Dark Acoustic Oscillations in Structure Formation: The High Redshift Universe

Sebastian Bohr

Faculty of Physical Sciences
University of Iceland
2021

Dark Acoustic Oscillations in Structure Formation: The High Redshift Universe

Sebastian Bohr

Dissertation submitted in partial fulfillment of a
Philosophiae Doctor degree in Physics

Advisor
Assoc. Prof. Jesús Zavala Franco

PhD Committee
Prof. Páll Jakobsson
Prof. Steen H. Hansen

Opponents
Prof. Céline Bœhm
Asst. Prof. Aurel Schneider

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, July 2021

Dark Acoustic Oscillations in Structure Formation: The High Redshift Universe
Dissertation submitted in partial fulfillment of a *Philosophiae Doctor* degree in Physics

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Dunhagi 5
107, Reykjavik
Iceland

Telephone: 525-4000

Bibliographic information:
Sebastian Bohr, 2021, *Dark Acoustic Oscillations in Structure Formation: The High Redshift Universe*, PhD dissertation, Faculty of Physical Sciences, University of Iceland, 154 pp.

ISBN 978-9935-9564-6-0

Printing: Háskólaprent
Reykjavik, Iceland, July 2021

Abstract

The nature of dark matter (DM) as a particle is still an unresolved mystery in Physics. Therefore, a vast amount of competing particle models have been proposed. A promising category among these models are those that include relevant collisional damping in the primordial power spectrum due to the interactions with relativistic particles in the early Universe. This damping is reflected in the DM distribution as dark acoustic oscillations (DAOs) before the onset of structure formation (analogous to the baryonic acoustic oscillations in the photon-baryon plasma, but at smaller, galactic scales) which are potentially observable.

In this Thesis, two effective parameters are proposed that fully describe DAO models based on their key features in the linear power spectrum: the amplitude/height (relative to the Cold Dark Matter expectations) and scale of their primary DAO peak (effectively setting the cut-off scale for structure formation). In the limit of a peak height of zero, this parametrization also includes warm dark matter (WDM), which has a very different particle origin with a collisionless damping and a featureless cut-off in the power spectrum. A large suite of tailored N-body zoom-in simulations is used to cover the DAO parameter space that is still unconstrained, but relevant for galaxy formation. A novel (scale-dependent) way to compare different structure formation is introduced that makes it possible to identify the regions of distinct non-linear structure formation at high redshifts $z \gtrsim 5$ based on statistical measures such as the non-linear power spectrum and the halo mass function. It is found that for a large part of the DAO parameter space, the non-linear power spectrum is actually indistinguishable from WDM models and only a small region of the models with the strongest DAOs has a distinct power spectrum (this region shrinks as the redshift becomes lower). However, the halo mass function breaks this WDM-DAO degeneracy and even the weakest DAO models show a distinct slope in the halo mass function for low-mass haloes, as long as the DAO scale is large enough. With these results, the proposed parametrization offers a quick way to connect a specific DM particle model to its linear power spectrum and from there (using the suite of simulations performed in this Thesis) to the non-linear power spectrum and halo mass function. It is also shown that the properties of DAO haloes can be well described by the extended Press-Schechter (EPS) formalism using a smooth-k filter. On the other hand, the structure of haloes within the DAO cosmology is well described by the well-known Navarro-Frenk-White profile (widely used in Cold Dark Matter, CDM). Relative to CDM, low-mass haloes in DAO models have a a lower concentration, which is also well approximated by the concentration-mass relation predicted by the EPS model and a simple mass assembly model based on hierarchical structure formation.

These results can be used to perform inexpensive calculations of the (high-redshift) halo mass function and concentration-mass relation instead of computationally expensive N-body simulations for virtually all the DAO parameter space explored in this Thesis. Finally, we show that truly distinct strong DAO features can potentially survive in the 1D Flux power spectrum down to redshifts probed by the Lyman-α forest ($z = 3 - 5.4$) and upcoming 21-cm observations at the cosmic dawn ($z = 10 - 25$). Future dedicated simulations including baryonic physics within the template provided in this Thesis should be able to give a detailed prediction for these possible observational signatures, to be searched for in future observations.

Útdráttur

Eðli hulduefnis er enn óleyst ráðgáta innan eðlisfræðinnar og því hefur verið stungið upp á ógrynni líkana til að lýsa því. Eitt þeirra sem lofar góðu, inniheldur hulduefni sem víxlverkar við afstæðilegar eindir í hinum unga alheimi. Þessi víxlverkun endurspeglast í dreifingu hulduefnis sem þrýstingssveiflur (e. DAO) áður en uppbygging alheimsins hefst og er hugsanlega mælanleg.

Í þessari ritgerð eru kynntar tvær breytur sem lýsa að fullu DAO líkönum út frá sérkennum þeirra í aflrófinu: útslagi og hæð megintopps DAO. Sérsniðin hermilíkön eru síðan nýtt til að rannsaka breyturými DAO sem tengist myndun vetrarbrauta. Kynnt er ný aðferð til að bera saman mismunandi formgerðir alheimsins. Hún gerir það mögulegt að bera kennsl á svæði með ólínulegri uppbyggingu formgerðar við há rauðvik ($z \gtrsim 5$); er þetta byggt á tölfræðilegum mælikvörðum eins og ólínulega aflrófinu og massadreifingarreglu hjúpsins.

Ein af niðurstöðunum er sú að fyrir stóran hluta breyturýmis DAO er ólínulega aflrófið í raun óaðgreinanlegt frá líkönum sem innihalda svokallað volgt hulduefni. Auk þess hefur aðeins lítill hluti af þeim líkönum með kröftugustu DAO, auðgreinanlegt aflróf. Hins vegar brýtur massadreifingregla hjúpsins margfeldnina á milli DAO og volgs hulduefnis. Þessar niðurstöður sýna að hægt er að nota áðurnefndar breytur á skjótan hátt til að tengja ákveðið hulduefnislíkan við línulega aflrófið og þaðan (með því að nota hermilíkönin í þessri ritgerð) við ólínulega aflrófið og massadreifingarreglu hjúpsins.

Einnig er sýnt fram á hægt er að lýsa eiginleikum DAO hjúpa með útvíkkaðri aðferðarfræði Press-Schechter með því að nota slétta k síu. Aftur á móti er formgerð hjúpa innan DAO-heimsfræði lýst vel með hinu þekkta Navarro-Frenk-White sniði (mikið notað í köldu hulduefni). Miðað við kalt hulduefni, er samansöfnun hjúpa með lágan massa minni en í DAO líkönum. Þessar niðurstöður er hægt að nýta til að framkvæma hagkvæma útreikninga, við hátt rauðvik, á massadreifingarreglu hjúpsins og massa-samansöfnunar sambandinu, í stað keyrslu hermilíkana sem kosta mikinn reiknitíma í tölvum. Þetta á við um nánast allt DAO breyturýmið sem kannað er í þessari ritgerð.

Að lokum sýnum við að ákveðnir eiginleikar DAO geta hugsanlega lifað af í 1D aflrófinu, niður í rauðvik sem kannað er með Lyman-α skóginum ($z = 3 - 5.4$) og komandi 21-cm athugunum ($z = 10 - 25$). Væntanleg hermilíkön sem taka eðlisfræði þungeinda með í reikninginn, innan þess ramma sem settur er fram í þessari ritgerð, ættu að geta spáð ítarlega fyrir um væntanlegar mæliniðurstöður sem leitað verður í náinni framtíð.

Table of Contents

Acknowledgments

First of all, I would like to thank my supervisor Jesús for his guidance and support throughout the years of my doctoral studies.

I would also like to thank the Astrophysics group at Háskóli Íslands for all the lunch discussions and talks. Special thanks go to Palli for translating the abstract of this thesis into Icelandic, and to my fellow PhD student Jan for our shared office and experiences.

For making my stay in Iceland more fun, I have to thank the Bandý and hiking groups. I will miss our games and hikes.

Last but not least, I am thankful to my family, especially my parents, for supporting me throughout my studies, and to Barbora for staying in Iceland with me and accompanying me through my PhD studies.

This work was supported by a Grant of Excellence from the Icelandic Research Fund (grant number 173929). The simulations that are part of this Thesis were performed on resources provided by the Icelandic High Performance Computing Centre at the University of Iceland, and the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

Chapter 1

Introduction

Dark matter (DM) is a key ingredient in Cosmology and Astrophysics needed to describe the evolution and current state of the Universe at various scales: from galaxies (~ 1 kpc) to clusters of galaxies (~ 1 Mpc) to the large scale structure of the Universe ($\gtrsim 10$ Mpc). Arguably, the evidence that solidified the DM hypothesis came from the observation of galaxy rotation curves by Rubin et al. (1980). If you only consider the mass from the stars and gas (baryons) that you can observe, the rotational velocity of baryons around the center of a galaxy is expected to be highest in the inner regions and then decrease with the distance to the center (this is because baryons are more concentrated in the center of galaxies, thus, this expectation is analogous to what would be expected in a solar-system-like system where most of the mass is located in the center). However, observations showed that the velocity did not decrease significantly in the outer regions, but remained high. The rotation curve (radial profile) was thus observed to be nearly flat. Stars with a velocity this high could not be gravitationally bound to the galaxy by the mass of the baryons alone. Therefore, the results of Rubin and collaborators in the 1970's and 1980's seemed to confirm the independent results originally found by Zwicky (1933) in the 1930's, who analysing the dynamics of galaxies within galaxy clusters concluded that in addition to baryons, there must also exist a vast amount of matter, which does not interact with light, hence, Zwicky named it *dark* matter.

A large number of other independent observations support the DM hypothesis. The following is an incomplete list:

- Baryon acoustic oscillations (BAOs) are density fluctuations that are created through frequent scattering in the baryon-photon fluid in the early Universe, which leave an observed signature in the large scale statistical distribution of galaxies in the late Universe. The scale and amplitude of the observed BAOs require the presence of DM (Cole et al., 2005; Eisenstein et al., 2005; Anderson et al., 2012).

- Gravitational lensing of background sources caused by matter bending space-time around foreground systems along the line of sight is stronger than can be explained by the mass of visible matter alone, requiring additional mass in the form of DM. On large scales, this can be observed as tomographic weak lensing and cosmic shear of millions of galaxies (Hildebrandt et al., 2017). On cluster scales, weak lensing arcs (Hoekstra et al., 1998) and multiple images through strong lensing (Tyson et al., 1998) of background objects require DM.

