UV}}$  values flatten relative to CDM at fainter magnitudes. Importantly for our considerations, WDM predictions for reionization will be less sensitive to the adopted faint-end cutoff than CDM, owing to the lack of small galaxies in these models.

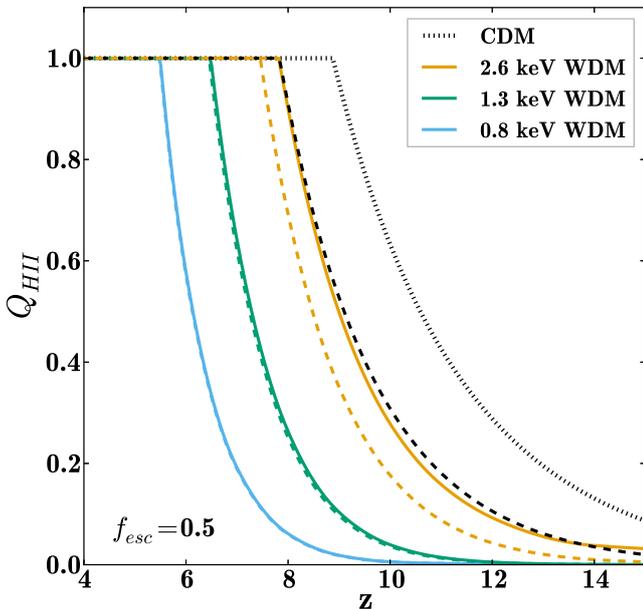
The points where the lines change to dashed in Fig. 9 mark the resolution limit in the simulations. The two WDM dashed lines bracket the following extreme cases: one, a power-law fit to the faint end, and, two, the constant value at the faintest point resolved in the simulations. The actual luminosity density would be somewhere between these two extremes, though for the 0.8 and 1.3 keV models the difference is negligible. All our analysis utilizes the power law extrapolations to get conservative estimates.

With the luminosity density in hand, the reionization history can now be determined by virtue of equation (8). Fig. 10 presents the volume filling fraction  $Q_{\text{H II}}$  as a function of redshift for two choices of limiting magnitude in calculating the luminosity density. The fiducial line types (shown in the legend) correspond to a limiting magnitude of  $M_{AB} = -10$  while dashed lines cutoff at a brighter limit of  $M_{AB} = -13$ . We used the initial condition  $Q_{\text{H II}} = 0$  at  $z = 20$  and integrated forward in time. We choose an optimistic escape fraction of  $f_{\text{esc}} = 0.5$ , higher than assumed in both Robertson et al. (2013) and Kuhlen & Faucher-Giguere (2012), and therefore more conservative with respect to WDM model constraints since we use the same  $\zeta_{\text{ion}}$  as in Kuhlen & Faucher-Giguere (2012). Robertson et al. (2013) uses a lower  $\zeta_{\text{ion}}$ , but instead assumes an almost constant luminosity density at the cutoff scale at high redshift.

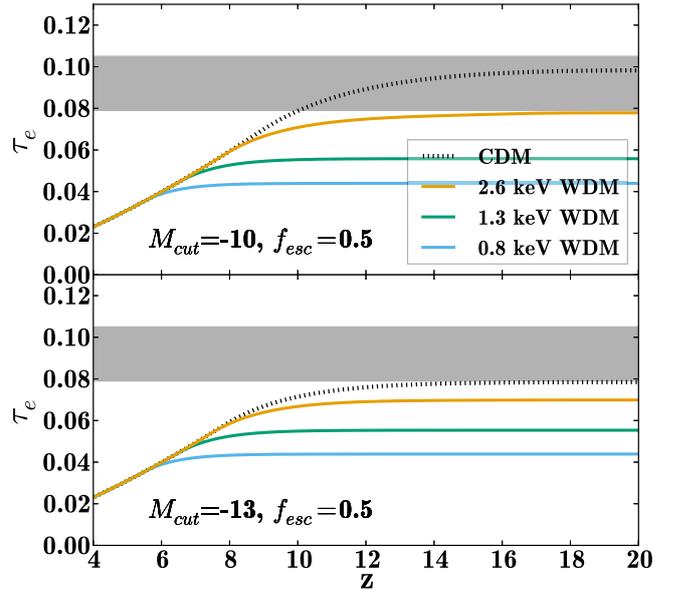
All of the models in Fig. 10 except 0.8 keV WDM have completed reionization by  $z \sim 5.8$  as required by results inferred from the kinematic Sunyaev-Zel'dovich effect and CMB polarization observations (Zahn et al. 2012). For the 0.8 keV model shown,



**Figure 9.** The cumulative UV luminosity density (in units  $\text{erg s}^{-1} \text{Hz}^{-1} \text{Mpc}^{-3}$ ) as a function of magnitude cutoff at selected redshifts for the dark matter models considered. The vertical lines mark the two cutoff scales we consider in this paper as plausible for extending the galaxy luminosity function. The change from solid to dashed lines occurs at the magnitude corresponding to the resolution limit of the simulation, beyond which we rely on extrapolations (dashed) to predict faint galaxy contributions. For a given WDM model, the upper dashed lines extend the best-fitting power law of the resolved function. The lower dashed line marks the constant value at the faintest simulated point, as would be expected if the WDM halo mass function drops dramatically beyond this point. These two extremes bracket reasonable expectations.



**Figure 10.** Build-up in volume filling fraction of ionized hydrogen as a function of redshift for our CDM and WDM models assuming  $f_{\text{esc}} = 0.5$  and with limiting integration magnitudes of  $M_{AB} = -10$  (fiducial lines, as in caption) and  $M_{AB} = -13$  (dashed). The WDM models produce more rapid build-ups at later times compared to CDM. The 0.8 keV WDM model does not complete reionization by  $z \sim 6$  (as seems to be required by observations) regardless of the assumed faint-end cutoff.

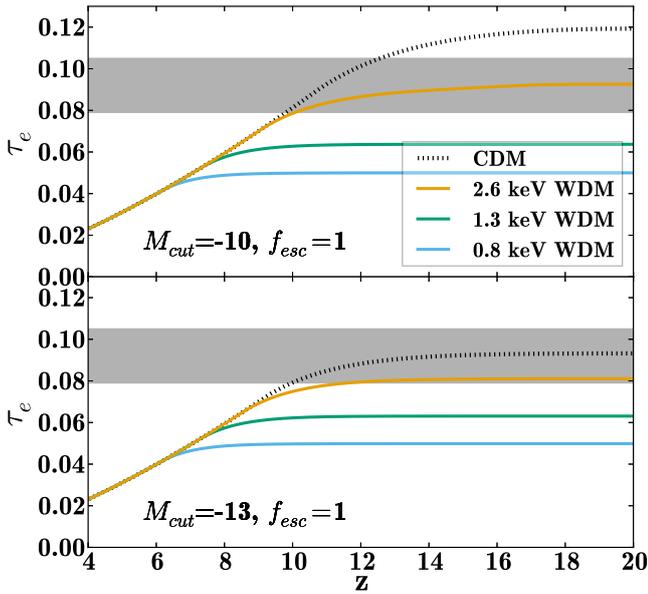


**Figure 11.** The electron-scattering optical depth of the CMB predicted for our CDM and WDM models, with the contribution shown cumulatively as a function of redshift. The top panel shows results for an assumed luminosity function cutoff at  $M_{AB} = -13$  and the bottom panel extends this cut to very faint luminosities  $M_{AB} = -10$ . In all case, we assume  $f_{\text{esc}} = 0.5$  and our fiducial value of  $\zeta_{\text{ion}}$ . The bands are the 68 per cent confidence limit on the most recent *Planck* results Ade et al. (2013). Note that none of the WDM models reach within the 68 per cent confidence band from *Planck*.

reionization is complete at  $z = 5.5$ . In general, the WDM cases produce a more rapid late-time build-up of ionized hydrogen due to the high-redshift suppression of haloes. It can also be seen that the difference between CDM and 2.6 keV WDM is larger when the fainter limiting magnitude is used, simply because the difference between the models is much larger here. Of course, these results are sensitive to the escape fraction. For example, if an escape fraction of  $f_{\text{esc}} = 0.2$  is used for the 1.3 keV model full reionization is not reached until  $z = 5.4$  (not shown on figure), so fairly high escape fractions seem to be required for 1.3 keV to reach full reionization by  $z \sim 6$ .

Another important probe of reionization is the integrated optical depth of electron scattering from the CMB. The shaded bands in Fig. 11 show the CMB optical depth range of  $0.092 \pm 0.013$  from the most recent *Planck* results (Ade et al. 2013). The lines show predictions for the optical depth as contributed as a function of redshift for our WDM and CDM models assuming an escape fraction  $f_{\text{esc}} = 0.5$ , with the two panels corresponding to different limiting magnitudes. Interestingly, with this choice of (fairly high) escape fraction, none of our WDM models can reproduce the measured optical depth, and even CDM requires a luminosity function extrapolation to a very faint limiting magnitude. This is consistent with the findings of Robertson et al. (2013).

In Fig. 12, we show results for the optical depth, now assuming  $f_{\text{esc}} = 1$ . In this case, the 2.6 keV model can reproduce the *Planck* value, though a fairly faint limiting magnitude seems to be required, even in this extreme case. Unsurprisingly, CDM severely overshoots the optical depth with these (rather high) reionization parameters. It is noteworthy that neither of the low-mass WDM models can reproduce the *Planck* optical depth within its 68 per cent confidence interval, even with very optimistic choices. If these WDM models are to be viable in the face of reionization constraints, they would require either significant contribution to the ionizing flux from



**Figure 12.** Electron-scattering optical depth as in Fig. 11 except now assuming  $f_{\text{esc}} = 1.0$ . Even with fairly extreme assumptions, neither of the two lightest WDM models are able to reach the 68 per cent confidence range (bands) reported by *Planck*.

non-stellar sources, or a significantly larger combination of  $f_{\text{esc}}\zeta_{\text{ion}}$  than what is currently believed to be realistic. A smaller  $C_{\text{HI}}$  could also help in this respect. Future observations will likely better constrain these parameters.

## 5 DISCUSSION

Future measurements of the luminosity function of faint galaxies at high- $z$ , particularly those from *JWST*, and possibly with *HST* via the Frontier Fields initiative, will significantly improve the sensitivity to WDM models and the halo mass cutoff effects presented here. For a direct number count comparison, the mass resolution of our simulations is sufficient to connect with observations down to plausible detection limits with *JWST*.

In contrast, our results on reionization specifically for the 2.6 keV model would improve with greater mass resolution. This is because of our conservatively approximated increasing faint end luminosity density function, which likely flattens at low luminosities in this model, but remains unresolved in our simulations. This improvement can also be made, though to a lesser degree, for the lower WDM particle mass models. Our results for these particle mass models indicate that the luminosity density reaches an approximately constant level at the faint end, and thus the ionizing flux will remain constant at fainter luminosities. Deep observations with *JWST* certainly will much better determine the faint end slope  $\alpha$  of the luminosity function. The reionization history is highly sensitive to this parameter. A future analysis could possibly circumvent the uncertainty stemming from fitting the evolution of the Schechter parameters, instead relying on direct observations of  $\alpha$  for most redshifts.

We have neglected any evolution in the escape fraction with redshift. Naïvely, one might expect the escape fraction to decrease with redshift since the overall density scales as  $(1+z)^3$ . However, observations seem to indicate the opposite: the escape fraction increases with redshift. Mitra, Ferrara & Choudhury (2013) found that  $f_{\text{esc}} \sim 0.06$  at  $z \sim 6$ , and increases to at least  $f_{\text{esc}} \gtrsim 0.146$  by  $z \sim 10$ .

This could be caused by an initial mass function for star formation favouring high mass stars in the early universe. Alternatively, Ferrara & Loeb (2012) proposes a mechanism where mini haloes close to the cooling limit contributes appreciably to the ionizing flux, but their contribution diminishes over time due to feedback mechanisms. Since these small haloes have a relatively larger escape fraction the overall escape fraction decreases with time. This mechanism is especially interesting from a WDM perspective, since the lower abundance of small haloes directly counteracts this. The evolution of the escape fraction remains uncertain, and a typical constant value in previous work was  $f_{\text{esc}} \sim 20$  per cent, and our conservative choice of  $f_{\text{esc}} > 0.5$  will very unlikely overestimate the ionizing flux.

Non-stellar processes can potentially contribute to the reionization history. Quasars might play an important role at high redshift, although current results seem to indicate the contribution from quasars is sub-dominant (Volonteri & Gnedin 2009). For example, Willott et al. (2010) found that at  $z \sim 6$  the ionizing flux from quasars is 20–100 times lower than what is needed for continued reionization. X-ray emission by black holes may also contribute appreciably. Ricotti & Ostriker (2004) and Ricotti, Ostriker & Gnedin (2005) analysed a pre-ionization contribution from a top-heavy initial mass function for the Population III stars. These Population III stars then collapse into black holes and subsequently accrete at nearly the Eddington limit. Accretion on to the black holes could partially reionize the IGM, although the primary effect would be heating the IGM. The Population III phase is rapidly self-limiting due to pollution by heavy elements and pair instability supernovae causing strong outflows, and thus the metal-poor Population III stars needs to collapse into black holes. This pre-ionization phase is then followed by a period of stellar reionization.

Bovill & Ricotti (2011) and Katz & Ricotti (2013) present a mechanism where the bulk of the star formation in the first dwarf galaxies happened in protoglobular clusters that were subsequently tidally stripped from the dwarf galaxy. Since the tidal stripping of globular clusters in the halo outskirts precede the stripping of the dark matter halo this can break the assumption of the abundance-matching technique since the dwarf galaxies would be stripped off much of their luminous matter. Consequently, the mass–luminosity mapping would no longer be strictly monotonic, and the scatter on any mass–luminosity relation would likely increase. Importantly, for mass scales larger than the masses of these dark dwarf galaxies, however, the abundance matching would still give a meaningful average mass–luminosity mapping. In any case, the abundance matching we employ here would be an upper bound on the luminosity of a dwarf galaxy and, therefore, our approach is highly conservative.