- Hot gas is the dominant form of baryonic matter in galaxy clusters (directly observable in X-ray emission). The measured distribution of hot gas suggests a system near hydrostatic equilibrium containing in addition to baryonic matter (hot gas corona and galaxies) a significant amount of DM (e.g. Vikhlinin et al., 2006).
- The Bullet Cluster is the result of a collision of two galaxy clusters. The dominant X-ray emitting hot gas of each cluster clearly experienced drag during the collision which separated the center of mass of the baryonic matter from the center of mass of the total gravitational matter in the cluster (measured directly through gravitational lensing). The Bullet Cluster can be explained by a vast amount of DM present in the cluster, but being unaffected by the collision (Clowe et al., 2006; Randall et al., 2008).
- Dwarf spheroidals, such as those in the Milky Way, have mass to light ratios inferred from their stellar kinematics, which are significantly higher than in larger galaxies, suggesting that they are strongly dominated by DM (the random velocities of their constituent stars are too large to be supported by the mass in the stellar component alone; Walker et al. 2009).

One of the strongest and more precise pieces of evidence on the existence of DM is provided by measurements of the cosmic microwave background (CMB; see Planck Collaboration et al. 2020 for the most recent observational results). The CMB is light that was emitted at the time of recombination ($z \sim 1100$; ~ 370000 years after the Big Bang), when the Universe expanded and cooled down enough for protons and electrons to combine into hydrogen and remain stable. While photons frequently scattered through Thomson scattering with free electrons keeping the mean free path short before recombination, after recombination baryonic matter became almost completely neutral and therefore, the mean free path of photons became effectively infinite. As these photons were previously in thermal equilibrium through frequent scattering, they are distributed as black body radiation with the temperature of the photon-baryon plasma at the time of photon decoupling. We observe this radiation redshifted into the microwave range today as the CMB. Although the CMB has a featureless (black body) energy spectrum, it is not completely isotropic due to acoustic oscillations in the baryon-photon fluid before recombination, which are imprinted in the angular power spectrum of the CMB. This phenomenon is briefly described below.

In the tightly coupled baryon-photon fluid, gravity tries to compress the fluid, while photon (radiation) pressure acts against the compression and expands the fluid. The oscillation between compression and expansion phases form pressure waves, which propagate at the speed of sound. When baryons and photons decouple from each other around the time of recombination, radiation pressure can no longer sustain gravity and the oscillations stop, imprinting the rarefaction and compression pattern of the oscillatory stage at recombination into the density fluctuations over the background, which is reflected in the temperature fluctuations observed in the CMB. The maximum

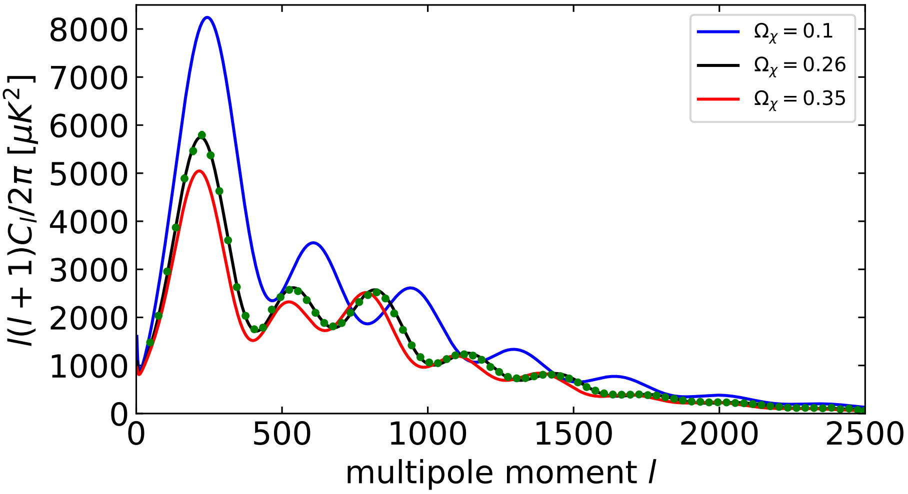

Figure 1.0.1. CMB temperature angular power spectrum showing the imprint of acoustic oscillations in the photon-baryon fluid. The green dots are the latest Planck 2018 measurements and the black line corresponds to the best-fit model, which corresponds to a global abundance of DM given by $\Omega_\chi = 0.26$*. With a smaller amount of DM, the acoustic peaks would be less damped and would follow a common damping envelope (blue line), Both of these features would make the model inconsistent with observations. A higher amount of DM would dampen the peaks below observations, especially the even peaks that correspond to the expansion phase (red line). Additionally, a change on the DM amount also changes the angular scale of the peaks by shifting the time of matter-radiation equality.*

distance a sound wave could travel by recombination is called the sound horizon, and at the angle corresponding to twice the sound horizon a strong correlation of the temperature fluctuations is indeed observed in the CMB, as the first and most prominent peak in the angular power spectrum. At multiples of the corresponding wavenumber/angle, we can also observe the harmonics of the acoustic oscillations in the higher order peaks (see Fig. 1.0.1). As DM also contributes to the gravity well driving the compression phase of the oscillations, a higher amount of DM dampens the peaks in the CMB power spectrum and the ratio of the peaks is highly influenced by the amount of DM compared to the amount of baryons. The additional mass in the gravity well from a higher DM density enhances the compression phase of acoustic oscillations and weakens the expansion phase. Therefore, the even peaks, that correspond to the expansion phase, are damped stronger than the odd peaks. This is most clear in the second and third peak in Fig. 1.0.1.

1.1 The Universe at large scales: the ΛCDM paradigm

All of the evidence mentioned above have in common that baryonic[1] matter alone cannot account for the total matter needed to explain the properties of the observed cosmic structure. In cosmology, the established standard model of structure formation and evolution is known as ΛCDM, which requires general relativity and the standard model of particle physics in an expanding Universe, plus two still unknown additional components: a cosmological constant Λ driving the *accelerated* expansion of the Universe and CDM standing for cold dark matter, which is a non-relativistic and collisionless matter component beyond the standard model of particle physics. The main parameters of the ΛCDM model are $\Omega_\Lambda = 0.6889$, $\Omega_\chi = 0.2619$, $\Omega_\mathrm{b} = 0.0492$, $H_0 = 67.66\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\sigma_8 = 0.8102$, and $n_\mathrm{s} = 0.9665$. The overdensities $\Omega_\mathrm{x} = \rho_\mathrm{x}/\rho_\mathrm{c}$ are measured relative to the critical density of the Universe and correspond to the the cosmological constant, DM, and baryons, respectively. The critical density ρ_c separates positive and negative spatial curvature of the Universe, and the Universe is flat if $\sum \Omega_\mathrm{x} = 1$. The Hubble rate H_0 represents the current rate of expansion of the Universe, σ_8 is the mass variance on $8\,\mathrm{Mpc}\,h^{-1}$scales and gives a normalization for the power spectrum, while the spectral index n_s sets the slope of the primordial power spectrum.

The ΛCDM model is very successful at describing the large scale structure of the Universe. It can explain observations like the CMB and BAOs with high accuracy. In addition, numerical simulations of cosmological structure formation and evolution (see Section 1.2) agree remarkably well with large-scale observations of the cosmic web as given by large galaxy surveys (see Fig. 1.1.2). It is crucial to remark that in the ΛCDM model, dark matter has only three defining characteristics – being cold (non-relativistic), dark (collisionless), and classical (non-quantum) matter – that are unchallenged, *but only on the very large* scales relevant for cosmology. However, the validity of these DM characteristics on smaller (galactic) scales remains debatable. A DM nature distinct to CDM is a viable possibility, which is completely compatible with the ΛCDM model, despite of what the name seems to imply.

---

[1] Note that baryons in cosmology and astrophysics include everything consisting of standard model particles (baryons, mesons, leptons), of which baryons make up the majority by mass

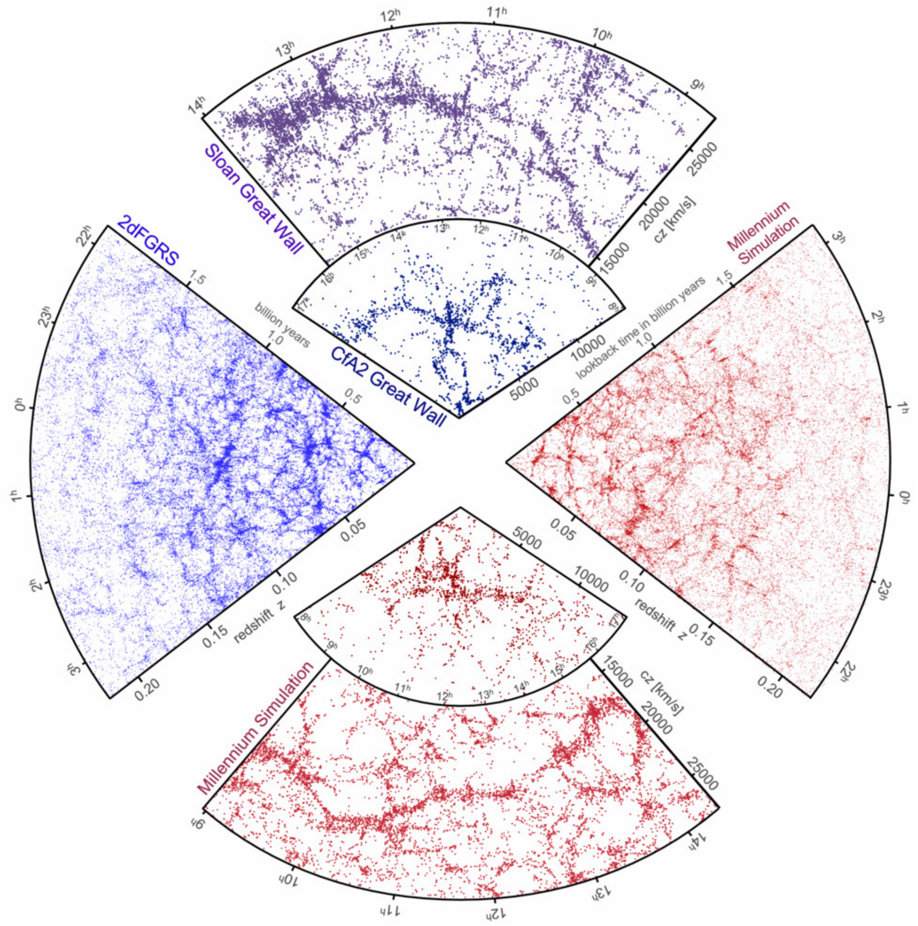

Figure 1.1.2. *The distribution of galaxies from observations and simulations (Fig. 1 of Springel et al. 2006). The top slice shows the observations of the CfA Great Wall (Geller & Huchra, 1989) and Sloan Great Wall (Gott et al., 2005), two largest structures in the local Universe. The left slice shows one half of the 2dFGRS survey (Colless et al., 2001). The bottom and right slices show similar structures obtained from mock surveys of the Millennium simulation (Springel et al., 2005) within the ΛCDM cosmology. (Reprinted by permission from Springer Customer Service Centre GmbH: Nature; The large-scale structure of the Universe. Springel V., Frenk C. & White S, ©2006)*