In the case of sterile neutrino WDM, it has been shown that the effects of the radiative decay of sterile neutrino WDM to X-ray photons may catalyse the formation of  $\text{H}_2$  and star formation (Biermann & Kusenko 2006). This does not affect the results presented here, because in this case the sterile neutrino WDM cosmology is constrained to produce the same observed high- $z$  luminosity functions from which the reionization history is inferred. The only method by which such radiative decays would enhance the reionization rate is if they preferentially enhanced star formation for small mass haloes below the luminosity function cutoff magnitude. We are aware of no mechanism in the literature that would produce such an enhancement for low-mass haloes. Moreover, this X-ray photon catalyzed process would enhance star formation preferentially in metal free low-mass haloes that rely on  $\text{H}_2$  to cool. More massive haloes which are involved in reionization are insensitive to this mechanism as they can cool by  $\text{Ly}\alpha$  line emission and are more likely

to be metal rich. Therefore, since halo formation is suppressed at these low masses, we believe our results apply for the case of sterile neutrino WDM, with the 3 and 6 keV mass scales disfavoured at  $>10\sigma$  and 98.6 per cent C.L., respectively.

Recall that the 1.3 keV (thermal; 6 keV sterile) model we have considered corresponds to model discussed by Lovell et al. (2012) as a solution to the too-big-to-fail problem (Boylan-Kolchin et al. 2012) and the M2L25 model studied in Boyarsky et al. (2009). We have demonstrated that this model is disfavoured at 98.6 per cent C.L. by direct galaxy counts at high redshift and is unable to reproduce the CMB optical depth even with extreme assumptions about the escape fraction.

## 6 CONCLUSION

We have shown that the Lyman-break technique for galaxy surveys at high redshift can provide a direct method for constraining the nature of dark matter and its clustering at small scales, with sensitivity to the structure formation suppression present in WDM models. We have analysed CDM and WDM cosmological simulations in order to test WDM models using the luminosity function observations at high- $z$  as well as a new analysis of cosmological reionization limits. Given the assumptions that the luminosity function of a  $\Lambda$ CDM universe is modelled by a Schechter function down to faint magnitudes and that the mass–luminosity relation of galaxies is independent of the dark matter model employed, we have modelled the luminosity function for several dark matter models to analyse the sensitivities to WDM dark matter models.

Using an approximate  $\chi^2$  test of the faint end of the luminosity function, direct number counts of galaxies significantly disfavours a 0.8 keV WDM model at greater than  $10\sigma$ , and a 1.3 keV model is disfavoured at approximately 98.6 per cent C.L. ( $2.2\sigma$ ). Further, with highly optimistic values for the parameters that translate high-redshift galaxy luminosity to ionizing flux, the 0.8 and 1.3 keV models are inconsistent with the CMB optical depth at greater than 68 per cent C.L. Furthermore, for the conservative case of a limiting luminosity of  $M_{AB} = -13$ , a 2.6 keV WDM model is only marginally consistent with the 68 per cent confidence region of the optical depth from *Planck*. Wherever possible, we have used conservative values on parameters, making WDM behave more like CDM. For this reason, we feel confident concluding that neither the 0.8 or 1.3 keV models are consistent at more than 68 per cent C.L. with reionization, even with the large uncertainty on the reionization process.

We expect upcoming deep surveys with *JWST* (and possibly *HST* via the Frontier Fields) to be able to reach luminosities and redshifts that can fully discern between a CDM model and a 1.3 keV model by direct number counts. Even 2.6 keV WDM might prove discernible if the observations are deep enough. Additionally, if the constraints on reionization parameters are improved, a 2.6 keV WDM model can be distinguished from CDM by its different reionization history. The study of galaxy formation and reionization in the high- $z$  universe adds a complementary and competitive probe to the nature of dark matter.

## ACKNOWLEDGEMENTS

We thank useful discussions with John Beacom and Richard Ellis. We would also like to thank the referee Massimo Ricotti for helpful feedback on the paper. KNA is partially supported by NSF CAREER Grant no. PHY-11-59224. JO and JSB were supported by grants from the NSF and NASA.

## REFERENCES

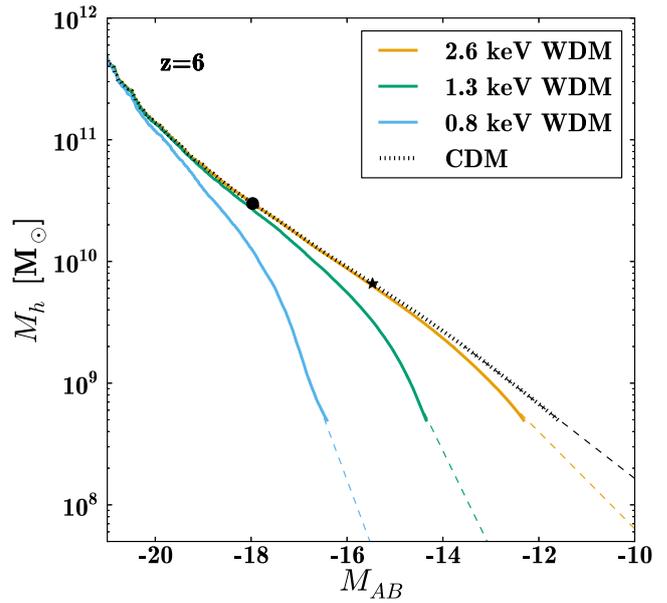
- Abazajian K., 2006, *Phys. Rev. D*, 73, 063506  
 Abazajian K. et al., 2011, *Astropart. Phys.*, 35, 177  
 Ade P. A. R. et al., 2013, preprint (arXiv:1303.5076)  
 Angulo R. E., Hahn O., Abel T., 2013, *MNRAS*, 434, 337  
 Barkana R., Haiman Z., Ostriker J. P., 2001, *Astrophys. J.*, 558, 482  
 Behroozi P. S., Wechsler R. H., Lu Y., Hahn O., Busha M. T., Klypin A., Primack J. R., 2014, *ApJ*, 787, 156  
 Biermann P. L., Kusenko A., 2006, *Phys. Rev. Lett.*, 96, 091301  
 Bode P., Ostriker J. P., Turok N., 2001, *ApJ*, 556, 93  
 Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, *ApJ*, 670, 928  
 Bouwens R. et al., 2011, *ApJ*, 752, L5  
 Bovill M. S., Ricotti M., 2011, *ApJ*, 741, 18  
 Boyarsky A., Lesgourgues J., Ruchayskiy O., Viel M., 2009, *Phys. Rev. Lett.*, 102, 201304  
 Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, *MNRAS*, 422, 1203  
 Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80  
 Calvi V., Pizzella A., Stiavelli M., Morelli L., Corsini E., Dalla Bontà E., Bradley L., Koekemoer A. M., 2013, *MNRAS*, 432, 3474  
 Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, 647, 201  
 de Souza R. S., Mesinger A., Ferrara A., Haiman Z., Perna R., Yoshida N., 2013, *MNRAS*, 432, 3218  
 Dodelson S., Widrow L. M., 1994, *Phys. Rev. Lett.*, 72, 17  
 Ellis R. S. et al., 2013, *ApJ*, 763, L7  
 Fan X. et al., 2006, *AJ*, 132, 117  
 Ferrara A., Loeb A., 2013, *MNRAS*, 431, 2826  
 Finlator K., Oh S. P., Ozel F., Dave R., 2012, preprint (arXiv:1209.2489)  
 Garrison-Kimmel S., Boylan-Kolchin M., Bullock J., Lee K., 2014, *MNRAS*, 438, 2578  
 Hahn O., Abel T., 2011, *MNRAS*, 415, 2101  
 Herpich J., Stinson G. S., Macciò A. V., Brook C., Wadsley J., Couchman H. M. P., Quinn T., 2014, *MNRAS*, 437, 293  
 Horiuchi S., Humphrey P. J., Onorbe J., Abazajian K. N., Kaplinghat M., Garrison-Kimmel S., 2014, *Phys. Rev. D*, 89, 025017  
 Howlett C., Lewis A., Hall A., Challinor A., 2012, *J. Cosmol. Astropart. Phys.*, 4, 27  
 Hu W., Kravtsov A. V., 2003, *ApJ*, 584, 702  
 Jenkins A., 2010, *MNRAS*, 403, 1859  
 Kamionkowski M., Liddle A. R., 2000, *Phys. Rev. Lett.*, 84, 4525  
 Katz H., Ricotti M., 2013, *MNRAS*, 432, 3250  
 Kistler M. D., Yuksel H., Beacom J. F., Hopkins A. M., Wyithe J. S. B., 2009, *ApJ*, 705, L104  
 Knebe A., Wagner C., Knollmann S., Diekershoff T., Krause F., 2009, *ApJ*, 698, 266  
 Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608  
 Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, *ApJ*, 609, 35  
 Kuhlen M., Faucher-Giguere C., 2012, *MNRAS*, 423, 862  
 Kusenko A., 2006, *Phys. Rev. Lett.*, 97, 241301  
 Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473  
 Lovell M. R. et al., 2012, *MNRAS*, 420, 2318  
 Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, *MNRAS*, 439, 300  
 Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, *ApJ*, 671, 1160  
 McLure R. J. et al., 2013, *MNRAS*, 432, 2696  
 Marsh D. J. E., Silk J., 2014, *MNRAS*, 437, 2652  
 Mitra S., Ferrara A., Choudhury T. R., 2013, *MNRAS*, 428, L61  
 Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121  
 Oesch P. A. et al., 2013, *ApJ*, 773, 75  
 Pacucci F., Mesinger A., Haiman Z., 2013, *MNRAS*, 435, L53  
 Polisensky E., Ricotti M., 2011, *Phys. Rev. D*, 83, 043506  
 Polisensky E., Ricotti M., 2014, *MNRAS*, 437, 2922  
 Prunet S., Pichon C., Aubert D., Pogosyan D., Teyssier R., Gottlöber S., 2008, *ApJs*, 178, 179

- Reed D. S., Smith R. E., Potter D., Schneider A., Stadel J., Moore B., 2013, MNRAS, 431, 1866
- Ricotti M., Ostriker J. P., 2004, MNRAS, 352, 547
- Ricotti M., Ostriker J. P., Gnedin N. Y., 2005, MNRAS, 357, 207
- Robertson B. E. et al., 2013, ApJ, 768, 71
- Schneider A., Smith R. E., Reed D., 2013a, MNRAS, 433, 1573
- Schenker M. A. et al., 2013b, ApJ, 768, 196
- Shaposhnikov M., Tkachev I., 2006, Phys. Lett. B, 639, 414
- Shi X.-d., Fuller G. M., 1999, Phys. Rev. Lett., 82, 2832
- Sobacchi E., Mesinger A., 2013, MNRAS, 432, 3340
- Somerville R. S., Bullock J. S., Livio M., 2003, ApJ, 593, 616
- Springel V., 2005, MNRAS, 364, 1105
- Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16
- Vale A., Ostriker J., 2004, MNRAS, 353, 189
- Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534
- Viel M., Becker G., Bolton J., Haehnelt M., 2013, Phys. Rev. D, 88, 043502
- Villaescusa-Navarro F., Dalal N., 2011, J. Cosmol. Astropart. Phys., 1103, 024
- Volonteri M., Gnedin N., 2009, ApJ, 703, 2113
- Wang J., White S. D., 2007, MNRAS, 380, 93
- Willott C. J. et al., 2010, AJ, 139, 906
- Windhorst R. A., Cohen S. H., Jansen R. A., Conselice C., Yan H.-J., 2006, New Astron. Rev., 50, 113
- Zahn O. et al., 2012, ApJ, 756, 65
- Zentner A. R., Bullock J. S., 2002, Phys. Rev. D, 66, 043003

## APPENDIX A: ABUNDANCE MATCHING WITH WDM

Herpich et al. (2014) found that the star formation in low- $z$  Milky Way-like galaxies is slightly suppressed in WDM cosmologies. This seems to fit well with what one would expect: the small-scale cutoff in the WDM transfer function postpones the formation of dwarf galaxy haloes, and therefore the potential wells that act as seeds for the first galaxies are shallower. The effect is relatively small, only a factor of 2 for their most extreme WDM model at  $z = 0$ .

The star formation efficiency for different dark matter models can readily be inferred from abundance matching. Fig. A1 shows the halo mass–luminosity relation by using the different dark matter halo catalogues. For a fixed mass, a WDM halo is seen to be more luminous than a CDM halo, and thus WDM would, unsurprisingly, need to have an enhanced star formation efficiency relative to CDM



**Figure A1.** Abundance matching utilizing the different halo catalogues. Abundance matching with the CDM catalogue gives a power law down to faint magnitudes. The dashed lines are power-law extrapolations to the faint end. Clearly, the faint end in the WDM models diverge from the CDM power law towards more efficient star formation: lower mass haloes have a larger luminosity relative to CDM. However, the WDM models must have roughly the same or a slightly lower star formation rate than CDM, hence our assumption of the CDM power-law behaviour is a very conservative estimate. The circle indicates the current HUDF magnitude limit, the asterisk is the expected *JWST* limits.

in order to match observations. This is counterintuitive, and more importantly contradicts the low- $z$  results of Herpich et al. (2014). A realistic WDM halo mass–luminosity mapping would give a slightly lower star formation efficiency: that is, a flatter slope than CDM in Fig. A1 instead of a steeper slope. In our analysis, we therefore conservatively assume that the halo mass–luminosity is a power law. Therefore, a halo mass can uniquely be mapped to the same luminosity independent of WDM model.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.

### 5.3 The future of warm dark matter

In light of the myriad of complementary methods utilised to constrain the WDM particle mass, it is tempting to conclude that WDM models do not work. Schneider et al. (2013a) shows that a WDM model with a mass of  $m_{\text{thermal}} = 4$  keV does not do significantly better than a CDM model in solving the small scale crisis. Nevertheless, alternatives to the thermal or thermal-like (Dodelson-Widrow sterile neutrino) WDM particle could potentially be viable due to their significantly different transfer function. A two-component cold/warm dark matter scheme or a more gradual power cut-off could also potentially solve the small scale crisis in standard  $\Lambda$ CDM. In any case, due to the largely unknown impact of baryonic feedback, state of the art hydrodynamical simulations should be a priority, especially in the light of Occam's Razor. It should be noted, however, that the too-big-to-fail problem is not likely to be solved by supernovae feedback as it would require extraordinarily frequent supernovae (Garrison-Kimmel et al., 2013), but it could possibly be solved by using more recent values of normalisation of the matter power spectrum  $\sigma_8$  and the primordial spectral index  $n_s$  (Polisensky and Ricotti, 2014). There is thus still no clear indication of what might be the solution to the CDM small scale crisis, and the full picture could possibly involve non-thermal WDM models. Explaining the small scale crisis could therefore also potentially yield information about the nature of the dark matter component.



## NEUTRINOS IN $N$ -BODY SIMULATIONS - A NOVEL APPROACH

Neutrinos are known to have a small but non-negligible impact on the matter power spectrum. At the linear level this effect is roughly of the order

$$\left. \frac{\Delta P_{\text{matter}}}{P_{\text{matter}}} \right|_{\text{max}} \sim -8 \frac{\Omega_{\nu}}{\Omega_m} \equiv -8f_{\nu}. \quad (6.1)$$

Naïvely one might have expected a suppression of the order  $\frac{\Omega_{\nu}}{\Omega_m}$ . The importance of neutrinos come from two factors: Background effects on the Friedmann law, and gravitational back-reactions on the perturbation level. In the absence of neutrinos, the gravitational potentials  $\phi$  and  $\psi$  in eqs. (1.21)-(1.24) are equal and constant during matter domination, and the CDM density contrast grows linearly with  $a$ . This is easy to see from the Poisson equation (eq. (1.21) in the limit  $ck \gg aH$ )

$$-\frac{c^2 k^2}{a^2} \psi = 4\pi G \delta \rho \quad (6.2)$$

and the fact that  $\bar{\rho} \propto a^{-3}$ . Introducing neutrinos, however, does not affect  $\delta \rho$  appreciably. The added matter component does impact the homogeneous expansion of the background nevertheless, thus causing a slight decay of the potential  $\phi$  and  $\psi$ . In fact, introducing the neutrinos in the background causes  $\delta_{CDM}$  to grow proportional to  $a^{1-3/5f_{\nu}}$ . Another background effect is that, for a fixed matter density in the present universe, the presence of neutrinos moves the matter-radiation equality to slightly later times, and thus delays the epoch where matter overdensities can grow appreciably. This is a simple consequence of the neutrinos being relativistic at the time of

matter-radiation equality. As for the neutrino induced back-reactions resulting in a different metric evolution, they can be attributed both to the fact that the shear term is no longer exactly zero (thus the potentials  $\phi$  and  $\psi$  are no longer exactly equal) and neutrino kinematical effects on gravitational collapse and subsequent virialisation. This effect is mostly the result of the neutrino free-streaming acting as a smoothing of the potential. Thus properly accounting for the neutrino impact on clustering is an important task, potentially also at the non-linear level where perturbation theory breaks down. See Lesgourgues and Pastor (2006) for a comprehensive review.

For a non-relativistic neutrino with temperature  $T_\nu(a)$  and mass  $m_\nu$  the thermal velocity decays as  $1+z$  and be written as

$$v_{\text{thermal}} \sim 3 \frac{T_\nu(a)}{m_\nu} = 3 \left( \frac{4}{11} \right)^{1/3} T_\gamma(a) \sim 150(1+z) \left( \frac{1 \text{ eV}}{m_\nu} \right) \text{ km/s.} \quad (6.3)$$

Converted to a free-streaming length this is

$$\lambda_{FS} = 2\pi \sqrt{\frac{3}{2}} \frac{v_{\text{thermal}}}{H} \sim 8 \frac{1+z}{\sqrt{\Omega_{\Lambda,0} + (1+z)^3 \Omega_{m,0}}} \left( \frac{1 \text{ eV}}{m_\nu} \right) h^{-1} \text{ Mpc} \quad (6.4)$$

For a neutrino mass close to the currently allowed upper bound on the sum of the neutrino masses  $m_\nu \sim 0.2 - 0.3 \text{ eV}$  (Ade et al., 2013) this is on the cluster scale thus potentially warranting a description at the non-linear level.<sup>a</sup> On the other hand, if the neutrino mass are close to the lower bound of the currently allowed sum of the neutrino masses  $m_\nu \sim 0.06 \text{ eV}$  in the standard hierarchy (Gonzalez-Garcia et al., 2012), a linear treatment of neutrinos is justified.

## 6.1 Neutrinos in $N$ -body simulations

Before presenting the different ways to implement neutrinos in  $N$ -body simulations we will first briefly go through the changes to the non-linear *total* matter power spectrum. The presence of neutrinos increases the maximal suppression relative to linear theory. This suppression is of the order

$$\left. \frac{\Delta P_{\text{matter}}}{P_{\text{matter}}} \right|_{\text{max}} \sim -9.8 f_\nu. \quad (6.5)$$

which is approximately 25 percent larger than the linear suppression in eq. (6.1). Furthermore, the *relative* matter power spectrum has a distinct through with the exact depth and position dependent on the neutrino mass. An interesting result is that the

<sup>a</sup>However, the upper bound is very dependent on the model assumptions, e.g., the dark energy equation of state.

suppression at the smallest scales is less than what is predicted from linear theory. This seemingly counter-intuitive consequence is a result of the presence of neutrinos delaying the onset of non-linear growth and increasing the non-linear scale (i.e. larger  $k_{\text{non-linear}}$ ) due to free-streaming smoothing. This means that the spectra with neutrinos are skewed with respect to  $\Lambda$ CDM resulting in a different mapping between scales. Thus naïvely calculating the relative power spectrum expectedly introduces a confusion of scales since the non-linear scales are shifted with respect to each other in the individual spectra.

The first self-consistent treatment of neutrinos in  $N$ -body simulations was done by Brandbyge et al. (2008). Here the neutrino component was included as a separate particle type into GADGET-2, and thermal velocities were properly included. The main difficulty in treating the neutrinos as particles is Poissonian shot noise from the large thermal velocities. Since the thermal velocities of the neutrinos are large comparable to the gravitational flow velocities, only a coarse sampling of the velocity distribution is feasible, and this introduces a white noise term in the neutrino spectrum. The white noise term will over time have a large impact on the full matter spectrum, and thus starting the simulations as late as possible is necessary, especially for a small neutrino mass  $m_\nu \lesssim 0.3$  eV. Furthermore, at large redshifts the true spectrum may even be dwarfed by the white noise contribution for small scales. For this reason the low-mass simulations were started at a redshift of  $z_i = 4$  with the initial conditions set up by the 2LPT formalism, with the neutrinos having a thermal velocity Fermi-Dirac contribution added. Viel et al. (2010) also performed  $N$ -body simulations with neutrino particles, and they analysed the impact on the Lyman- $\alpha$  forest. Bird et al. (2012b) found the corrections to the halofit model by Smith et al. (2003) resulting from neutrino clustering based on their implementation of a neutrino particle approach. Besides being computationally demanding, a major drawback of particle based neutrino simulations is the fact that it cannot reliably model the small scale neutrino spectrum. Even though the neutrinos are free-streaming, the low-momenta part of the distribution would cluster in haloes. The noise stemming from the velocity sampling, however, completely dwarfs this effect, rendering the small scale spectrum unreliable.

The shot noise from particle based neutrino simulations is especially difficult to treat for a small neutrino mass. A small neutrinos mass, however, also makes the neutrino component less non-linear. For this reason Brandbyge and Hannestad (2009) developed a method where the neutrino part was treated linearly independent of the  $N$ -body simulation by solving the full Boltzmann hierarchy, e.g. in CAMB, and subsequently modifying the PM grid by the resulting transfer functions. This grid approach speeds up neutrino simulations by an order of magnitude compared to a particle approach. Furthermore, for realistic neutrino masses, the difference between a particle and a grid based approach is at the sub-percent level. Since the neutri-

nos are not themselves included in the simulation, the actual non-linear neutrino spectrum can not be extracted. It should be noted that the grid approach is not fully consistent as neutrinos are evolved independently, and thus there is no neutrino-CDM back-reactions in the potential. For the matter power spectrum these two issues are not important, however, as the dominant effect is not the clustering properties of the neutrino component, it is the time integrated smoothing of the potential through free-streaming that is the main contribution. Ali-Haïmoud and Bird (2013)<sup>b</sup> expanded the grid based approach to be self-consistent by solving the linearised Boltzmann equation directly in the simulation, fully incorporating the gravitational interaction between the non-linear CDM component and the linear neutrino grid. An interesting result is how the linearised neutrino density contrast relates to the possibly non-linear density contrast for scales much smaller than the free-streaming length

$$\delta_v(k) \sim \left( \frac{k}{k_{FS}} \right)^{-2} \delta_{\text{matter}}(k) \quad (6.6)$$

Naturally this semi-linear treatment still fails to predict the non-linear clustering properties of the small momenta neutrinos. Typically neutrinos with momentum  $cp_v \lesssim T_v$  will be captured in haloes. The results of Ali-Haïmoud and Bird (2013) agrees with those in Brandbyge and Hannestad (2009) with a slightly better precision when compared to a particle approach. The advantage of a grid approach over a particle approach is the significantly reduced computational costs.

A full treatment of structure formation in a neutrino cosmology was done in Brandbyge and Hannestad (2010). This work combines the advantages of the grid and particle approaches by a hybrid approach. The basic idea is to keep the high-momenta neutrinos on a grid, and once the neutrino thermal velocity decreases to a level comparable to the gravitational flow velocities the grid is partly converted to neutrino particles. In practice the neutrino component is followed through several momenta bins, and low-momenta bins are converted to particles. Since a neutrino particle might eventually end up belonging to a higher momenta bin several bins are converted simultaneously to minimize this leakage effect. This method was then further utilised to provide a highly detailed analysis of the neutrino halo clustering in Brandbyge et al. (2010). Good agreement was found with the  $N$ -one-body method of Ringwald and Wong (2004), where the neutrinos are treated as independent test particles moving in the external gravitational potential of CDM halos. Unsurprisingly, it was found that tidal effects were relatively more important for neutrinos, especially in the low mass limit. Furthermore, an NFW-profile remained a good fit to the total

---

<sup>b</sup>Here a small inconsistency was also found in the analysis of Bird et al. (2012b) which originally claimed a larger discrepancy between the grid-based and particle-based approaches than found by Brandbyge and Hannestad (2009). After this inconsistency was fixed, the two works agreed on the discrepancy being small.

matter distribution in a halo, although a slightly flatter profile was found in the central regions for a constant total mass. A suppression of the (differential-)halo mass function was also found for large mass haloes, and the suppression increases with neutrino mass. This suppression was found to mainly be an effect of linear effects from the transfer function.

## 6.2 Introduction to the paper

It is with the difficulties of properly accounting for non-linearities in the neutrino spectrum in mind that our work Hannestad et al. (2012) should be viewed. We here develop a novel approach to implement neutrinos in  $N$ -body simulations building on the fluid approach of Shoji and Komatsu (2010). In this work we demonstrate that a neutrino component can easily be implemented as an isothermal baryonic gas with a Jeans scale comparable to the free-streaming scale, and our approach gives results consistent with a full, non-linear treatment of the neutrino component itself (the hybrid approach of Brandbyge and Hannestad (2010) discussed previously). While this method, much like the grid methods, expectedly fails to accurately track the actual neutrino clustering, the impact on the *total* non-linear matter clustering is accurately calculated. Since the total matter power spectrum is the quantity predicted by galaxy surveys, whereas the actual neutrino spectrum is extremely difficult to observe, this is not a severe limitation. The conclusion that a precise treatment of neutrino clustering is not important is also consistent with the results of the grid based approaches discussed in the previous section: The main effect of neutrinos on the total matter power spectrum is not neutrino clustering since they make up a negligible fraction of the total matter density. Instead it is the time-integrated smoothing effect on the potentials that has significant impact, and particularly the delayed onset of non-linear growth, that affects the total matter spectrum. The fact that the neutrino-effect can be modelled as a baryonic gas is fortunate, as most modern  $N$ -body simulations have the capability to implement baryons. With very minimal extensions, this can be changed to mimic the neutrino impact on the non-linear matter power spectrum. Furthermore, it also seems likely that our fluid approach could be implemented in a grid like manner by adapting the linear framework presented in chapter 4.

# Neutrinos in non-linear structure formation — a simple SPH approach

Steen Hannestad,<sup>a</sup> Troels Haugbølle<sup>b</sup> and Christian Schultz<sup>a</sup>

<sup>a</sup>Department of Physics and Astronomy, Aarhus University,  
Ny Munkegade, DK-8000 Aarhus C, Denmark

<sup>b</sup>Centre for Star and Planet Formation, Natural History Museum of Denmark,  
University of Copenhagen,  
Øster Voldgade 5-7, DK-1350 Copenhagen, Denmark

E-mail: [sth@phys.au.dk](mailto:sth@phys.au.dk), [haugboel@nbi.dk](mailto:haugboel@nbi.dk), [cs06@phys.au.dk](mailto:cs06@phys.au.dk)

Received October 13, 2011

Revised January 29, 2012

Accepted February 13, 2012

Published February 29, 2012

**Abstract.** We present a novel method for implementing massive neutrinos in  $N$ -body simulations. Instead of sampling the neutrino velocity distribution by individual point particles we take neutrino free-streaming into account by treating it as an effective redshift dependent sound speed in a perfect isothermal fluid, and assume a relation between the sound speed and velocity dispersion of the neutrinos. Although the method fails to accurately model the true neutrino power spectrum, it is able to calculate the total matter power spectrum to the same accuracy as more complex hybrid neutrino methods, except on very small scales. We also present an easy way to update the publicly available GADGET-2 version with this neutrino approximation.

**Keywords:** power spectrum, cosmological simulations, neutrino masses from cosmology

**ArXiv ePrint:** [1110.1257](https://arxiv.org/abs/1110.1257)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theory</b>	<b>2</b>
2.1	The Boltzmann equation	2
<b>3</b>	<b>Implementation in <math>N</math>-body simulations — <math>\nu</math>SPH</b>	<b>4</b>
3.1	Parameters and initial conditions	4
3.2	Neutrinos	5
<b>4</b>	<b>Results</b>	<b>5</b>
4.1	Power spectra	5
4.1.1	Evolution with redshift	6
4.1.2	Softening	7
4.1.3	Effective sound speed	7
<b>5</b>	<b>Discussions and conclusions</b>	<b>9</b>
<b>A</b>	<b>Including SPH neutrinos into GADGET-2</b>	<b>10</b>

---

## 1 Introduction

Over the past few years cosmological measurements have become much more precise and the need for correspondingly accurate theoretical predictions is rapidly growing. Upcoming measurements of weak gravitational lensing and galaxy clustering at high redshift are expected to reach percent level accuracy on scales where non-linear gravitational effects are important. On these scales there are a number of difficult challenges that must be overcome before percent level precision in the theoretical models can reliably be obtained. For example, baryonic effects must be properly taken into account (see e.g. [1] for a recent discussion of the required accuracy). Another important issue is the presence of massive neutrinos in the universe. Even though neutrinos are very light, they do contribute to cosmic structure formation and have an impact which is certainly larger than the accuracy goal. This has led to a number of analytic or semi-analytic studies of non-linear gravitational clustering with neutrinos, using either higher order perturbation theory or renormalisation group methods inspired by field theory [2–5]. Even though such methods are extremely important, in particular in the pseudo linear regime, they must be checked and/or calibrated using high precision  $N$ -body simulations, and they are hard to combine self consistently with  $N$ -body models including baryon physics.