1.2 The non-linear regime of structure formation: N-body simulations

Structure formation theory can be divided into two regimes depending on the strength of density perturbations. The perturbations on the large scales probed by observations of the CMB and BAOs discussed above are still small compared to the average density. Therefore, they can be described using linear perturbation theory (see also Section 1.5). When the perturbations become too large however, linear perturbation theory breaks down and thus a different approach is needed. Higher order perturbation theory can be used for the quasi-linear regime (e.g. Carrasco et al. 2012), while simplified analytical models can be used to predict from the properties of the linear density field (power spectrum) some of the statistical properties of the end state of non-linear evolution: dark matter haloes (e.g. spherical collapse, Gunn & Gott 1972, or ellipsoidal collapse, Sheth et al. 2001). However, the most general and powerful approach are N-body simulations, which in essence compute the dynamical evolution of a sample of N particles by computing the gravitational forces for all particles at each timestep in the simulation.

For collisionless DM particles, the system is fully described statistically by a phase-space distribution function $f(\vec{x},\vec{\mathrm{v}},t)$, whose evolution is given by the Vlasov-Poisson equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + v_i \frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial f}{\partial \mathrm{v}_i} = 0, \tag{1.1}$$

$$\rho_\chi(\vec{x},t) = \int f(\vec{x},\vec{\mathrm{v}},t)d^3\vec{\mathrm{v}}, \tag{1.2}$$

$$\Delta\Phi(\vec{x}) = 4\pi G\rho_\chi(\vec{x}), \tag{1.3}$$

where ρ_χ is the DM density field, Φ is the gravitational field, and G is the gravitational constant. In the case of N-body simulations, the phase-space distribution is discretized by a set of N (macro) particles with a mass that is many orders of magnitude larger than the actual (micro) DM particles. Thus, the system of N particles provides a statistical representation of the true system averaged at the scales probed (resolved) by the macro particles whose evolution is given by a discretized form of Eqs. (1.1)-(1.3):

$$\tilde{f}(\vec{x},\vec{\mathrm{v}}) = \sum_i m_i W(|\vec{x}-\vec{x}_i|,\varepsilon)\delta^3(\vec{\mathrm{v}}-\vec{\mathrm{v}}_i), \tag{1.4}$$

$$\tilde{\rho}_\chi(\vec{x}) = \sum_i m_i W(|\vec{x}-\vec{x}_i|,\varepsilon), \tag{1.5}$$

$$\tilde{\Phi} = \int g(\vec{x}-\vec{x}')\tilde{\rho}_\chi(\vec{x}')d^3\vec{x}', \tag{1.6}$$

where m_i is the mass of the (macro) particle i, δ^3 is the three-dimensional Dirac delta function, W is a softening kernel with softening length ε, and g is a Green's function for the Poisson equation. The softening kernel smoothes the point-like masses of N discrete particles effectively extending each particle to represent a finite volume. This removes the gravitational singularities resulting from point-like density spikes, which would require tiny time-steps in the numerical integration of close encounters due to a diverging force (see e.g. Dehnen & Read, 2011). Additionally, the softening kernel suppresses two body interactions, which would be artificial as every simulation particle represents a large amount of DM particles. Cosmological N-body simulations use a comoving reference frame for the N particles and include the expansion of the Universe in the scale factor, which is obtained by solving the Friedmann equations.

N-body method. Solving the Boltzmann equation directly in phase space for a large number of particles in a full cosmological setting is computationally impractical. Instead, N-body simulations rely on solving the equations of motion for each of the N (macro) particles directly. The simplest approach to solve the dynamics of the N-body system is numerically integrating the N^2 gravitational interactions, but this approach is highly inefficient and too computational expensive for a reasonably large N (required by cosmic structures). Another method is the particle-mesh (PM) technique, which first creates a density mesh from the N particles. A Fourier transform of the density is then performed in order to solve the Poisson equation (Eq. 1.3) for the potential Φ in Fourier space by simple multiplication. Finally, the potential and force at the location of each particle can be found after an inverse Fourier transform and an accurate interpolation (Klypin & Shandarin, 1983). This method is however limited by the mesh size restricting the effective resolution. A third algorithm is the hierarchical tree method, which divides the simulation volume recursively into cubes in a tree structure (Barnes & Hut, 1986). For distant particles, the tree structure does not get resolved fully and all particles within the respective cube are grouped together to act as a single bigger (macro) particle for the gravitational force, instead of computing each individual force pair. For nearby particles, the tree gets fully resolved and each force pair is computed exactly. Modern N-body codes actually use a combination of these methods to overcome their individual shortcomings. For example, the code AREPO (Springel, 2010) used in this Thesis uses the treePM algorithm, which utilizes the tree method for short-range interactions and the PM method to compute the long-range forces in Fourier space (this hybrid approach is based on the widely known N-body code GADGET, Springel 2005). *It is important to remark that if different DM models have no other interactions besides gravitation that are relevant during the non-linear evolution, it is sufficient to only change the initial conditions for the N-body code, the rest of the evolution is treated in the same way as in the CDM case. This is precisely the type of models studied in this Thesis.*

Initial conditions. For cosmological simulations, the initial conditions of the N particle system have to be a statistical realization of the DM phase space distribution as given by the linear regime of cosmological structure formation, i.e. the N-body approach should start before linear perturbation theory can no longer be trusted. For Gaussian random fields, both the density and velocity fields in the linear regime are given by the linear power spectrum $P(k)$, where k corresponds to a mode with wavenumber k; this quantity contains (statistically) all the information on the phase space clustering of DM. The power spectrum $P(k)$ is the Fourier transform of the two-point correlation function $\xi(r)$, which in the limit $r \to 0$ reduces to the variance of the density field. The power spectrum is often written as the dimensionless power spectrum $\Delta^2 = k^3 P(k)/(2\pi^2)$, which gives the contribution to the variance of the density field per bin of $\ln k$. For the linear regime, Fourier space is especially useful, as each k mode can be treated independently from each other. The N-body code can therefore also treat the simulated volume as a periodic box if its size is sufficiently large to be considered linear over the full simulation time (cosmological principle).

The procedure to construct a statistical realization of the linear density field is as follows. First, all particles are homogeneously distributed in the simulation box. In the next step, perturbations of wavenumber k are introduced according to a Gaussian distribution whose variance is determined by the linear power spectrum. The corresponding real space density field is then used to compute displacements for the N particles with second-order Lagrangian perturbation theory (Jenkins, 2010). After these displacements, the particles are distributed (statistically) according to the input power spectrum (see top left and right of Fig. 1.2.3). The initial conditions used in this Thesis were constructed with the code MUSIC (Hahn & Abel, 2011).

The N-body approach is limited in two ways. First, the size of the simulation box, beyond which periodicity is assumed, sets the lower limit for wavenumbers that can be probed to $k_{\rm min} = 2\pi/L$ (left vertical line in the top left of Fig. 1.2.3), where L is the side length of the simulation volume. L has to be chosen large enough, so that the minimum wavenumber remains mostly linear over the course of the simulation. Otherwise, the simulation would miss power transferred from larger scales to small scales, i.e., it would lack the gravitational coupling of the clustering on the largest scales (smallest wavenumbers) to the smallest scales, which is relevant for the non-linear evolution of the small scales. Second, the spatial and mass resolution given by the number N of (macro) particles determines the maximum wavenumber, roughly given by the Nyquist frequency $k_{\rm max} \sim k_{\rm Ny} = \pi/d$ (right vertical line in top left of Fig. 1.2.3), where d is the interparticle separation, and the minimum mass scale that can be probed $m_{\rm part} = \rho_\chi L^3/N$. At lower redshifts, Poisson noise (due to the finite discretization) is introduced into the particle distribution and sets a minimum level of (artificial/numerical) irreducible power $\Delta^2 = (kL)^3/(2\pi^2 N)$ (dashed line in center left of Fig. 1.2.3). For a given simulation size L, the number of particles will determine the softening length for the softening kernel W in Eq. (1.4)-(1.5) and with that the smallest scales that can be studied. The softening length is chosen as a fraction of the interparticle separation, usually between 1/100 and 1/10 – a common choice which will be used in later chapters is $\varepsilon = L/(45N^{1/3})$.

DM haloes. The small DM over-densities that are present in the initial conditions grow over time and eventually disconnect from the expansion of the Universe, when they collapse due to gravity into self-bound structures called haloes. These haloes are the (non-linear) virialized final state of evolution of the primordial density perturbations seen in the CMB. After their initial "formation" (i.e. when they first become virialized structures), haloes continue growing and increasing their mass over time in two main modes: continuously by accreting particles from its surroundings and abruptly by merging with other haloes. The smallest (least massive) merging haloes become subhaloes of the host halo, and are gravitationally bound to it. The exact mass and size of a halo is rather ambiguous given the difficulty of properly defining the boundary of a halo. A common choice is to approximate the halo as a sphere of radius R_{200} within which the average density is given by 200 times the critical density ρ_c. This choice is inspired by the (analytical) spherical collapse model, which predicts that the overdensity of a spherical region that collapses and virializes in an Einstein$-$de Sitter Universe is ~ 180 times the critical density. The corresponding virial mass is then given as $M_{200} = 4\pi/3\rho_c R_{200}^3$. Unless otherwise stated, this is the halo mass used in this Thesis.