Early simulations including massive neutrinos were exclusively done in mixed dark matter models with high neutrino masses and no dark energy (see e.g. [30]). However, observations strongly suggest that the underlying cosmological model is a variant of the  $\Lambda$ CDM model with relatively modest neutrino masses of approximately 1 eV or less. Of course it should be noted that current data actually favours more than the 3 neutrino degrees of freedom predicted by the standard model (see e.g. [9–15, 18]), and that an additional sterile neutrino

with a mass around 1 eV might explain this, as well as the observed short baseline neutrino oscillation anomalies (see e.g. [6–8] for recent discussions of the short baseline anomaly).

Neutrinos are well known to be difficult to handle in structure formation simulations. The reason is that for reasonable values of the neutrino mass, the thermal velocities of typical neutrinos vastly exceed the gravitational flow velocities until very late in the evolution of the universe. This in turn leads to very short time steps and thermal poisson noise in the simulations unless very large numbers of neutrino particles are used. This problem has been addressed in a number of different ways, for example by evolving the neutrino density on a grid, or by a combination of grid and particle methods. While such simulations can achieve high accuracy they are cumbersome and complex to run. We refer the reader to a number of papers discussing these issues [19–25].

Here we present a much simpler approach, based on the assumption that neutrinos can be treated as a perfect fluid with a finite sound speed. While this assumption is manifestly wrong, we demonstrate that quantities such as the matter power spectrum can be calculated with high accuracy (reaching an accuracy of 1-2% over the relevant range of scales) in such simulations which have the advantage of being very fast and simple to run. As in our previous studies of neutrino clustering we have not included baryonic effects in our simulations. While this will affect absolute values of power spectra to some extent, the ratio of power spectra with and without neutrino mass should be almost unchanged on the scales of interest (i.e.  $k \sim 1 - 2 h/\text{Mpc}$ ).

In the next section we briefly outline the theoretical framework for treating neutrinos in structure formation. In section 3 we describe our implementation on massive neutrinos as a fluid in the publicly available GADGET-2 code, and in section 4 we provide a detailed description of our results, including convergence tests. Finally section 5 contains a discussion and our conclusions. Appendix A describes the small changes necessary to add neutrino particles to the latest public version of the GADGET-2 code. A link to a working version is also provided.

## 2 Theory

### 2.1 The Boltzmann equation

In this subsection we will briefly outline how the evolution of massive neutrinos is followed in linear theory. The notation is identical to [17]. We use the metric in the conformal Newtonian gauge

$$ds^2 = -a^2(1 + 2\psi)d\tau^2 + a^2(1 - 2\phi)dx^2. \quad (2.1)$$

In a perturbed universe the phase-space distribution function is expanded to first order as follows

$$f = f_0 + \frac{\partial f_0}{\partial T} \delta T = f_0(1 + \Psi), \quad (2.2)$$

with the perturbation parametrised by  $\Psi = -d \ln f_0 / d \ln q \delta T / T$ , and the zeroth-order Fermi-Dirac phase-space distribution function given by

$$f_0(q) = \frac{1}{e^{q/T} + 1}. \quad (2.3)$$

Here  $q_i = ap_i$ , where  $p_i$  is the proper redshifting momentum and  $T$  is the neutrino temperature today.

After neutrino decoupling the collisionless Boltzmann equation evolves the distribution function as

$$\frac{df}{d\tau} = \frac{\partial f}{\partial \tau} + \frac{dx^i}{d\tau} \frac{\partial f}{\partial x^i} + \frac{dq}{d\tau} \frac{\partial f}{\partial q} + \frac{dn_i}{d\tau} \frac{\partial f}{\partial n_i} = 0. \quad (2.4)$$

Expanding the perturbation  $\Psi$  in a Legendre series the perturbed neutrino energy density is given by a weighted sum over neutrino momentum states

$$\delta\rho_\nu(k) = 4\pi a^{-4} \int q^2 dq \epsilon f_0 \Psi_0, \quad (2.5)$$

where  $\epsilon = (q^2 + a^2 m^2)^{1/2}$ .

From the Boltzmann equation the  $\Psi_l$ 's are related to each other and the metric potentials by

$$\dot{\Psi}_0 = -\frac{qk}{\epsilon} \Psi_1 - \dot{\phi} \frac{d \ln f_0}{d \ln q}, \quad (2.6)$$

$$\dot{\Psi}_1 = \frac{qk}{3\epsilon} \left( \Psi_0 - 2\Psi_2 \right) - \frac{\epsilon k}{3q} \psi \frac{d \ln f_0}{d \ln q}, \quad (2.7)$$

$$\dot{\Psi}_l = \frac{qk}{\epsilon(2l+1)} \left( l\Psi_{l-1} - (l+1)\Psi_{l+1} \right), \quad l \geq 2. \quad (2.8)$$

The second term on the right hand side of the equation for  $\dot{\Psi}_0$  encodes the effect of structure formation in evolving gravitational potentials, whereas the first term incorporates the effect of velocity on structure formation. The change in velocity is affected by the three terms on the right hand side of the equation for  $\dot{\Psi}_1$ . The first term encodes the effect of flow velocities redshifting in an expanding Universe, whereas the second term, found from the hierarchy of  $\dot{\Psi}_l$ 's, incorporates the effect of momentum (and redshift) dependent neutrino free-streaming. These two terms give rise to less structure, whereas the last term in the equation for  $\dot{\Psi}_1$  gives the acceleration as a gradient of the gravitational potential.

The above set of hierarchy equations can formally be integrated to give the fluid equations [17]

$$\dot{\delta} = -(1+w)(\theta - 3\dot{\phi}) - 3\frac{\dot{a}}{a} \left( \frac{\delta P}{\delta \rho} - w \right) \delta \quad (2.9)$$

$$\dot{\theta} = -\frac{\dot{a}}{a}(1-3w)\theta - \frac{\dot{w}}{1+w}\theta + \frac{\delta P/\delta \rho}{1+w} k^2 \delta - k^2 \sigma + k^2 \psi, \quad (2.10)$$

where  $\delta$  is the density contrast,  $\theta$  is the fluid velocity perturbation,  $w = P/\rho$  is the equation of state, and  $\sigma$  is the anisotropic stress.

For a perfect fluid we have  $\sigma = 0$  and the hierarchy can be truncated at  $l = 1$ . This reduces the Boltzmann equation to the usual continuity and Euler equations. Even though this is formally not allowed for massive neutrinos the system of equations could still be truncated by assuming a relation between  $\sigma$  and the effective sound speed,  $c_s$ , in the system. This approach was taken in [16] and shown to provide reasonable accuracy when tracking the linear evolution of neutrino perturbations. Since neutrinos are a subdominant component in structure formation (except in sterile neutrino warm dark matter scenarios), a relatively small error in the neutrino component translates into a much smaller error in such quantities as the total matter power spectrum.

In the neutrino fluid approach it is assumed that neutrinos behave like a perfect fluid with sound speed squared,  $c_s^2$ , given by

$$c_s^2 = \alpha F(z) \sigma_\nu^2, \quad (2.11)$$

where  $\sigma_\nu$  is the velocity dispersion of neutrinos, given by their effective temperature, and  $F(z)$  is some function of redshift (but not of position).  $\alpha$  is a constant of order 1. In the non-relativistic limit we have  $c_s^2 \ll 1$ ,  $w \sim 0$ , and the above equations reduce to

$$\dot{\delta} = -(\theta - 3\dot{\phi}) \quad (2.12)$$

$$\dot{\theta} = -\frac{\dot{a}}{a}\theta + c_s^2 k^2 \delta + k^2 \psi. \quad (2.13)$$

In the non-relativistic limit the neutrino velocity dispersion can be written as

$$\sigma_\nu^2 = \frac{\int p^2 / m_\nu^2 f_0(p) d^3 p}{\int f_0(p) d^3 p} \quad (2.14)$$

which in the same limit reduces to the following simple form [16]

$$\sigma_\nu^2 = \frac{15\zeta(5)}{\zeta(3)} \left(\frac{4}{11}\right)^{2/3} \frac{T_\gamma^2 (1+z)^2}{m_\nu^2}, \quad (2.15)$$

and  $F(z) \simeq (5/9)$ .

Again, we stress that this approach is manifestly not correct in the sense that a perfect fluid with finite sound speed exhibits acoustic oscillations on scales smaller than the Jeans scale because of the  $c_s^2 k^2 \delta$  term in the equation for  $\dot{\theta}$  while the true neutrino equations have a  $k^2 \sigma$  damping term which is not proportional to the local density. Since the Jeans scale for collisional particles is given as

$$k_J = \sqrt{\frac{3}{2}} \frac{H}{c_s (1+z)} \quad (2.16)$$

and the free-streaming scale for neutrinos is given as

$$k_{FS} = \sqrt{\frac{3}{2}} \frac{H}{\sigma_\nu (1+z)} \quad (2.17)$$

the relation in eq. (2.11) ensures that the Jeans scale is close to the free-streaming scale provided that  $\alpha F(z) \sim \mathcal{O}(1)$ .

However, since the neutrino transfer function at the starting point of the  $N$ -body simulation has been evolved using the exact Boltzmann equation hierarchy up to a point where the neutrino transfer function on small scales is orders of magnitude below the CDM transfer function, the most important point is *not* the difference between dissipation and oscillations, but just the fact that the neutrino transfer function cannot grow on small scales.

### 3 Implementation in $N$ -body simulations — $\nu$ SPH

#### 3.1 Parameters and initial conditions

Our simulations use the same cosmological parameters as in our previous studies, i.e. they assume a flat,  $\Lambda$ CDM model with  $\Omega_b = 0.05$ ,  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $h = 70$ . The

scalar fluctuation amplitude is assumed to be  $A_s = 2.3 \times 10^{-9}$ . While these values are not preferred by current cosmological data, they are sufficiently close to the best fit values to produce qualitatively identical results. We have used a modified version of the GADGET-2 Particle-Mesh N-body code [26, 27] to perform the simulations. Unless otherwise stated, our simulations are performed in boxes of size  $256 \text{ Mpc}/h$ , using  $512^3$  particles and a  $512^3$  force mesh. This is more than adequate for our purposes and we have carefully checked that resolution is not a problem in our simulations.

Initial conditions are generated using CAMB to evolve the linear transfer functions until a redshift of  $z = 49$ , at which point a 2nd order Lagrangian method is used to generate the perturbation field, as described in [19]. We assume adiabatic fluctuations.

### 3.2 Neutrinos

We implement neutrinos in GADGET-2 as a new SPH particle type. However, instead of the normal SPH equations we assume an isothermal neutrino gas where the only free parameter is the redshift dependent sound speed. This parameter is simply found by using the relation in eq. (2.11). Given that in the normal gas formulation the isothermal sound speed is redshift independent this has to be encoded in the equations of motion for the SPH particles. In GADGET-2 a constant entropy formulation is used that allows for change in entropy only at shocks or due to external physics. We could use the explicit equations (2.11) and (2.15), but for simplicity we instead use them to set the initial entropy, and then evolve the entropy forward using the simple evolution equation

$$\frac{dA}{dT} = -2A/T, \quad (3.1)$$

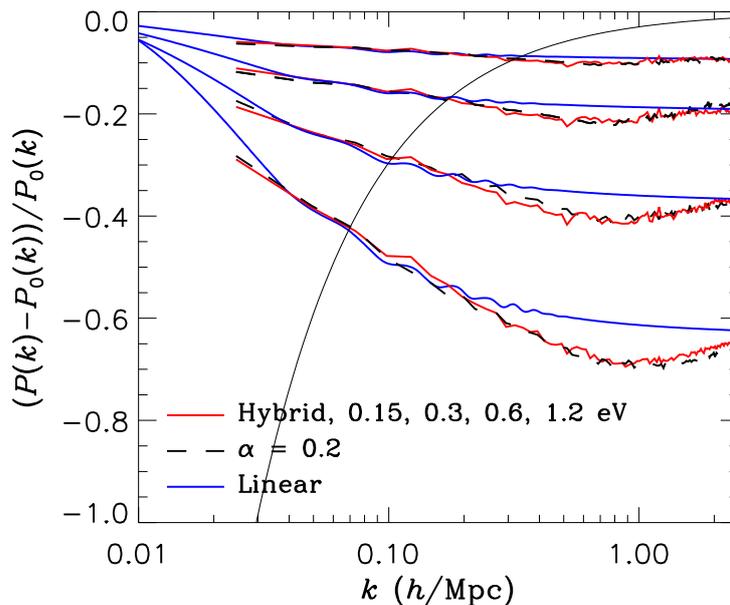
where  $T = a$  is the time variable and  $A = c_s^2$  is the effective entropy in GADGET-2. The treatment of neutrinos as a fluid induces acoustic oscillations at small scales. We do not want the oscillations to be reflected in the gravitational potential, and therefore it is essential to set the gravitational smoothing large enough that they do not impact the CDM. This is discussed below. Finally we note that although we have used SPH as a framework for treating neutrinos in the non-linear regime, a mesh implementation should work equally well.

## 4 Results

### 4.1 Power spectra

The total matter power spectra can reproduce results from the detailed hybrid code [21] to a precision of approximately 1-2% up to  $k \sim 1 - 2$ , more than sufficient for analysing most upcoming data sets.<sup>1</sup> In figure 1 we show the suppression of fluctuation power as a function of different neutrino masses up to 1.2 eV. The results are compared to the exact results from the hybrid code, and to results from linear perturbation theory. To provide an estimate of the maximum deviation tolerable we also show the statistical error on a power spectrum measurement from a galaxy survey with an effective volume of  $10 (\text{Gpc}/h)^3$ . From the figure it can be seen that deviations between the  $\nu$ SPH code and the hybrid codes are smaller than the statistical uncertainty at least up to  $k \sim 1 - 2$ . For illustration purposes we show CDM density plots from simulations with  $m_\nu = 0.15 \text{ eV}$  and  $1.2 \text{ eV}$  respectively in figure 5.