Halo mass function. The simplest statistical property of haloes is their abundance as a function of mass, which is called the halo mass function (see bottom left of Fig. 1.2.3). The CDM model predicts a hierarchical formation of DM haloes, with low mass haloes forming first and more massive haloes later, mostly through the merger of

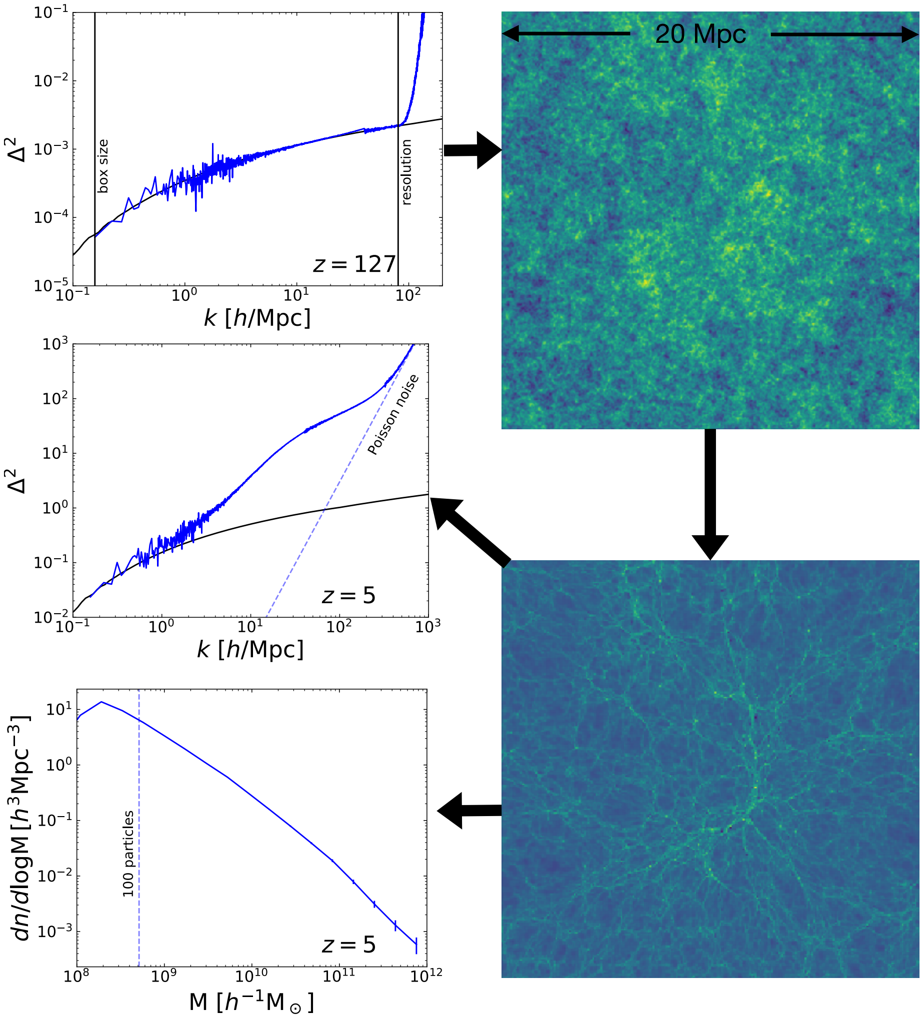

Figure 1.2.3. Cosmological simulations in a nutshell. An initial conditions code is used that takes the linear power spectrum as an input (in this case at $z = 127$; black curve in top left) to create a discretized realization of N particles which contains the statistical properties of the input spectrum (top right; colour scale is exaggerated for better visibility). The power spectrum computed from the N particle realization (blue line in top left) follows the input power spectrum very closely between the limiting wavenumbers set by the box size (beyond which the simulation is periodic) and the resolution limit. At low k, the reconstructed power spectrum shows some noise due to the limited number of modes that can be fit into the finite simulation volume, while at high k the simulation is limited by the finite number of particles. The simulation evolves the distribution of the particles down to a lower (non-linear) redshift ($z = 5$; bottom right), at which the DM clusters together into clear structures. From a snapshot like this, statistical properties like the power spectrum (center left; compared to the linear spectrum at the same redshift in black) and halo mass function (bottom left) can be extracted and used to compare varying DM models. The images on the right were created using the Pynbody library (Pontzen et al., 2013).

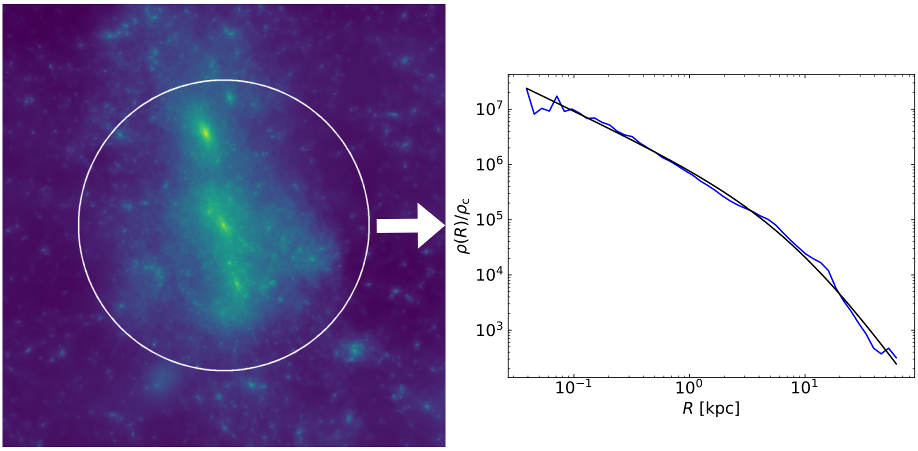

Figure 1.2.4. The DM distribution in a halo. The left panel shows a DM halo and its immediate environment at $z = 5$; the circle indicates the virial radius (boundary) of the halo. The spherically averaged density profile can be extracted from the (simulated) particle data in the snapshot (blue line on the right) and it is well described by a NFW profile (black line on the right, Eq. 1.8). The image on the left was created with the Pynbody library (Pontzen et al., 2013).

small haloes. One of the outcomes of this hierarchical scenario is that the CDM model predicts an ever increasing amount of haloes towards smaller masses, which can be approximated by a power law (see e.g. Boylan-Kolchin et al., 2009)

$$\frac{dn}{dM} \propto M^{-1.9} \tag{1.7}$$

At the high-mass end, the halo mass function eventually reaches a cut-off at the scale at which structures have had the time to become non-linear and collapse into haloes (massive clusters today). The gravitational collapse into haloes is also well described by the extended Press-Schechter (EPS) formalism using spherical or ellipsoidal collapse from which the halo mass function can be computed (Press & Schechter, 1974; Sheth et al., 2001), with quite a good agreement with simulations. This formalism assumes that areas with a density larger than a critical density $\delta_c(t) = 1.686/D(t)$, where $D(t)$ is the linear growth factor, will have collapsed into haloes at a cosmic time t. These models only require the linear power spectrum as an input to compute the mass variance and give a probabilistic estimate of the number of haloes based on the collapse barrier. In Chapter 3, we apply the EPS formalism to DAO models.

Inner halo structure. Despite the complexity of the phenomenon of gravitational clustering in a cosmological setting, the internal structure of DM haloes shows some remarkably simple properties. For instance, the spherically averaged density profile of CDM haloes of all sizes are very well approximated by a common (universal) shape, the so called Navarro-Frenk-White (NFW) profile (Navarro et al., 1996, 1997, see

Fig. 1.2.4)

$$\frac{\rho(r)}{\rho_c} = \frac{\delta_c}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}, \tag{1.8}$$

where δ_c, and the scale radius r_s are free parameters. However, the concentration parameter $c = R_{200}/r_s$ offers a better comparison between haloes, where a low concentration means lower central densities. Requiring that the mass within R_{200} equals M_{200}, δ_c is given by c:

$$\delta_c = \frac{200}{3}\frac{c^3}{\ln(1+c) - c/(1+c)}. \tag{1.9}$$

Furthermore, the concentration of haloes is tightly correlated with mass, with low mass haloes being more concentrated than more massive haloes. The concentration-mass relation can also be understood from the hierarchical process of mass assembly and the extended Press-Schechter formalism, assuming that the mean inner density within the scale radius is proportional to the critical density of the Universe at the assembly redshift of the halo (see e.g. Ludlow et al. 2016). In Chapter 3, we apply the model of Ludlow et al. (2016) to DAO models.

1.3 Challenges to the CDM model at galactic and sub-galactic scales

Besides being part of the standard ΛCDM model of cosmology, CDM is the established ingredient for the theory of galaxy formation and evolution. As mentioned above, its main characteristics are that DM is assumed to be made of classical (with negligible quantum effects) particles that are *cold* and *collisionless*. *Cold* in the broadest sense means that DM is non-relativistic, but in CDM it means more specifically that DM does not have any significant primordial thermal/random velocities that could affect galactic-scale structure formation. In other words, CDM does not experience collisionless damping through the phenomenon of free-streaming. *Collisionless* specifies that DM only interacts gravitationally; it does not interact otherwise with other species or with itself. However, the degree to which DM can be considered cold and/or collisionless is not tightly constrained by observations. As mentioned above, on large scales, CDM successfully explains the structure of the Universe from its early stages as imprinted in the CMB, to the cosmic web formed by the large scale distribution of galaxies today. On small (galactic and sub-galactic) scales however, it faces a number of significant challenges, which could possibly be related to the CDM assumptions. The following are among the most relevant challenges in low mass galaxies (see Bullock & Boylan-Kolchin 2017 for a review):

- *under-abundance of low-mass galaxies*: Although the minimum halo mass for a galaxy to form remains uncertain, at the mass scales corresponding to (low-mass) dwarf galaxies, CDM predicts a vast abundance of low-mass haloes (see Eq. 1.7), each of which could potentially host a galaxy, while the observed number of dwarf galaxies is way below the naively expected number in CDM. For instance, the observed satellites of the Milky Way and M31 have been repeatedly claimed to be too few to be compatible with the predictions from CDM simulations since two decades ago (Klypin et al., 1999; Moore et al., 1999). Similarly, low-mass field galaxies are also seemingly underabundant (Zavala et al., 2009; Papastergis et al., 2011; Klypin et al., 2015; Schneider et al., 2017; Trujillo-Gomez et al., 2018). The under-abundance of satellites is mostly solved by correcting for the detection efficiency of the SDSS survey (see Kim et al., 2018).