<sup>1</sup>Note that in previous papers, power spectra have generally been compared with results from purely particle based simulations. However, the hybrid code has been extensively tested and results match those from the particle code to better than % precision on all scales. Since the hybrid code is much faster we have therefore switched to using that as our reference model.



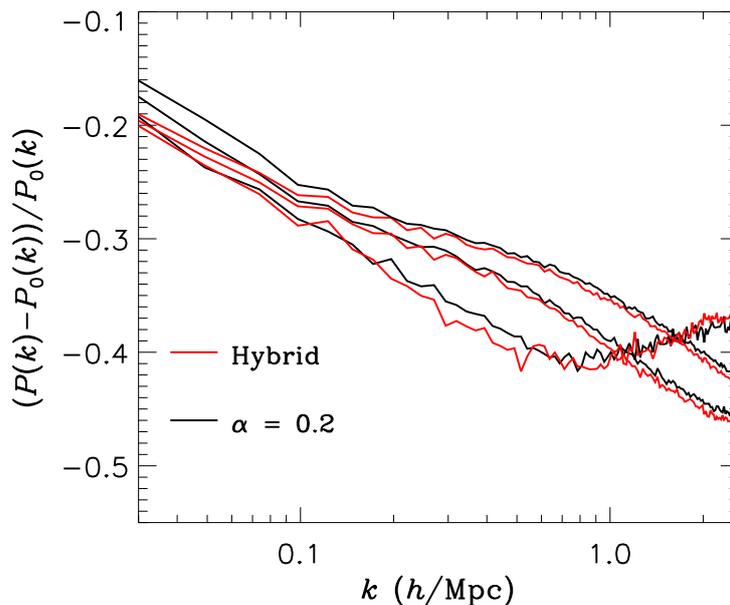
**Figure 1.** Power spectrum damping for various neutrino masses. Thin black line is the estimated  $1\sigma$  uncertainty on  $P(k)$  from a galaxy survey with effective volume  $V_{\text{eff}} = 10 \text{ (Gpc/h)}^3$ , using the simple prescription from eq. (7) in [28].

However, the neutrino clustering is not reproduced to the same level of accuracy. As is the case for grid-based simulations, an important ingredient has been removed. Decoupled neutrinos have a distribution given by a relativistic Fermi-Dirac function. This means that even though the mean momentum is large, a small proportion of neutrinos will always have low initial momenta and be subject to gravitational clustering. This phenomenon was studied in great detail in [22] and the neutrino density profiles in typical dark matter halos extracted. However, in the  $\nu$ SPH approach the neutrino gas is treated as isothermal which in particle language is approximately equal to providing all neutrinos with the same initial thermal velocity. While this is a reasonable approximation to the real physical situation it fails to follow the small proportion of low momentum neutrinos that actually cluster in halos.

In terms of the neutrino power spectrum this can be seen as a lack of power at high  $k$ . Again, we stress that this is not a problem in most cases, i.e. where the total matter power spectrum is studied: weak lensing, large scale structure surveys, Lyman- $\alpha$  surveys etc. However, the code is not well suited for studies of neutrino behaviour, for example in the context of relic neutrino detection. In practise the capability to follow the detailed neutrino distribution is rarely needed and the ease with which the  $\nu$ SPH method can be implemented far outweighs its limitations.

#### 4.1.1 Evolution with redshift

The pronounced dip in suppression is located at the point where the  $\Lambda$ CDM model has gone non-linear while the  $\nu\Lambda$ CDM model is still in the linear regime. At higher redshifts this point is located at higher values of  $k$ , as can be seen in figure 2. The amplitude of the suppression also increases because the overall power spectrum normalisation is lower at high redshift. We have checked explicitly that the  $\nu$ SPH code produces results which are compatible with the



**Figure 2.** Power spectrum suppression at  $z = 4, 7/3, 0$  (in order of decreasing amplitude on large scales) for  $0.6\text{ eV}$  neutrinos in the  $\nu\text{SPH}$  approximation.

hybrid code. Figure 2 clearly shows that the agreement between the two codes is even better at higher redshifts than at  $z = 0$ .

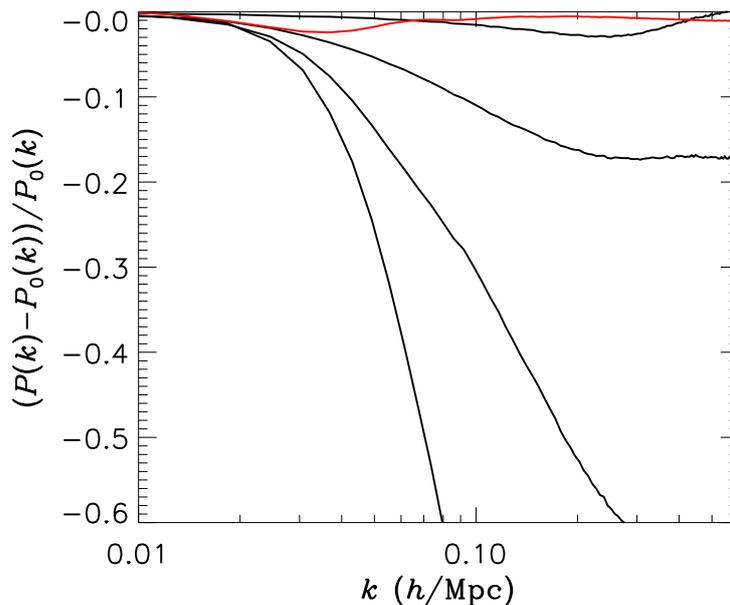
#### 4.1.2 Softening

One important effect which must be taken into account is that if the effective softening length for SPH particles is set too low, the system fails to properly track evolution not only of the SPH system, but also the dominant CDM component. This can be seen in the total matter power spectrum in figure 3. The matter power spectrum drops steeply on scales smaller than the effective sound propagation scale. That this is indeed due to waves can be seen in figure 6 where the right side panels clearly show oscillations in both the CDM and the neutrino components. What happens is that the strong waves in the neutrino component wipe out the CDM fluctuations before they become highly non-linear, except in a few very high density regions.

When the softening length is increased the spurious acoustic waves vanish and the CDM system behaves as expected. We stress that in the simulations actually used to compare against the hybrid code, this is not an issue. We merely show figures 3 and 6 for illustration purposes. Any value of the neutrino softening length fulfilling the criterion that  $r_{\text{soft}} \gtrsim r_{\text{side}}/N^{1/3}$  is generally safe to use. For CDM we have used a softening length of  $32\text{ kpc}/h$ , but our results do not depend on this assumption at all.

#### 4.1.3 Effective sound speed

In figure 4 we show the relative suppression for  $0.6\text{ eV}$  for different values of the effective sound speed. For low values the overall suppression is systematically underestimated. However, for values around  $0.2$  and up, the small scale structure fits extremely well. There is a simple reason for this: On small scales neutrinos contribute almost nothing to the overall structure. Even if they do cluster in halos, the fraction of halo masses in neutrinos is tiny. Therefore it is



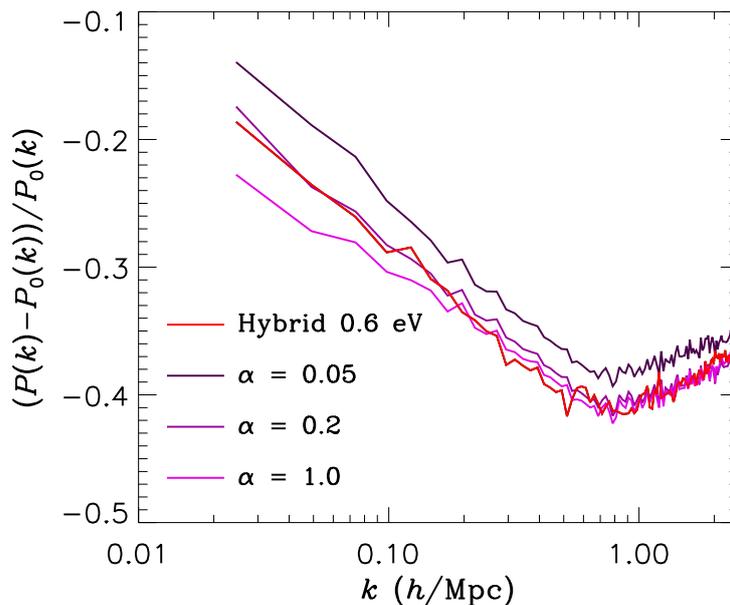
**Figure 3.** The effect of neutrino softening in the  $\nu$ SPH simulations. Black lines are  $256^3$  simulations in a  $1024 \text{ Mpc}/h$  box. The lines show the matter power spectrum for  $r_{\text{soft}} = 2000, 1000, 500, 250 \text{ kpc}$ , relative to a model with  $r_{\text{soft}} = 4000 \text{ kpc}$ . The red line shows the same for  $r_{\text{soft}} = 1000 \text{ kpc}$  relative to  $r_{\text{soft}} = 4000 \text{ kpc}$ , but for a  $512^3$  simulation. As can be seen, the effect diminishes dramatically when a larger number of particles is used.

a good approximation when calculating global quantities to treat neutrinos as non-clustering on sufficiently small scales. This behaviour is captured by our treatment and even though the method significantly under predicts the amount of power in the neutrino component on small scales it has no effect on quantities such as the power spectrum.

On larger scales, the effective velocity must be tuned to fit with the linear theory prediction. Simulations with too high a sound speed underestimate power on large scales, while the opposite is true for neutrino components too similar to CDM. The reason for this is relatively simple to understand. In the SPH implementation, neutrinos are modeled as a single fluid with a single sound speed. However, the real neutrino distribution always contains particles with velocities very different from the average. On large scales the small fraction of neutrinos with large momenta are important and if not included the power spectrum suppression becomes too small on large scales. This is not a serious failing, however, because the problem only occurs where structures are highly linear and if high accuracy is needed at all  $k$ , the SPH spectrum can be replaced by the linear power spectrum at small  $k$ .

In conclusion: Quantities such as the power spectrum are relatively robust against variations in the actual value of the effective sound speed on small scales, while on large scales it should be tuned to fit the linear theory prediction. In practise we find a value of  $\alpha$  around 0.2-0.3 to be accurate for any neutrino mass up to  $1.2 \text{ eV}$ .

One might worry that for a given simulation this parameter has to be tuned to achieve sufficient accuracy. However, the semi-analytic estimate in ref. [31] predicts a value of roughly of approximately 0.3, close to the result we find. Furthermore we find that our results are not strongly dependent on the exact value of  $\alpha$  as long as it is in the 0.2-0.4 range. The best fit value of  $\alpha$  will not in general be cosmology dependent to any significant extent as long as the dominant matter component is CDM.



**Figure 4.** The relative suppression in a model with 0.6 eV neutrinos.

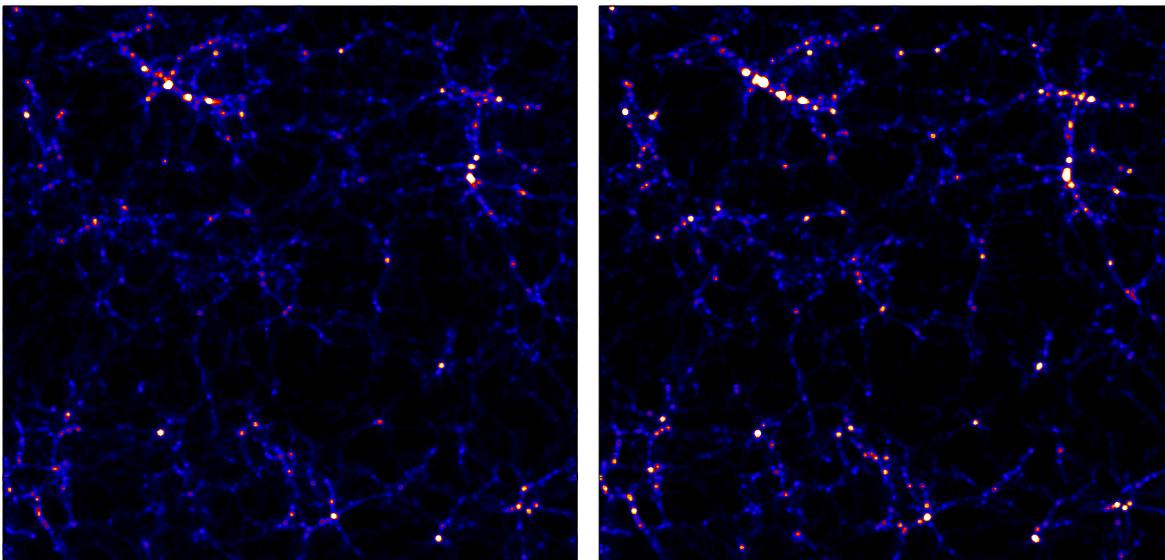
## 5 Discussions and conclusions

We have presented a very simple scheme for implementing neutrinos in  $N$ -body simulations using SPH and the assumption that neutrinos behave as an ideal, isothermal fluid. While this assumption is wrong because neutrinos in reality are free-streaming, we have demonstrated that errors in quantities such as the matter power spectrum can be kept under control at the 1-2% level over all relevant scales, sufficient for analysing for example data from the upcoming EUCLID satellite mission. The reason is that although neutrino small scale clustering is severely underestimated in the  $\nu$ SPH code, the contribution of neutrinos to small scale matter clustering is negligible and quantities such as the power spectrum can therefore be estimated quite accurately.

The SPH implementation presented here has a number of advantages over the hybrid neutrino code presented in [21]: It is very easy to implement in existing codes, such as GADGET-2 (see appendix A for how to implement it). The only requirement is the addition of a new SPH particle type with a modified time-temperature relation. As for initial conditions, the only required input is a set of transfer functions for CDM, baryons and neutrinos. This is in contrast to the hybrid code which requires solving the linear theory neutrino Boltzmann equations either separately or in the  $N$ -body code. The hybrid code also necessitates solving the momentum dependent Boltzmann equations and keeping this information as the  $N$ -body code is evolved forward in time.

While the  $\nu$ SPH code fails to accurately map neutrino structures on small scales, it is more than adequate for power spectrum calculation. The difference compared to an exact particle based simulation (or the hybrid implementation) can be expected to be less than 1-2% on all scales for any  $k$  (at very high  $k$  the exact behaviour of neutrinos have no impact because they do not contribute to clustering at all).

This in turn means that it is well suited for calculations of e.g. weak lensing. Cluster abundances can also be reliably calculated. As was demonstrated in [22], the halo mass



**Figure 5.** Density in the CDM component for a neutrino mass of 1.2 eV (left) and 0.15 eV (right). Simulations are made with  $512^3$  particles in a  $256 \text{ Mpc}/h$  box, and the figures show slices of thickness  $10 \text{ Mpc}/h$ .

function in models with neutrinos can be accurately calculated by using the correct power spectrum and making the assumption that neutrinos do not contribute significantly to halo masses, i.e. exactly the assumption needed for the SPH approach to work.

The only real drawback of the code is that it cannot follow neutrino structures. However, this feature is only really needed for calculations of quantities such as the local neutrino density.

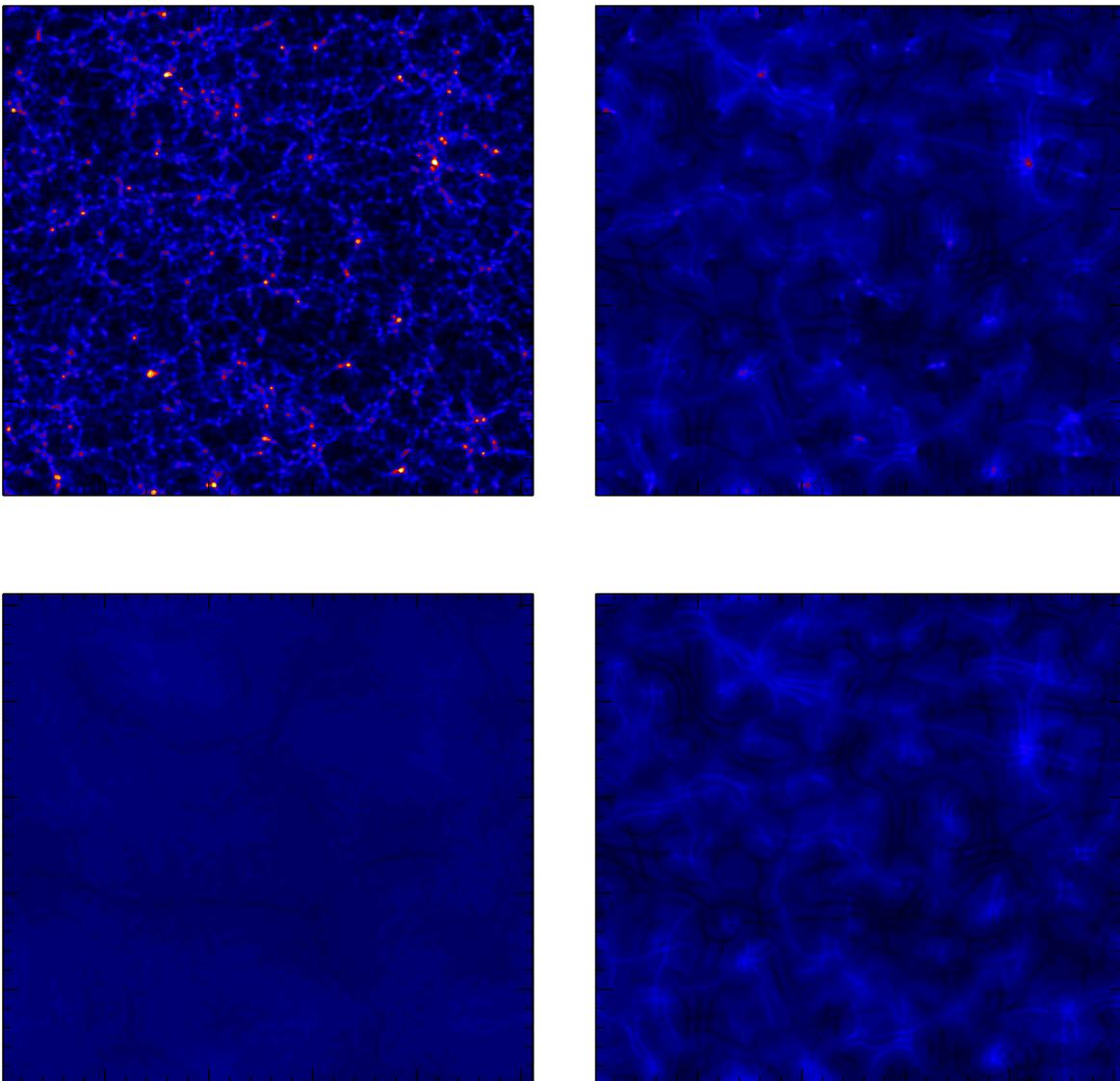
In summary, the code presented here is more than sufficiently accurate and extremely easy to implement in existing  $N$ -body codes. For this reason the  $\nu$ SPH method could well become the standard method for implementing neutrinos in structure formation simulations because of its simplicity.

## Acknowledgments

We acknowledge computing resources from the Danish Center for Scientific Computing (DCSC).

## A Including SPH neutrinos into GADGET-2

Patching a stock version of GADGET-2 with SPH neutrinos is very easy. It is only a matter of inserting  $\simeq 100$  lines of code. A stock version of GADGET-2 patched with SPH neutrinos can be found on <http://phys.au.dk/~steen/nusph.tar.gz> — but note that here the neutrinos replace baryons as particle type 0. This means that in this version it is not possible to have both baryons and neutrinos, however this is easy to ameliorate should one wish to. A normal GADGET-2 makefile can be used to compile this code, the only modification is to add `OPT+=DNEUTRINO_FLUID`. Without this option the code behaves like the stock version. It is also necessary to add an `OmegaNeutrino` value in the parameter file.



**Figure 6.** Density in simulations with  $0.6\text{ eV}$  neutrinos in a  $1024\text{ Mpc}/h$  box using  $256^3$  resolution. The plot shows a slice of thickness  $40\text{ Mpc}/h$ . Top row is CDM, bottom neutrinos. The left column is with an SPH softening of  $4000\text{ kpc}/h$ , and the right with a softening of  $250\text{ kpc}/h$ . The acoustic waves are seen in both the neutrino and the CDM components in the right column.

The above minimal SPH neutrino code is based on GADGET-2 version 2.0.7, so to see the changes it is possible to do a simple

```
diff -r /path/to/Gadget2-2.0.7_nusph/Gadget2 /path/to/Gadget2-2.0.7/Gadget2
```

between this version and a stock version of the GADGET-2 code. It is not too much work to put in the extra code by hand in a modified GADGET-2 code. It should also be possible to run a three-way merge, e.g. diff3, on each of the files `allvars.h`, `begrun.c`, `global.c`, `hydra.c`, `init.c`, `io.c` and `read_ic.c` and handle any conflicts.

## References

- [1] K. Heitmann, M. White, C. Wagner, S. Habib and D. Higdon, *The Coyote Universe I: Precision Determination of the Nonlinear Matter Power Spectrum*, *Astrophys. J.* **715** (2010) 104 [[arXiv:0812.1052](#)] [[INSPIRE](#)].
- [2] J. Lesgourgues, S. Matarrese, M. Pietroni and A. Riotto, *Non-linear Power Spectrum including Massive Neutrinos: the Time-RG Flow Approach*, *JCAP* **06** (2009) 017 [[arXiv:0901.4550](#)] [[INSPIRE](#)].
- [3] S. Saito, M. Takada and A. Taruya, *Nonlinear power spectrum in the presence of massive neutrinos: perturbation theory approach, galaxy bias and parameter forecasts*, *Phys. Rev. D* **80** (2009) 083528 [[arXiv:0907.2922](#)] [[INSPIRE](#)].
- [4] Y.Y. Wong, *Higher order corrections to the large scale matter power spectrum in the presence of massive neutrinos*, *JCAP* **10** (2008) 035 [[arXiv:0809.0693](#)] [[INSPIRE](#)].
- [5] S. Saito, M. Takada and A. Taruya, *Impact of massive neutrinos on nonlinear matter power spectrum*, *Phys. Rev. Lett.* **100** (2008) 191301 [[arXiv:0801.0607](#)] [[INSPIRE](#)].
- [6] C. Giunti and M. Laveder, *3+1 and 3+2 Sterile Neutrino Fits*, *Phys. Rev. D* **84** (2011) 073008 [[arXiv:1107.1452](#)] [[INSPIRE](#)].
- [7] C. Giunti and M. Laveder, *Status of 3+1 Neutrino Mixing*, *Phys. Rev. D* **84** (2011) 093006 [[arXiv:1109.4033](#)] [[INSPIRE](#)].
- [8] J. Kopp, M. Maltoni and T. Schwetz, *Are there sterile neutrinos at the eV scale?*, *Phys. Rev. Lett.* **107** (2011) 091801 [[arXiv:1103.4570](#)] [[INSPIRE](#)].
- [9] J. Hamann, S. Hannestad, G.G. Raffelt and Y.Y. Wong, *Sterile neutrinos with eV masses in cosmology: How disfavoured exactly?*, *JCAP* **09** (2011) 034 [[arXiv:1108.4136](#)] [[INSPIRE](#)].
- [10] J. Hamann, S. Hannestad, G.G. Raffelt, I. Tamborra and Y.Y. Wong, *Cosmology seeking friendship with sterile neutrinos*, *Phys. Rev. Lett.* **105** (2010) 181301 [[arXiv:1006.5276](#)] [[INSPIRE](#)].
- [11] E. Giusarma et al., *Constraints on massive sterile neutrino species from current and future cosmological data*, *Phys. Rev. D* **83** (2011) 115023 [[arXiv:1102.4774](#)] [[INSPIRE](#)].
- [12] Z. Hou, R. Keisler, L. Knox, M. Millea and C. Reichardt, *How Additional Massless Neutrinos Affect the Cosmic Microwave Background Damping Tail*, [arXiv:1104.2333](#) [[INSPIRE](#)].
- [13] R. Keisler et al., *A Measurement of the Damping Tail of the Cosmic Microwave Background Power Spectrum with the South Pole Telescope*, *Astrophys. J.* **743** (2011) 28 [[arXiv:1105.3182](#)] [[INSPIRE](#)].
- [14] A.X. Gonzalez-Morales, R. Poltis, B.D. Sherwin and L. Verde, *Are priors responsible for cosmology favoring additional neutrino species?*, [arXiv:1106.5052](#) [[INSPIRE](#)].
- [15] J.R. Kristiansen and O. Elgaroy, *Reactor sterile neutrinos, dark energy and the age of the universe*, [arXiv:1104.0704](#) [[INSPIRE](#)].
- [16] M. Shoji and E. Komatsu, *Massive Neutrinos in Cosmology: Analytic Solutions and Fluid Approximation*, *Phys. Rev. D* **81** (2010) 123516 [*Erratum ibid.* **D 82** (2010) 089901] [[arXiv:1003.0942](#)] [[INSPIRE](#)].
- [17] C.-P. Ma and E. Bertschinger, *Cosmological perturbation theory in the synchronous and conformal Newtonian gauges*, *Astrophys. J.* **455** (1995) 7 [[astro-ph/9506072](#)] [[INSPIRE](#)].
- [18] WMAP collaboration, E. Komatsu et al., *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophys. J. Suppl.* **192** (2011) 18 [[arXiv:1001.4538](#)] [[INSPIRE](#)].

### 6.3 The future and outlook

The recent advances in the neutrino  $N$ -body simulations have taught us that to understand the observable non-linear structure of the cosmos an exact treatment of neutrinos is not necessary. Therefore a fluid approach is well justified and requires a negligible amount of modification to existing simulations capable of treating baryons. A grid based approach requires a somewhat larger modification, but has the advantage that it is computationally faster than a fluid simulation. Whether grid methods are still faster if the linear equations are solved directly in the simulation, and thus including the non-linear potential, is not obvious. This will necessarily introduce another layer of computations and give some overhead. A particle based approach, while possible, does not provide significant benefits over the other methods. If one requires an accurate treatment of the clustering properties of neutrinos the hybrid method, where neutrinos are split into different momenta bins and evolved separately, is the best approach. In fact it is the only approach, as particle simulations have significant noise at the interesting scales for any reasonable number of particles. In any case, if the sum of the neutrino masses is somewhat smaller than the upper bound of  $0.2 - 0.3$  eV, the neutrinos will not cluster non-linearly, and while they will still have an effect on the non-linear matter power spectrum, their clustering properties can be understood within linear theory.



**Part IV**

**Outlook**



## CONCLUSION AND OUTLOOK

The work presented in this dissertation touches upon the model dependence of cosmological structure formation. A method for ray-tracing  $N$ -body simulations have been presented, and while the complete implementation is not entirely finished, it could realistically be put into a working condition with a few months' work. We have also presented a framework to track structure formation in a dynamical dark energy cosmology directly in an  $N$ -body simulation. This will likely result in a publication shortly. While none of the codes mentioned are public as of writing, the aim has all along been to make them public, so any one expressing interest will be given access.

The paper Schultz et al. (2014) used a conservative approach to constrain WDM models. We found that single component WDM models with a mass of  $m_{\text{thermal}} \sim 0.5 - 1.5$  keV are not able to match low- $z$  and high- $z$  observations simultaneously. This is interesting since these models are usually employed to solve the small scale crisis of  $\Lambda$ CDM. The results presented adds to the increasing pile of evidence suggesting that thermal and thermal-like warm dark matter models are not viable. Furthermore, our analysis can also be generalised to other models predicting different abundances of haloes. Our conclusions are only based on the relative halo mass functions between models, assuming that the mass-luminosity relation is invariant between these. While this assumption may not necessarily be the case for WDM cosmologies compared to a fiducial  $\Lambda$ CDM model, it only serves to make the conclusions more conservative to the extent that the  $\Lambda$ CDM model gives an upper bound. As long as an invariant mass-luminosity relation between models is approximately true, our method should be applicable irrespective of the initial transfer function.

In the paper Hannestad et al. (2012) we introduced a new technique to imple-

ment neutrinos in  $N$ -body simulations. The advantage of this method is the fact that it is very easy to implement and produces robust results. By ignoring the true neutrino distribution and only considering the contribution from the free-streaming property of neutrinos we found good agreement with the much more elaborate hybrid method that properly accounts for neutrino clustering. The method can be implemented into existing  $N$ -body simulations capable of treating baryons with very little work, thus making it an attractive alternative to other methods. In the future it would be interesting to combine this method with the linear Fourier based approach to treat dynamical dark energy that was presented in chapter 4. A linear treatment of neutrinos in  $N$ -body simulations has been shown to give good results for small neutrino masses, and solving the linear fluid equations on the run should only introduce a small computational overhead.