- *core-cusp problem*: Although it is challenging to infer the inner DM distribution of haloes due to the incomplete kinematical information we have from gravitational tracers (stars, gas), rotation curves of low surface brightness galaxies consistently show the presence of less dense and more extended DM haloes than naively expected, seemingly implying the presence of constant DM density “cores" in the inner parts of the galaxies (e.g. de Blok & McGaugh, 1997). Stellar kinematics in some dwarf Spheroidals in the Milky Way also show “underdense” possible cored haloes (Walker & Peñarrubia, 2011). In contrast and as mentioned in Section 1.2 above, DM haloes in CDM simulations show a steep power law density profile towards their center (see Eq. 1.8), called a “cusp”. The higher DM densities and cuspy profiles predicted by CDM in low-mass haloes have been challenged by observations of low-mass galaxies.

- *too-big-to-fail problem*: Simulations of Milky Way sized haloes have massive subhaloes that due to their mass should be expected to be the hosts of the satellite galaxies observed in the Milky Way. However, the internal kinematics of most of the MW satellites is inconsistent with inhabiting such subhaloes due to their high DM densities. This challenge known as the “too big to fail” problem, since these subhaloes should be too massive to fail to form stars (Boylan-Kolchin et al., 2011, 2012), has also been identified in the Local Group (Garrison-Kimmel et al., 2014) and in the field (Papastergis et al., 2015).

- *diversity problem*: The rotation curves of simulated dwarf galaxies within CDM show a high degree of similarity between galaxies of the same mass. This is at odds with a large diversity in observed rotation curves, which seemingly implies that galaxies of the same mass can inhabit very cuspy or very cored DM haloes; the latter deviate more strongly from the the simulated rotation curves in CDM (Oman et al., 2015). A similar diversity is observed in the inner DM densities of the MW satellites, which augments the too-big-to-fail problem (Zavala et al., 2019).

There are different approaches to provide solutions to these problems. First, as there are always limitations to observations, the underlying observational data can be checked for incompleteness, biases and systematic uncertainties (see e.g. for the under-abundance problem Koposov et al. 2008; Kim et al. 2018; Chauhan et al. 2019; Dutton et al. 2019; for diversity problem Oman et al. 2019; for too-big-to-fail and core-cusp problem Campbell et al. 2017). Second, uncertain baryonic physics can alleviate some of these problems. For example, impulsive supernova feedback can explain the presence of DM cores if the energy released is large enough (Pontzen & Governato, 2012; Peñarrubia et al., 2012; Di Cintio et al., 2014; Burger & Zavala, 2021), tidal effects from the Milky Way disk can address the too-big-to-fail problem (Garrison-Kimmel et al., 2019), and cosmic reionization can suppress galaxy formation at dwarf galaxy scales (Gnedin, 2000; Sawala et al., 2016b). And third, deviations from the CDM hypothesis through additional DM physics can also alleviate the small scale issues; which is the approach focused on in this Thesis.

1.4 Particle models

As the nature of DM still remains a mystery, a wide variety of different DM models has been proposed and studied (for a recent review, see e.g. Buckley & Peter, 2018). Particle models for CDM include axions, a new elementary particle that could solve the strong CP problem in particle physics (Preskill et al., 1983). Another and perhaps favourite CDM candidate are weakly interacting massive particles (WIMPs), that are just coupled weakly, i.e. through the weak force, with standard model particles and are massive enough to be cold. One of the reasons for their appeal is that for typical weak-scale cross-sections and a particle mass of $\mathscr{O}(100)$ GeV, similar to that of W/Z-bosons, the thermal relic density of symmetric WIMPs, i.e. there is the same amount of WIMPs and anti-WIMPs, matches the observed DM density of the Universe, which is called the “WIMP miracle". A suitable WIMP candidate (the light neutralino) is also predicted by supersymmetric theories, which extend the standard model of particle physics. Although current constraints reject large parts of the possible WIMP parameter space, there are still significant regions that remain to be explored (see e.g. Arcadi et al., 2018; Roszkowski et al., 2018, for an overview of the WIMP paradigm and current constraints on WIMPs).

Among the many possible classifications of alternative DM models, one could focus on their impact on cosmological structure formation at scales that are relevant for galaxy formation and evolution. In particular, there are two main DM-physics effects on the linear matter power spectrum: collisionless and collisional damping. Both of these effects suppress the power on small scales, but the physical mechanism and resulting power spectrum is rather different. In the case of collisionless damping, the requirement of DM being cold gets relaxed and DM is then generically called warm dark matter (WDM). This type of DM has a non-negligible velocity in the early Universe leading to a collisionless damping in the linear power spectrum caused by the free-streaming mechanism with a characteristic length set by the random motions of the DM particles, which flow from over- to under-dense regions. On scales up to the free-streaming length, structure formation is highly suppressed in the early universe, while on larger scales WDM behaves identical to CDM. Therefore, WDM with a galactic-scale free-streaming length can alleviate the problems of CDM at the scale of low-mass galaxies, while keeping the success of CDM on large scales. A possible WDM candidate are sterile neutrinos, which are hypothetical right-handed neutrinos that can interact with the active neutrinos of the standard model, but that are effectively collisionless in structure formation. Sterile neutrinos are additionally attractive since they offer a mechanism to explain the small masses of the standard model neutrinos (for a recent review on sterile neutrions see Boyarsky et al., 2019). Another WDM candidate is the gravitino, a supersymmetric partner of the graviton from supergravity theories, which is produced from the decay of heavier superpartners (see e.g. Pagels & Primack, 1982; Feng et al., 2010). Note that WIMPs actually have a free-streaming scale of order of Earth mass (Bringmann, 2009). However, as this is well below the scales relevant for galaxy formation, they are not considered WDM but CDM.

In contrast, collisional damping in the early Universe occurs when DM is strongly coupled with a relativistic species through additional interactions besides gravity. At early times, pressure waves can then propagate through the tightly coupled DM-radiation fluid, which introduces dark acoustic oscillations (DAO) similar to BAOs in the early universe. These DAOs get imprinted in the distribution of matter as oscillations and a suppression of power (Silk-like damping) on small scales with an amplitude that depends on the scattering rate. The small scale suppression below the cut-off scale is weaker than in the WDM (free streaming) case, as the oscillations can even reach the power of CDM. For the potential additional interactions driving this mechanism, there are two types of models. First, the interactions can be between DM and standard model particles like photons and/or neutrinos (Bœhm et al., 2002; Bœhm & Schaeffer, 2005). Second, the interactions can be part of a more complex dark sector decoupled from ordinary matter with more additional dark/hidden particles. Possible scenarios are interactions with massless sterile neutrinos through a massive vector boson (van den Aarssen et al., 2012), atomic DM, which is a dark equivalent of the proton-electron-photon interactions of the standard model including a dark recombination (Cyr-Racine & Sigurdson, 2013), or non-Abelian DM and dark radiation (Buen-Abad et al., 2015).

There are other physical mechanism to produce a galactic-scale cutoff in the power spectrum, e.g. ultra-light bosons with a de Broglie wavelength of the order of 1 kpc, which in addition introduce quantum effects on galactic scales. This DM model is called fuzzy dark matter (e.g. Hui et al., 2017). In this Thesis, we will only consider the two damping mechanisms described in the previous paragraph, generically, WDM and DAOs.

The damping of small scale power in the linear power spectrum survives into the non-linear regime leading to a clear suppression of the non-linear power spectrum compared to CDM, but to a weaker degree than in the linear power spectrum. Additionally, it also delays the formation of haloes and leads to i) fewer small mass haloes (potentially explaining the dearth of low-mass galaxies, see e.g. Bode et al. 2001 for WDM), and ii) lower inner DM densities (see e.g. Avila-Reese et al., 2001, for WDM). Recent high-resolution simulations for WDM have shown that haloes within this cosmology still follow a NFW profile (Eq. 1.8), but the concentration of small mass haloes is lower

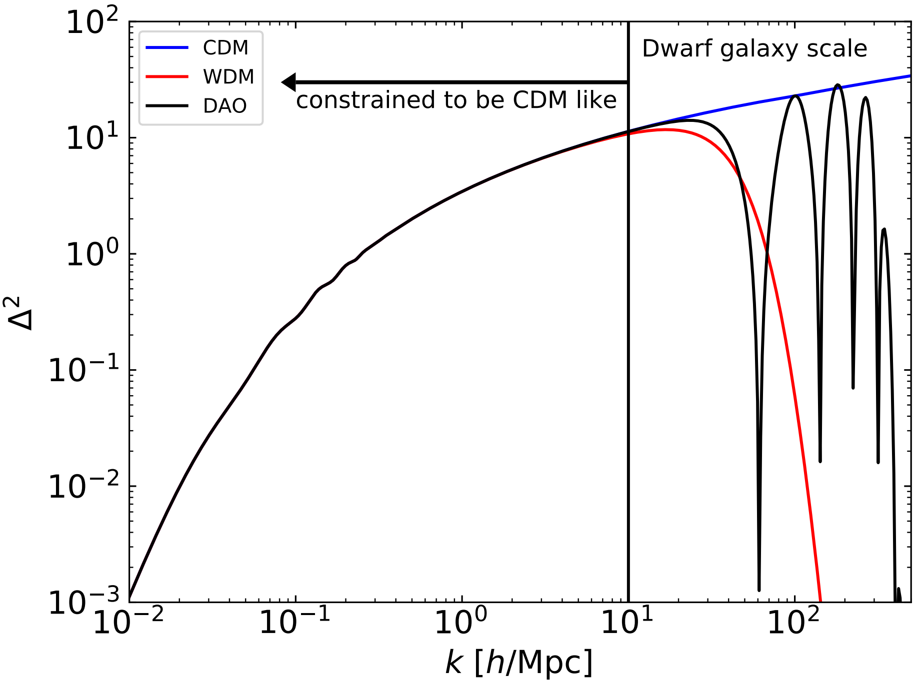

Figure 1.4.5. Dimensionless linear matter power spectrum of different DM models. Observations of the distribution of galaxies and the Lyman-α forest constrain DM to behave like CDM below $10\,h\,Mpc^{-1}$. However, above $10\,h\,Mpc^{-1}$the linear matter power spectrum remains unconstrained. In the case of CDM, the power spectrum keeps rising at larger k. However, WDM and DAO models introduce a cut-off at smaller scales and in the case of DAOs also subsequent oscillations.

than for CDM (see e.g. Lovell et al., 2014; Ludlow et al., 2016). Simulations with models with DAOs are scarce, but recent examples show similar features as that in WDM with the suppression of small scale power leading to less low-mass haloes and to less dense halo centers (e.g. Buckley et al., 2014; Vogelsberger et al., 2016).

Figure 1.4.5 shows the effects of the different classes of DM models on the linear power spectrum. While the power of CDM (blue line) keeps rising towards high k, the collisionless damping of WDM (red line) erases all power at small scales and introduces a sharp cut-off at a k corresponding to the free-streaming scale. The collisional damping of DAOs (black line) also leads to a suppression of the small scale power, but in an oscillatory pattern due to the acoustic nature of the damping which is clearly visible in the power spectrum. Observations of the Lyman-α forest (see also Sec. 1.6) have constrained the linear power spectrum to be CDM-like below $k \sim 10\,h\,\mathrm{Mpc}^{-1}$(e.g. Viel et al., 2013; Iršič et al., 2017). The different possibilities for DM below this scale and their consequences on non-linear quantities are the focus of this Thesis.

1.5 The ETHOS framework in the linear regime

The effective theory of structure formation (ETHOS) was introduced in Cyr-Racine et al. (2016) to offer a framework to explore cosmological structure formation in a more general way than CDM, encompassing different DM models having DM physics that deviate in both the linear and non-linear regimes of structure formation. In particular for the interest of this Thesis, ETHOS incorporates a variety of models with a cut-off in the linear power spectrum, as explained in Sec. 1.4 and shown in Fig. 1.4.5, that can have a significant impact in the low-end of the relevant scales for galaxy formation and evolution. The original objective of ETHOS in Cyr-Racine et al. (2016) was to extend the equations of cosmological perturbation theory with a generalized and effective parametric model to study particle models with DAOs. Instead of implementing each particle DM model separately into a Boltzmann solver, they can be grouped into classes according to a set of effective parameters that are sufficient to characterize, with precision, the linear power spectrum. In this way, the linear power spectrum of a new model can be quickly computed with the ETHOS implementation in the Boltzmann solver, and constraints on the effective parameters can simultaneously constrain multiple particle models that have similar ETHOS parameters. The ETHOS parameters can then be converted back to the particle physics parameters of individual DM models. This creates a direct link between the parameters of particle DM models and those that are more relevant (directly connected) to cosmological structure formation.

In the ETHOS framework, the DAO features in the linear power spectrum are fundamentally caused by DM interacting with a new relativistic species called dark radiation (DR) in the early Universe. To obtain the DM and DR density perturbations, δ_χ and $\delta_{\rm DR}$, it is necessary to modify the Boltzmann equations for the DM perturbation, which are coupled to the DR ones in this case. The evolution of DM perturbations is described by a couple of equations in Fourier space *for each* Fourier mode with comoving wave number k:

$$\dot{\delta}_\chi + \theta_\chi - 3\dot{\Phi} = 0, \tag{1.10}$$

$$\dot{\theta}_\chi - c_\chi^2 k^2 \delta_\chi + \mathscr{H}\theta_\chi - k^2\Psi = \dot{\kappa}_\chi \left[\theta_\chi - \theta_{\rm DR}\right], \tag{1.11}$$

where θ_χ ($\theta_{\rm DR}$) is the divergence of the DM (DR) bulk velocity, Φ and Ψ are the two gravitational potentials in the conformal Newtonian gauge, $\mathscr{H}$ is the conformal Hubble parameter, and $\dot{\kappa}_\chi$ is the DM drag opacity. Overhead dots denote derivative with respect to conformal time. The DM sound speed c_χ^2 is given by

$$c_\chi^2 = \frac{T_\chi}{m_\chi}\left(1 - \frac{\dot{T}_\chi}{3\mathscr{H}T_\chi}\right), \tag{1.12}$$

where T_χ is the DM temperature, and m_χ is the DM particle mass. For non-relativistic DM, the sound speed only amounts to a small contribution in the perturbative equations. In the case of CDM, the sound speed $c_\chi = 0$ and drag opacity $\dot{\kappa}_\chi = 0$ vanish and the equations return to their well known form for CDM (see e.g. Ma & Bertschinger, 1995).

The interaction partner (DR) follows the relativistic Boltzmann equations

$$\dot{\delta}_{\mathrm{DR}} + \frac{4}{3}\theta_{\mathrm{DR}} - 4\dot{\Phi} = 0, \tag{1.13}$$

$$\dot{\theta}_{\mathrm{DR}} + k^2\left(\sigma_{\mathrm{DR}} - \frac{1}{4}\delta_{\mathrm{DR}}\right) - k^2\Psi = \dot{\kappa}_{\mathrm{DR}}(\theta_{\mathrm{DR}} - \theta_\chi), \tag{1.14}$$

$$\dot{\Pi}_{\mathrm{DR},l} + \frac{k}{2l+1}((l+1)\Pi_{\mathrm{DR},l+1} - l\Pi_{\mathrm{DR},l-1}) = (\alpha_l\dot{\kappa}_{\mathrm{DR}} + \beta_l\dot{\kappa}_{\mathrm{DR-DR}})\Pi_{\mathrm{DR},l}, \tag{1.15}$$

where σ_{DR} is the DR shear stress, $\dot{\kappa}_{\mathrm{DR}}$ is the DR opacity, α_l are the angular coefficients for DM-DR scattering, $\dot{\kappa}_{\mathrm{DR-DR}}$ is the DR self-scattering, β_l are the angular coefficients for DR-DR scattering, and $\Pi_{\mathrm{DR},l}$ is the l-th moment of the DR multipole hierarchy. In this Thesis, the DR self-scattering $\dot{\kappa}_{\mathrm{DR-DR}}$ will be set to zero, as it only directly affects the DR and it is a higher-order effect for the DM perturbation.

So far, Eqs. (1.10)-(1.15) are general and independent of the exact nature of the DM-DR interactions. For a specific DM model, the opacities $\dot{\kappa}_\chi$ and $\dot{\kappa}_{\mathrm{DR}}$, and the angular coefficients α_l can be computed from the matrix element for the respective DM-DR scattering process. For instance:

$$\dot{\kappa}_\chi = \frac{4\rho_{\mathrm{DR}}}{3\rho_\chi}\dot{\kappa}_{\mathrm{DR}} \propto a\int dp\,p^4\frac{\partial f^{(0)}_{\mathrm{DR}}(p)}{\partial p}(A_0(p) - A_1(p)), \tag{1.16}$$

where a is the scale factor, p is the magnitude of the three-momentum, $f^{(0)}_{\mathrm{DR}}$ is the homogeneous part of the DR phase-space distribution function, and the A_l are the projection of the squared matrix element on the l-th Legendre polynomial (for the exact expression see Cyr-Racine et al. 2016).

For several relevant DM particle models, the squared matrix element is a power law of the incoming DR momentum p_{DR}

$$|M|^2 \propto \left(\frac{p_{\mathrm{DR}}}{m_\chi}\right)^{n-2}, \tag{1.17}$$

where n is an integer. Therefore, it is a good assumption to model the DR and DM opacities as power laws in redshift. In ETHOS, the opacities are then approximated as[2]:

$$\dot{\kappa}_{\mathrm{DR}} = -(\Omega_\chi h^2)a_n\left(\frac{1+z}{1+z_{\mathrm{D}}}\right)^n, \tag{1.18}$$

$$\dot{\kappa}_\chi = -\frac{4}{3}(\Omega_{\mathrm{DR}}h^2)a_n\frac{(1+z)^{n+1}}{(1+z_{\mathrm{D}})^n}, \tag{1.19}$$

where a_n and the corresponding n are the effective ETHOS parameters, Ω_χ and Ω_{DR} are the DM and DR fraction of the total matter-energy density of the Universe today, h is the dimensionless Hubble constant, and $z_{\mathrm{D}} = 10^7$ is an arbitrary redshift to normalize the value of a_n. The parameter n controls the redshift scaling of the opacities and therefore

---

[2] Note that a sum of different power laws is also possible for more general models in ETHOS. However, in most cases a single power law is dominant during the relevant times and is therefore sufficient for the cases studied in this Thesis.

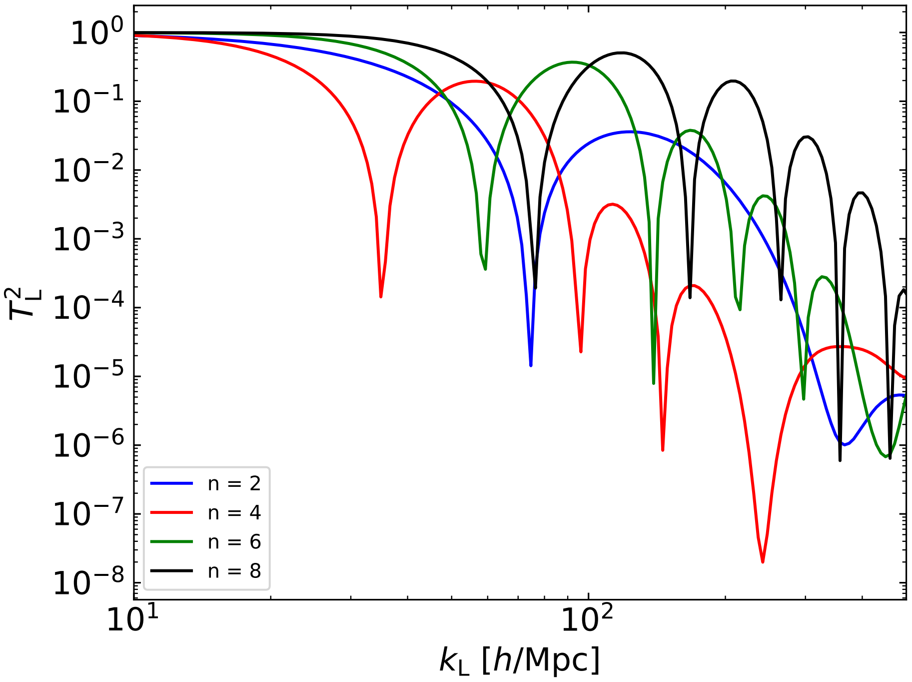

Figure 1.5.6. Linear transfer function $T_L^2 = P_{\mathrm{ETHOS}}/P_{\mathrm{CDM}}$ *of ETHOS models with different power n for the opacity redshift power law. A steeper redshift scaling of the DM-DR opacities (larger n) leads to a quicker transition from the coupled to decoupled regime and therefore less damping during the mildly coupled regime.*

also how quickly DM and DR decouple from each other. A slow transition from tightly coupled to decoupled (small n) leads to a stronger suppression of the power spectrum through collisional damping during the weakly coupled regime (e.g. in the models of Bœhm et al. 2002; van den Aarssen et al. 2012). A quick transition on the other hand (large n) strengthens the DAO impact in the power spectrum (see Fig. 1.5.6; e.g. in the atomic DM model of Cyr-Racine & Sigurdson 2013). For a fixed n, a_n controls the strength of the interactions and thus the time of decoupling. Increasing a_n decreases the wavenumber at which the first DAO peak occurs, i.e., the cut-off in the power spectrum occurs at larger scales.