The progress in the field of cosmology in the last twenty-five years seems to retain its momentum. With upcoming space based missions like Euclid and the James Webb Space Telescope the near future looks very bright. These missions will help to a great extent to constrain the available parameter space for dark matter and dark energy models. While explaining the true nature of these dark components is probably still some decades down the road, the available freedom in models will gradually decrease. Furthermore, settling the BICEP2 controversy with upcoming Keck and Planck high frequency data will also be extremely interesting to follow. Also on the computational side important advances should be expected. With more realistic treatment of gas physics in  $N$ -body simulations the discrepancy between observations and theory may decrease. To this extent the exponential increase in computing power predicted by Moore's law will also allow further sophistication of the algorithms. For the space curious person it is indeed a very interesting time to be alive.

## ACKNOWLEDGEMENTS

I would like to thank my supervisor Steen Hannestad for supervision and the opportunity to spend some years in academia. I enjoyed the small talk and informal coffee breaks.

The cosmology group at Aarhus University provided a comfortable working environment: Ole Bjælde, Rasmus Sloth-Hansen, Io Odderskov and Thomas Tram. I would like to thank Tobias Basse in particular for proof-reading this dissertation and constructive professional discussions.

Jacob Brandbyge proved to be an inspiration at the very early stages of my ph.d., but also later acted as a mentor in figuring out the life post-graduation. For this I am thankful.

The cosmology group at University of California, Irvine also provided a creative environment that greatly enhanced productivity. Especially my close collaborators Jose Oñorbe, James Bullock and Kevork Abazajian deserve a special thanks.

I would also like to thank my students throughout the years for having been a valuable source of inspiration. I came to appreciate the fact that you do not truly understand a subject until you have taught it to others. I enjoyed both lecturing at the university level, but also the public lectures I was invited to give, gave a great deal of satisfaction.



## BIBLIOGRAPHY

- P. Ade et al. Planck 2013 results. XVI. Cosmological parameters. 2013.
- Y. Ali-Haïmoud and S. Bird. An efficient implementation of massive neutrinos in non-linear structure formation simulations. *Monthly Notices of the Royal Astronomical Society*, 428(4):3375–3389, 2013. doi: 10.1093/mnras/sts286. URL <http://mnras.oxfordjournals.org/content/428/4/3375.abstract>.
- J.-M. Alimi, A. Füzfa, V. Boucher, Y. Rasera, J. Courtin, and P.-S. Corasaniti. Imprints of dark energy on cosmic structure formation – i. realistic quintessence models and the non-linear matter power spectrum. *Monthly Notices of the Royal Astronomical Society*, 401(2):775–790, 2010. doi: 10.1111/j.1365-2966.2009.15712.x. URL <http://mnras.oxfordjournals.org/content/401/2/775.abstract>.
- R. E. Angulo, O. Hahn, and T. Abel. The Warm DM halo mass function below the cut-off scale. 2013. URL <http://inspirehep.net/record/1227608>.
- S. Anselmi, G. Ballesteros, and M. Pietroni. Non-linear dark energy clustering. *Journal of Cosmology and Astroparticle Physics*, 2011(11):014, 2011. URL <http://stacks.iop.org/1475-7516/2011/i=11/a=014>.
- M. Baldi. Dark energy simulations. *Physics of the Dark Universe*, 1(1–2): 162 – 193, 2012. ISSN 2212-6864. doi: <http://dx.doi.org/10.1016/j.dark.2012.10.004>. URL <http://www.sciencedirect.com/science/article/pii/S2212686412000088>. Next Decade in Dark Matter and Dark Energy Next Decade in Dark Matter and Dark Energy.
- R. Barkana, Z. Haiman, and J. P. Ostriker. Constraints on warm dark matter from cosmological reionization. *The Astrophysical Journal*, 558(2):482, 2001. URL <http://stacks.iop.org/0004-637X/558/i=2/a=482>.
- J. Barnes and P. Hut. A Hierarchical  $O(N \log N)$  Force Calculation Algorithm. *Nature*, 324:446–449, 1986. doi: 10.1038/324446a0.
- M. Bartelmann and P. Schneider. Weak gravitational lensing. *Phys.Rept.*, 340:291–472, 2001. doi: 10.1016/S0370-1573(00)00082-X.

- P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu. The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. *ApJ*, 762:109, Jan. 2013. doi: 10.1088/0004-637X/762/2/109.
- E. Bertschinger. Cosmological dynamics: Course 1. 1993.
- E. Bertschinger. Cosmological perturbation theory and structure formation. 2001.
- BICEP2 Collaboration, P. A. R. Ade, R. W. Aikin, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, J. A. Brevik, I. Buder, E. Bullock, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, M. Halpern, M. Hasselfield, S. R. Hildebrandt, G. C. Hilton, V. V. Hristov, K. D. Irwin, K. S. Karkare, J. P. Kaufman, B. G. Keating, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, E. M. Leitch, M. Lueker, P. Mason, C. B. Netterfield, H. T. Nguyen, R. O’Brien, R. W. Ogburn, IV, A. Orlando, C. Pryke, C. D. Reintsema, S. Richter, R. Schwarz, C. D. Sheehy, Z. K. Staniszewski, R. V. Sudiwala, G. P. Teply, J. E. Tolán, A. D. Turner, A. G. Vieregg, C. L. Wong, and K. W. Yoon. BICEP2 I: Detection Of B-mode Polarization at Degree Angular Scales. *ArXiv e-prints*, Mar. 2014.
- S. Bird, M. Viel, and M. G. Haehnelt. Massive neutrinos and the non-linear matter power spectrum. *Monthly Notices of the Royal Astronomical Society*, 420(3): 2551–2561, 2012a. doi: 10.1111/j.1365-2966.2011.20222.x. URL <http://mnras.oxfordjournals.org/content/420/3/2551.abstract>.
- S. Bird, M. Viel, and M. G. Haehnelt. Massive Neutrinos and the Non-linear Matter Power Spectrum. *Mon.Not.Roy.Astron.Soc.*, 420:2551–2561, 2012b. doi: 10.1111/j.1365-2966.2011.20222.x.
- D. Blas, J. Lesgourgues, and T. Tram. The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes. *J. Cosmology Astropart. Phys.*, 7:034, July 2011. doi: 10.1088/1475-7516/2011/07/034.
- P. Bode, J. P. Ostriker, and N. Turok. Halo formation in warm dark matter models. *The Astrophysical Journal*, 556(1):93, 2001. URL <http://stacks.iop.org/0004-637X/556/i=1/a=93>.
- M. S. Bovill and M. Ricotti. Where are the Fossils of the First Galaxies? II. True Fossils, Ghost Halos, and the Missing Bright Satellites. *Astrophys.J.*, 741:18, 2011. doi: 10.1088/0004-637X/741/1/18.
- A. Boyarsky, J. Lesgourgues, O. Ruchayskiy, and M. Viel. Lyman- $\alpha$  constraints on warm and on warm-plus-cold dark matter models. *JCAP*, 0905:012, 2009a. doi: 10.1088/1475-7516/2009/05/012.

- A. Boyarsky, O. Ruchayskiy, and M. Shaposhnikov. The Role of sterile neutrinos in cosmology and astrophysics. *Ann.Rev.Nucl.Part.Sci.*, 59:191–214, 2009b. doi: 10.1146/annurev.nucl.010909.083654.
- M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *Mon.Not.Roy.Astron.Soc.*, 415:L40, 2011.
- M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. The Milky Way’s bright satellites as an apparent failure of  $\Lambda$ CDM. *MNRAS*, 422:1203–1218, May 2012. doi: 10.1111/j.1365-2966.2012.20695.x.
- M. Boylan-Kolchin, J. S. Bullock, S. T. Sohn, G. Besla, and R. P. van der Marel. The space motion of leo i: The mass of the milky way’s dark matter halo. *The Astrophysical Journal*, 768(2):140, 2013. URL <http://stacks.iop.org/0004-637X/768/i=2/a=140>.
- J. Brandbyge and S. Hannestad. Grid Based Linear Neutrino Perturbations in Cosmological N-body Simulations. *JCAP*, 0905:002, 2009. doi: 10.1088/1475-7516/2009/05/002.
- J. Brandbyge and S. Hannestad. Resolving Cosmic Neutrino Structure: A Hybrid Neutrino N-body Scheme. *JCAP*, 1001:021, 2010. doi: 10.1088/1475-7516/2010/01/021.
- J. Brandbyge, S. Hannestad, T. Haugbolle, and B. Thomsen. The Effect of Thermal Neutrino Motion on the Non-linear Cosmological Matter Power Spectrum. *JCAP*, 0808:020, 2008. doi: 10.1088/1475-7516/2008/08/020.
- J. Brandbyge, S. Hannestad, T. Haugboelle, and Y. Y. Wong. Neutrinos in Non-linear Structure Formation - The Effect on Halo Properties. *JCAP*, 1009:014, 2010. doi: 10.1088/1475-7516/2010/09/014.
- A. M. Brooks, M. Kuhlen, A. Zolotov, and D. Hooper. A Baryonic Solution to the Missing Satellites Problem. *Astrophys.J.*, 765:22, 2013. doi: 10.1088/0004-637X/765/1/22.
- R. A. C. Croft, D. H. Weinberg, M. Bolte, S. Burles, L. Hernquist, N. Katz, D. Kirkman, and D. Tytler. Toward a precise measurement of matter clustering: Ly- $\alpha$  forest data at redshifts 2-4. *The Astrophysical Journal*, 581(1):20, 2002. URL <http://stacks.iop.org/0004-637X/581/i=1/a=20>.
- M. Davis, G. Efstathiou, and C. Frenk. k white, sdm, 1985. *ApJ*, 292:371.
- W. de Blok. The Core-Cusp Problem. *Adv.Astron.*, 2010:789293, 2010. doi: 10.1155/2010/789293.

- R. S. de Souza, A. Mesinger, A. Ferrara, Z. Haiman, R. Perna, et al. Constraints on Warm Dark Matter models from high-redshift long gamma-ray bursts. *Mon.Not.Roy.Astron.Soc.*, 432:3218, 2013. doi: 10.1093/mnras/stt674.
- S. Dodelson and L. M. Widrow. Sterile neutrinos as dark matter. *Phys. Rev. Lett.*, 72: 17–20, Jan 1994. doi: 10.1103/PhysRevLett.72.17. URL <http://link.aps.org/doi/10.1103/PhysRevLett.72.17>.
- R. Flauger, J. C. Hill, and D. N. Spergel. Toward an Understanding of Foreground Emission in the BICEP2 Region. 2014.
- P. Fosalba, E. Gaztanaga, F. Castander, and M. Manera. The onion universe: all sky light-cone simulations in shells. 2007.
- S. Garrison-Kimmel, M. Rocha, M. Boylan-Kolchin, J. S. Bullock, and J. Lally. Can feedback solve the too-big-to-fail problem? *Monthly Notices of the Royal Astronomical Society*, 433(4):3539–3546, 2013. doi: 10.1093/mnras/stt984. URL <http://mnras.oxfordjournals.org/content/433/4/3539.abstract>.
- M. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz. Global fit to three neutrino mixing: critical look at present precision. *JHEP*, 1212:123, 2012. doi: 10.1007/JHEP12(2012)123.
- T. Hamana. Raytrix. <http://th.nao.ac.jp/MEMBER/hamanatk/RAYTRIX/>, 2010. [Online; accessed 29-June-2014].
- S. Hannestad, T. Haugbolle, and C. Schultz. Neutrinos in Non-linear Structure Formation - a Simple SPH Approach. *JCAP*, 1202:045, 2012.
- S. Hilbert, J. Hartlap, S. White, and P. Schneider. Ray-tracing through the Millennium Simulation: Born corrections and lens-lens coupling in cosmic shear and galaxy-galaxy lensing. 2008.
- M. Ibison. On the conformal forms of the Robertson-Walker metric. *J.Math.Phys.*, 48: 122501, 2007. doi: 10.1063/1.2815811.
- N. Kaiser and G. Squires. Mapping the dark matter with weak gravitational lensing. *Astrophys.J.*, 404:441–450, 1993. doi: 10.1086/172297.
- X. Kang, A. V. Macciò, and A. A. Dutton. The effect of warm dark matter on galaxy properties: Constraints from the stellar mass function and the tully-fisher relation. *The Astrophysical Journal*, 767(1):22, 2013. URL <http://stacks.iop.org/0004-637X/767/i=1/a=22>.

- H. Katz and M. Ricotti. Two Epochs of Globular Cluster Formation from Deep Fields Luminosity Functions: Implications for Reionization and the Milky Way Satellites. 2012.
- A. Klypin, S. Gottlöber, A. V. Kravtsov, and A. M. Khokhlov. Galaxies in n-body simulations: overcoming the overmerging problem. *The Astrophysical Journal*, 516(2): 530, 1999.
- A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada. Where Are the Missing Galactic Satellites? *ApJ*, 522:82–92, Sept. 1999. doi: 10.1086/307643.
- A. Knebe, S. R. Knollmann, S. I. Muldrew, F. R. Pearce, M. A. Aragon-Calvo, Y. Ascasibar, P. S. Behroozi, D. Ceverino, S. Colombi, J. Diemand, K. Dolag, B. L. Falck, P. Fasel, J. Gardner, S. Gottlöber, C.-H. Hsu, F. Iannuzzi, A. Klypin, Z. Lukić, M. Maciejewski, C. McBride, M. C. Neyrinck, S. Planelles, D. Potter, V. Quilis, Y. Rasera, J. I. Read, P. M. Ricker, F. Roy, V. Springel, J. Stadel, G. Stinson, P. M. Sutter, V. Turchaninov, D. Tweed, G. Yepes, and M. Zemp. Haloes gone MAD: The Halo-Finder Comparison Project. *MNRAS*, 415:2293–2318, Aug. 2011. doi: 10.1111/j.1365-2966.2011.18858.x.
- A. Knebe, F. R. Pearce, H. Lux, Y. Ascasibar, P. Behroozi, J. Casado, C. C. Moran, J. Diemand, K. Dolag, R. Dominguez-Tenreiro, P. Elahi, B. Falck, S. Gottlöber, J. Han, A. Klypin, Z. Lukić, M. Maciejewski, C. K. McBride, M. E. Merchán, S. I. Muldrew, M. Neyrinck, J. Onions, S. Planelles, D. Potter, V. Quilis, Y. Rasera, P. M. Ricker, F. Roy, A. N. Ruiz, M. A. Sgró, V. Springel, J. Stadel, P. M. Sutter, D. Tweed, and M. Zemp. Structure finding in cosmological simulations: the state of affairs. *MNRAS*, 435:1618–1658, Oct. 2013. doi: 10.1093/mnras/stt1403.
- S. R. Knollmann and A. Knebe. AHF: Amiga’s Halo Finder. *ApJS*, 182:608–624, June 2009. doi: 10.1088/0067-0049/182/2/608.
- H. Kodama and M. Sasaki. Cosmological perturbation theory. *Progress of Theoretical Physics Supplement*, 78:1–166, 1984. doi: 10.1143/PTPS.78.1. URL <http://ptps.oxfordjournals.org/content/78/1.abstract>.
- E. Kolb and M. Turner. *The Early Universe*. Frontiers in Physics. Addison-Wesley Longman, Incorporated, 1990. ISBN 9780201116038. URL <http://books.google.dk/books?id=CHDFQgAACAAJ>.
- J. Kormendy and K. C. Freeman. Scaling Laws for Dark Matter Halos in Late-Type and Dwarf Spheroidal Galaxies. In S. Ryder, D. Pisano, M. Walker, and K. Freeman, editors, *Dark Matter in Galaxies*, volume 220 of *IAU Symposium*, page 377, July 2004.