The ETHOS framework in the linear regime can be summarized schematically by a mapping between particle physics and effective parameters

$$\{m_\chi, \{g_i\}, \{h_i\}\} \rightarrow \{\Omega_{\mathrm{DR}}, a_n, n, \alpha_l\}, \tag{1.20}$$

where the DM mass m_χ, coupling constants $\{g_i\}$ and internal parameters $\{h_i\}$ on the left describe the DM particle model and are connected to the ETHOS parameters on the right, which modulate the features in the linear power spectrum. As an example, the particle DM model of van den Aarssen et al. (2012) consists of DM interacting with a sterile neutrino ν via a light vector boson V and the resulting squared matrix element is

$$|M|^2 = 64 g_\chi^2 g_\nu^2 \frac{m_\chi^2 E_\nu^2}{m_V^4} \propto p_{\rm DR}^2, \tag{1.21}$$

where g_χ and g_ν are the coupling constants of DM and sterile neutrino with the V boson, respectively, m_χ and $m_{\rm V}$ are the DM and V mass, respectively, and E_ν is the energy of the scattering neutrino. From the $p_{\rm DR}^2$ scaling (see Eq. 1.21), it follows that $n = 4$ in this case and with Eqs. 1.16 and 1.19 it follows that

$$a_4 = (1+z_{\rm D})^4 \frac{3}{2} \frac{\pi g_\chi^2 g_\nu^2}{m_V^4} \frac{\rho_{\rm c}}{m_\chi} \left(\frac{310}{441}\right) T_{\rm DR}^2, \tag{1.22}$$

where $T_{\rm DR}$ is the DR temperature. For a detailed derivation, see Cyr-Racine et al. (2016).

1.6 Astrophysical probes of the DM nature at high-redshift

As DM does not interact with any particle of the standard model of particle physics in any significant way, the DM distribution in the Universe cannot be directly observed. Thus, it has to be inferred from observations of baryonic tracers of the underlying gravitational field, which receives contributions from both, DM and baryons. There is an abundant number of astrophysical systems that are observed to infer their DM distribution, some of which are mentioned at the beginning of this Chapter. In the context of this Thesis, there are two tracers of the matter power spectrum to highlight: the Lyman-α forest ($z = 3-5.4$) and the 21-cm signal in the even earlier Universe ($z = 6-30$).

The Lyman-α forest 1D flux spectrum. The Lyman-α transition is between the ground state and the first excited state of neutral hydrogen. A photon with the matching wavelength can therefore be absorbed by neutral hydrogen exciting its electron to a higher orbital. When light emitted over a broad spectrum from a bright and distant source (e.g. a quasar) travels through multiple intervening neutral hydrogen clouds along the line of sight, each cloud absorbs light at the Lyman-α wavelength creating an absorption line feature over the continuum spectrum. Each absorption line is however redshifted due to the expansion of the Universe, according to the location in redshift space of each of the intervening gas clouds. Therefore, when the light spectrum of the distant source is observed, it has a large number of absorption lines – the Lyman-α forest – at wavelengths above the Lyman-α rest wavelength and up to a $z_{\rm s}+1$ multiple thereof, where $z_{\rm s}$ is the redshift of the distant source. Thus, the wavelength and strength of each absorption line maps the density of neutral hydrogen within the line of sight. From combining multiple lines of sight, a 1D flux power spectrum can be constructed, which in turn can be used to infer the matter power spectrum (Croft et al., 1998, 1999). This offers a powerful probe for the nature of DM in the mildly non-linear regime ($z = 3 - 5.4$; $k = 0.1 - 10\,h\,\mathrm{Mpc}^{-1}$) and it has been used extensively to constrain a possible suppression of the linear matter power spectrum in the case of WDM (e.g. Viel et al., 2013; Iršič et al., 2017). The constraints of the WDM mass are relevant for Chapter 2 and the effects on the Lyman-α forest of sDAO models are discussed on Chapter 5.

21-cm signal from the cosmic dawn. The spin flip of an electron in the ground state of neutral hydrogen emits photons of 21-cm wavelength. As this transition has a very small transition rate (and small natural line width), it is highly unlikely to be observed in the laboratory and can only be observed on astronomical scales in gas clouds. The 21-cm signal does not require a background source as in the case of Lyman-α observations, but its emission or absorption can be detected against the CMB by comparing the observed 21-cm signal with the expected emission from the CMB. Therefore, it can probe the distribution of neutral hydrogen at the cosmic dawn, before the epoch of reionization ($z > 10$). These redshifts are especially interesting for DM physics, as DAO features are more prominent at higher redshifts, closer to the linear regime in structure formation theory. The population ratio of both spin states is determined by the spin temperature T_S and the change in the 21-cm brightness temperature T_b compared to the CMB temperature T_CMB is given by

$$T_{21} = T_\mathrm{b} - T_\mathrm{CMB} \propto \left(1 - \frac{T_\mathrm{CMB}}{T_\mathrm{S}}\right). \tag{1.23}$$

Therefore, if $T_\mathrm{S} = T_\mathrm{CMB}$, the absorption and emission of 21-cm photons are in equilibrium and there is no 21-cm signal. For larger spin temperatures however, there is net emission of 21-cm photons ($T_{21} > 0$), while a smaller spin temperature leads to net absorption ($T_{21} < 0$). The 21-cm signal develops in different phases. When the first generation of stars form, they emit Lyman-α photons which couple the spin temperature to the gas temperature through Wouthuysen-Field coupling, in which the electron gets excited from the ground state to the first excited state and then decays into the ground state with a spin flip (Wouthuysen, 1952; Field, 1959). This cools the spin temperature and leads to absorption. In the second phase, the X-rays of the first generation of stars heat the neutral hydrogen until $T_\mathrm{S} > T_\mathrm{CMB}$ leading to emission (Pritchard & Furlanetto, 2007; Pacucci et al., 2014). With the onset of reionization, the fraction of neutral hydrogen decreases reducing the 21-cm signal as well (Barkana & Loeb, 2001; Pritchard & Loeb, 2008). The exact timing of the different phases depends on the fraction of DM collapsed into haloes, and therefore on the DM model. In the case of WDM and models with DAOs, a delay in the collapse of the first haloes is predicted, and thus a 21-cm signal different to that predicted by CDM is expected, both in the global signal averaged over the full sky as well as in the fluctuation spectrum. The effects of DAO models on the 21-cm signal are shown in Chapter 4.

1.7 Thesis outline

In the following, a brief outline of the content of the different chapters is discussed. Chapters 2 and 3 are connected to two articles, one published and one under review. In both, I was the leading author. Chapters 4 and 5 are connected to two articles, both published; my contributions to these papers are mentioned below.

Chapter 2. The ETHOS framework described in Sec. 1.5 is very convenient to generalize linear perturbation theory and describe the *linear* evolution of alternative DM models, based (primarily) on the parameters that define the expansion of the DM opacity into power laws in redshift (see Eqs. 1.18-1.19 in Sec. 1.5). For the subsequent non-linear evolution however, the original ETHOS parametrization becomes less convenient, obscuring/limiting a physical interpretation of the evolution of the linear power spectrum into the non-linear regime, which should ideally be based on the relevant physical scale(s) of the linear damping mechanism(s). Therefore, one of the main purposes of this Thesis is to reparametrize the ETHOS framework in terms of parameters that have both, a more apparent physical interpretation and a simpler phenomenological connection to the features in the linear power spectrum of models with DAOs. This is done in Chapter 2 where two new effective and physically motivated parameters are introduced: $h_{\rm peak}$ and $k_{\rm peak}$ – the height and scale of the first DAO peak – that can fully describe the linear power spectrum of a variety of ETHOS models. Within this new parameter space, we performed a suite of N-body DM-only simulations down to $z = 5$, where we chose to stop the simulations due to several reasons. First, to avoid entering the evolutionary stage where DM self-interactions can have a significant impact on the inner structure of DM haloes (see e.g. Vogelsberger et al., 2012, 2014b, 2016; Rocha et al., 2013). Second, we performed DM-only simulations with the goal of isolating the impact of DM physics over that of baryonic physics. We note that by focusing in the high-redshift Universe, the impact of DAOs in the linear power spectrum becomes more apparent, while the role of baryonic physics is less entangled with a possible signature of new DM physics. Third, $z = 5$ is roughly the highest redshift probed by Lyman-α observations, which remain one of the most powerful observational constraints on the possibility of new DM physics in the power spectrum at small scales.

Chapter 3. We extend here the characterization of structure formation models with ETHOS developed in Chapter 2 by making a detailed analysis of the abundance and structure of haloes using the same simulation suite. In the first part, we apply the extended Press-Schechter (EPS) formalism (mentioned in Section 1.2 above) to describe the halo mass function of the different DM models, determining the appropriate parameters for the window function to match the simulation results. In the second part, we study the inner halo structure and compare the halo concentrations to the predictions of the model of Ludlow et al. (2016), which has been found to be remarkably accurate for CDM and WDM.