- H. Kragh. *Cosmology and Controversy: The Historical Development of Two Theories of the Universe*. Princeton University Press, 1999. ISBN 9780691005461. URL <http://books.google.dk/books?id=eq7TfxZ0zSEC>.
- M. Kuhlen, M. Vogelsberger, and R. Angulo. Numerical simulations of the dark universe: State of the art and the next decade. *Physics of the Dark Universe*, 1(1-2):50 – 93, 2012. ISSN 2212-6864. doi: <http://dx.doi.org/10.1016/j.dark.2012.10.002>. URL <http://www.sciencedirect.com/science/article/pii/S2212686412000064>. Next Decade in Dark Matter and Dark Energy Next Decade in Dark Matter and Dark Energy.
- M. Kunz and D. Sapone. Crossing the Phantom Divide. *Phys.Rev.*, D74:123503, 2006. doi: [10.1103/PhysRevD.74.123503](https://doi.org/10.1103/PhysRevD.74.123503).
- A. Kusenko. Sterile neutrinos: The Dark side of the light fermions. *Phys.Rept.*, 481: 1–28, 2009. doi: [10.1016/j.physrep.2009.07.004](https://doi.org/10.1016/j.physrep.2009.07.004).
- J. Lesgourgues. Inflationary cosmology. [http://lesgourg.web.cern.ch/lesgourg/Inflation\\_EPFL.pdf](http://lesgourg.web.cern.ch/lesgourg/Inflation_EPFL.pdf), 2006. [Online; accessed 10-June-2014].
- J. Lesgourgues. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. *ArXiv e-prints*, Apr. 2011.
- J. Lesgourgues and S. Pastor. Massive neutrinos and cosmology. *Phys.Rept.*, 429:307–379, 2006. doi: [10.1016/j.physrep.2006.04.001](https://doi.org/10.1016/j.physrep.2006.04.001).
- A. Lewis, A. Challinor, and A. Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *Astrophys.J.*, 538:473–476, 2000. doi: [10.1086/309179](https://doi.org/10.1086/309179). 4 pages, 1 figure. Corrected/updated references. Accepted by ApJ. For the F90 source code see <http://www.mrao.cam.ac.uk/~aml1005/cmb/> Journal-ref: *Astrophys.J.* 538 (2000) 473.
- B. Li, L. J. King, G.-B. Zhao, and H. Zhao. A Semi-analytic Ray-tracing Algorithm for Weak Lensing. 2010.
- A. Liddle and D. Lyth. *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, 2000. ISBN 9780521575980. URL <http://books.google.dk/books?id=XmWauPZSovMC>.
- E. Lifshitz. On the Gravitational stability of the expanding universe. *J.Phys.(USSR)*, 10:116, 1946.
- E. V. Linder and R. J. Scherrer. Aetherizing lambda: Barotropic fluids as dark energy. *Phys. Rev. D*, 80:023008, Jul 2009. doi: [10.1103/PhysRevD.80.023008](https://doi.org/10.1103/PhysRevD.80.023008). URL <http://link.aps.org/doi/10.1103/PhysRevD.80.023008>.

- M. R. Lovell, V. Eke, C. S. Frenk, L. Gao, A. Jenkins, et al. The Haloes of Bright Satellite Galaxies in a Warm Dark Matter Universe. *Mon.Not.Roy.Astron.Soc.*, 420:2318–2324, 2012. doi: 10.1111/j.1365-2966.2011.20200.x.
- M. R. Lovell, C. S. Frenk, V. R. Eke, A. Jenkins, L. Gao, et al. The properties of warm dark matter haloes. *Mon.Not.Roy.Astron.Soc.*, 439:300–317, 2014. doi: 10.1093/mnras/stt2431.
- C.-P. Ma and E. Bertschinger. Cosmological perturbation theory in the synchronous and conformal Newtonian gauges. *Astrophys.J.*, 455:7–25, 1995. doi: 10.1086/176550.
- C.-P. Ma and J. N. Fry. Deriving the nonlinear cosmological power spectrum and bispectrum from analytic dark matter halo profiles and mass functions. *The Astrophysical Journal*, 543(2):503, 2000. URL <http://stacks.iop.org/0004-637X/543/i=2/a=503>.
- C. Marinoni, J. Bel, and A. Buzzi. The Scale of Cosmic Isotropy. 2012.
- K. Markovič and M. Viel. Lyman- $\alpha$  Forest and Cosmic Weak Lensing in a Warm Dark Matter Universe. 2013. doi: 10.1017/pasa.2013.43.
- A. L. Melott. Comment on ‘Discreteness Effects in Simulations of Hot/Warm Dark Matter’ by J. Wang and; S.D.M. White. 2007.
- M. J. Mortonson and U. Seljak. A joint analysis of Planck and BICEP2 B modes including dust polarization uncertainty. 2014.
- V. K. Narayanan, D. N. Spergel, R. Dave, and C.-P. Ma. Constraints on the mass of warm dark matter particles and the shape of the linear power spectrum from the Ly $\alpha$  forest. *Astrophys.J.*, 543:L103–L106, 2000.
- P. Peebles. *The Large-scale Structure of the Universe*. Princeton series in physics. Princeton University Press, 1980. ISBN 9780691082400. URL [http://books.google.dk/books?id=0\\_BPaHFtX1YC](http://books.google.dk/books?id=0_BPaHFtX1YC).
- P. J. E. Peebles. *Principles of physical cosmology*. Princeton University Press, 1993.
- P. J. E. Peebles. The void phenomenon. *The Astrophysical Journal*, 557(2):495, 2001. URL <http://stacks.iop.org/0004-637X/557/i=2/a=495>.
- E. Polisensky and M. Ricotti. Constraints on the dark matter particle mass from the number of milky way satellites. *Phys. Rev. D*, 83:043506, Feb 2011. doi: 10.1103/PhysRevD.83.043506. URL <http://link.aps.org/doi/10.1103/PhysRevD.83.043506>.

- E. Polisensky and M. Ricotti. Massive milky way satellites in cold and warm dark matter: dependence on cosmology. *Monthly Notices of the Royal Astronomical Society*, 437(3):2922–2931, 2014. doi: 10.1093/mnras/stt2105. URL <http://mnras.oxfordjournals.org/content/437/3/2922.abstract>.
- W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, Feb. 1974. doi: 10.1086/152650.
- T. R. Quinn, N. Katz, J. Stadel, and G. Lake. Time stepping N body simulations. *Astrophys.J.*, 1997.
- A. Ringwald and Y. Y. Wong. Gravitational clustering of relic neutrinos and implications for their detection. *JCAP*, 0412:005, 2004. doi: 10.1088/1475-7516/2004/12/005.
- A. Schneider, R. E. Smith, A. V. Macciò, and B. Moore. Non-linear evolution of cosmological structures in warm dark matter models. *Monthly Notices of the Royal Astronomical Society*, 424(1):684–698, 2012. ISSN 1365-2966. doi: 10.1111/j.1365-2966.2012.21252.x. URL <http://dx.doi.org/10.1111/j.1365-2966.2012.21252.x>.
- A. Schneider, D. Anderhalden, A. Maccio, and J. Diemand. Warm dark matter does not do better than cold dark matter in solving small-scale inconsistencies. 2013a.
- A. Schneider, R. E. Smith, and D. Reed. Halo Mass Function and the Free Streaming Scale. 2013b. URL <http://inspirehep.net/record/1222433>.
- P. Schneider. Weak gravitational lensing. 2005.
- C. Schultz, J. Oñorbe, K. N. Abazajian, and J. S. Bullock. The high-z universe confronts warm dark matter: Galaxy counts, reionization and the nature of dark matter. *Monthly Notices of the Royal Astronomical Society*, 442(2):1597–1609, 2014. doi: 10.1093/mnras/stu976. URL <http://mnras.oxfordjournals.org/content/442/2/1597.abstract>.
- S. Seitz, P. Schneider, and J. Ehlers. Light propagation in arbitrary space-times and the gravitational lens approximation. *Class.Quant.Grav.*, 11:2345–2374, 1994. doi: 10.1088/0264-9381/11/9/016.
- U. c. v. Seljak, A. Makarov, P. McDonald, S. F. Anderson, N. A. Bahcall, J. Brinkmann, S. Burles, R. Cen, M. Doi, J. E. Gunn, i. c. v. Ivezić, S. Kent, J. Loveday, R. H. Lupton, J. A. Munn, R. C. Nichol, J. P. Ostriker, D. J. Schlegel, D. P. Schneider, M. Tegmark, D. E. V. Berk, D. H. Weinberg, and D. G. York. Cosmological parameter analysis including sdss  $ly\alpha$  forest and galaxy bias: Constraints on the primordial spectrum

- of fluctuations, neutrino mass, and dark energy. *Phys. Rev. D*, 71:103515, May 2005. doi: 10.1103/PhysRevD.71.103515. URL <http://link.aps.org/doi/10.1103/PhysRevD.71.103515>.
- S. Shao, L. Gao, T. Theuns, and C. S. Frenk. The phase-space density of fermionic dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 430(3):2346–2357, 2013. doi: 10.1093/mnras/stt053. URL <http://mnras.oxfordjournals.org/content/430/3/2346.abstract>.
- M. Shoji and E. Komatsu. Massive Neutrinos in Cosmology: Analytic Solutions and Fluid Approximation. *Phys.Rev.*, D81:123516, 2010. doi: 10.1103/PhysRevD.81.123516,10.1103/PhysRevD.82.089901,10.1103/PhysRevD.81.123516,10.1103/PhysRevD.82.089901.
- R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman. Stable clustering, the halo model and non-linear cosmological power spectra. *Monthly Notices of the Royal Astronomical Society*, 341(4):1311–1332, 2003. doi: 10.1046/j.1365-8711.2003.06503.x. URL <http://mnras.oxfordjournals.org/content/341/4/1311.abstract>.
- V. Springel. The Cosmological simulation code GADGET-2. *Mon.Not.Roy.Astron.Soc.*, 364:1105–1134, 2005. doi: 10.1111/j.1365-2966.2005.09655.x.
- V. Springel. E pur si muove: Galileian-invariant cosmological hydrodynamical simulations on a moving mesh. 2009.
- R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri. Revising the Halofit Model for the Nonlinear Matter Power Spectrum. *ApJ*, 761:152, Dec. 2012. doi: 10.1088/0004-637X/761/2/152.
- R. Teyssier, S. Pires, S. Prunet, D. Aubert, C. Pichon, et al. Full-Sky Weak Lensing Simulation with 70 Billion Particles. 2008.
- J. L. Tinker and C. Conroy. The Void Phenomenon Explained. *Astrophys.J.*, 691:633–639, 2009. doi: 10.1088/0004-637X/691/1/633.
- S. Tremaine and J. E. Gunn. Dynamical role of light neutral leptons in cosmology. *Phys. Rev. Lett.*, 42:407–410, Feb 1979. doi: 10.1103/PhysRevLett.42.407. URL <http://link.aps.org/doi/10.1103/PhysRevLett.42.407>.
- M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese, and A. Riotto. Constraining warm dark matter candidates including sterile neutrinos and light gravitinos with WMAP and the Lyman- $\alpha$  forest. *Phys.Rev.*, D71:063534, 2005. doi: 10.1103/PhysRevD.71.063534.

- M. Viel, M. G. Haehnelt, and V. Springel. The effect of neutrinos on the matter distribution as probed by the Intergalactic Medium. *JCAP*, 1006:015, 2010. doi: 10.1088/1475-7516/2010/06/015.
- M. Viel, G. D. Becker, J. S. Bolton, and M. G. Haehnelt. Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman- $\alpha$  forest data. *Phys.Rev.*, D88(4):043502, 2013. doi: 10.1103/PhysRevD.88.043502.
- R. M. Wald. *General Relativity*. The University of Chicago Press, first edition, 1984.
- J. Wang and S. D. M. White. Discreteness effects in simulations of hot/warm dark matter. *Monthly Notices of the Royal Astronomical Society*, 380(1):93–103, Aug. 2007. ISSN 00358711. doi: 10.1111/j.1365-2966.2007.12053.x. URL <http://arxiv.org/abs/astro-ph/0702575>.
- J. Wang, C. S. Frenk, J. F. Navarro, L. Gao, and T. Sawala. The missing massive satellites of the milky way. *Monthly Notices of the Royal Astronomical Society*, 424(4): 2715–2721, 2012. doi: 10.1111/j.1365-2966.2012.21357.x. URL <http://mnras.oxfordjournals.org/content/424/4/2715.abstract>.
- D. H. Weinberg, J. S. Bullock, F. Governato, R. K. de Naray, and A. H. G. Peter. Cold dark matter: controversies on small scales. 2013.
- M. J. White and W. Hu. A New algorithm for computing statistics of weak lensing by large scale structure. *Astrophys.J.*, 537:1–11, 2000. doi: 10.1086/309009.