Chapters 4−5. While the DM matter power spectrum and halo mass function are direct statistical measures of the DM distribution and DM halo abundance, they are not directly observable. In Chapters 4-5, we study observational signatures of DAO models. In Chapter 4, we explore the detectability of the DAOs in the 21-cm signal (both global and the fluctuations spectrum) in the full range of the DAO parameter space (h_{peak}, k_{peak}). My contribution to this project was the execution of the necessary N-body simulations and the preparation of their output for the 21-cm simulations, as well as contributing to the adaptation of the extended Press-Schechter formalism for DAO models into the 21-cm code. In Chapter 5, we explore a strong DAO model within a cosmological hydrodynamical simulation with the specific goal of simulating its Lyman-α 1D flux spectrum signal, and comparing it to a WDM model with identical cut-off scale. One of the purposes was to test to what extent the distinctive DAO features in the linear power spectrum can survive non-linear structure formation and remain truly distinct from the featureless WDM case. For this project, I performed part of the simulations.

Chapter 2

The non-linear extension of the ETHOS framework

This chapter is based on the following article:

ETHOS – An effective parametrization and classification for structure formation: the non-linear regime at $z \gtrsim 5$

Published in Monthly Notices of the Royal Astronomical Society, Volume 498, Issue 3, pp. 3403-3419 (2020), Oxford University Press

Authors:
Sebastian Bohr[1], Jesús Zavala[1], Francis-Yan Cyr-Racine[2], Mark Vogelsberger[3], Torsten Bringmann[4] and Christoph Pfrommer[5]
[1]Centre for Astrophysics and Cosmology, Science Institute, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland
[2]Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, USA
[3]Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Department of Physics, University of Oslo, Box 1048, N-0371 Oslo, Norway
[5]Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany

We propose two effective parameters that fully characterise galactic-scale structure formation at high redshifts ($z \gtrsim 5$) for a variety of dark matter (DM) models that have a primordial cutoff in the matter power spectrum. Our description is within the recently proposed ETHOS framework and includes standard thermal Warm DM (WDM) and models with dark acoustic oscillations (DAOs). To define and explore this parameter space, we use high-redshift zoom-in simulations that cover a wide range of non-linear scales from those where DM should behave as CDM ($k \sim 10\,h\,\mathrm{Mpc}^{-1}$), down to those

characterised by the onset of galaxy formation ($k \sim 500\,h\,\mathrm{Mpc}^{-1}$). We show that the two physically motivated parameters $h_{\rm peak}$ and $k_{\rm peak}$, the amplitude and scale of the first DAO peak, respectively, are sufficient to parametrize the linear matter power spectrum and classify the DM models as belonging to effective non-linear structure formation regions. These are defined by their relative departure from Cold DM ($k_{\rm peak} \to \infty$) and WDM ($h_{\rm peak} = 0$) according to the non-linear matter power spectrum and halo mass function. We identify a region where the DAOs still leave a distinct signature from WDM down to $z = 5$, while a large part of the DAO parameter space is shown to be degenerate with WDM. Our framework can then be used to seamlessly connect a broad class of particle DM models to their structure formation properties at high redshift without the need of additional N-body simulations.

2.1 Introduction

Dark matter (DM) is a crucial ingredient in the formation of structures in the Universe as it makes up the majority of its matter content. Although the most likely explanation for DM is the particle hypothesis, its specific nature remains a mystery. The CDM model of structure formation has now emerged as the standard paradigm, and it has been shown to be consistent with the observed large scale structure of the Universe (e.g. Springel et al., 2005). At smaller (galactic) scales however, the CDM model has faced a number of significant challenges over the last decades: (i) the underabundance of low-mass galaxies (either satellites or in the field) (Klypin et al., 1999; Moore et al., 1999; Zavala et al., 2009; Papastergis et al., 2011; Klypin et al., 2015), (ii) the core-cusp problem in low-surface brightness galaxies and possibly in dwarf spheroidals (de Blok & McGaugh, 1997; Walker & Peñarrubia, 2011), (iii) the "too-big-to-fail problem" (Boylan-Kolchin et al., 2011; Papastergis et al., 2015), (iv) the plane of satellites problem (Pawlowski et al., 2013), and (v) the diversity problem of rotation curves in dwarf galaxies (Oman et al., 2015). We note that with recent observations of ultra-faint galaxies, the too-big-to-fail problem becomes a diversity problem as well for the broad distribution of stellar kinematics in dwarf spheroidals in the Milky Way (Zavala et al., 2019).

There is a long history of attempts to provide a satisfactory solution to these issues based on either: (i) incompleteness, biases and systematic uncertainties in observations (e.g. Koposov et al., 2008; Kim et al., 2018, for the “missing satellites problem"), (ii) invoking a strong influence of uncertain baryonic physics in dwarf galaxies (e.g. impulsive supernova feedback to explain DM cores (Pontzen & Governato, 2012), tidal effects from the Milky-Way disk to alleviate the too-big-to-fail problem (Garrison-Kimmel et al., 2019), and suppression of galaxy formation at the dwarf mass scale due to cosmic reionisation (Gnedin, 2000; Sawala et al., 2016b) to explain the underabundance of low-mass galaxies); and (iii) additional DM physics, i.e. departures from the CDM hypothesis such as: Warm Dark Matter (WDM; for a review see Adhikari et al. 2017)

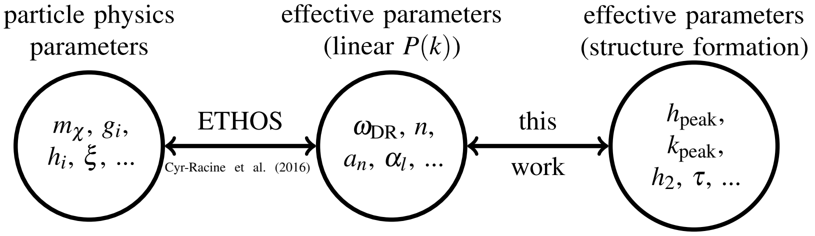

Figure 2.1.1. Diagram illustrating the different sets of parameters that characterise a given DM particle model and the connections between them within the ETHOS framework, both in Cyr-Racine et al. (2016) and in this work. The particle physics space parameters such as the DM particle mass m_χ, coupling constants g_i (e.g. DM-DR), internal parameters h_i such as the mediator mass and degrees of freedom and the present day DR to CDM temperature ratio ξ, were mapped in Cyr-Racine et al. (2016) into effective parameters fully describing the linear DM power spectrum (see section 2.3.1). In this work, we make a re-parametrization, defining new ETHOS parameters that have both a more straightforward interpretation in terms of the linear power spectrum and a clearer physical interpretation (amplitude $h_{\rm peak}$ and scale $k_{\rm peak}$ of the first DAO peak, amplitude of the second peak h_2, and damping of higher order peaks τ; see section 2.3). The redefined ETHOS parameter space can be connected naturally to that defined in section 2.3.1, and thus to the particle physics space. Crucially, it is also sufficient to characterise non-linear structure formation for a variety of relevant DM models (such as WDM and models with DAOs) in the high-redshift Universe.

where the relativistic motion of the DM particles in the early Universe reduces the abundance and inner DM densities of galactic-scale haloes relative to CDM (e.g. Colín et al., 2000; Lovell et al., 2012); self-interacting DM (SIDM; for a review see Tulin & Yu 2018) where DM particles have strong self-interactions redistributing energy in the centre of haloes, thus resulting in DM cores (e.g. Spergel & Steinhardt, 2000; Vogelsberger et al., 2012); and quantum effects at galactic scales if DM is made of extremely light bosons with $\mathcal{O}(1\,\mathrm{kpc})$ de Broglie wavelength (fuzzy DM; for a review see Hui et al. 2017), also giving rise to extended DM cores (Robles & Matos, 2012; Mocz et al., 2017).

Whether the CDM challenges are due to missing new DM physics, systematic uncertainties, or an inaccurate account of baryonic physics remains an open question (for a recent review on different DM models and their impact on structure formation see Zavala & Frenk 2019; for a review of the CDM challenges and possible solutions see Bullock & Boylan-Kolchin 2017). Regardless of the answer to these puzzles, the impact of the DM particle nature on the physics of galaxies remains a relevant factor that needs to be taken into account, not only because it causes a major and unavoidable uncertainty in structure formation, but also because the detailed properties of galaxies remain one of the most promising avenues to find clues about the DM identity. To incorporate new DM physics into structure formation theory, a novel framework has been proposed that aims at mapping a broad range of DM particle physics models into a set of effective parameters that fully characterise structure formation at galactic scales (ETHOS; Cyr-

Racine et al., 2016; Vogelsberger et al., 2016). Thus far, ETHOS covers two types of new DM physics: (i) a primordial cutoff in the linear matter power spectrum suppressing the growth of small density perturbations due to either collisionless damping (free-streaming) like in WDM, or due to collisional damping caused by interactions between DM and relativistic particles in the early Universe and resulting in Dark Acoustic Oscillations (DAOs; for a review see Bringmann 2009). DAOs are given explicitly in ETHOS by hidden sector DM-dark radiation (DR) interactions (van den Aarssen et al., 2012; Buckley et al., 2014) but DM interactions with photons or neutrinos lead to a similar damping, (e.g. Bœhm et al., 2002); (ii) DM self-interactions (SIDM) reducing the central density of haloes in the non-linear regime.

In this work we concentrate exclusively on (i) above, i.e., on the impact of a primordial DM cutoff with the objective of defining a parameter space that fully characterises structure formation within ETHOS at galactic scales (at high redshift $z > 5$; see below). Ours is then a continuation of the work done in Cyr-Racine et al. (2016) where a small set of effective parameters was defined that where sufficient to characterise the *linear* power spectrum in a variety of DM models with a cutoff. However, a large number of models that are different with respect to their linear power spectrum can in fact lead to identical structure formation. Therefore, we re-parametrize this effective ETHOS parameters, still being determined by the linear power spectrum, but with the goal of providing a full account of the non-linear evolution of galactic-scale structures (down to $z = 5$) using cosmological $N-$body simulations and a physical interpretation of the parameters; see Fig. 2.1.1. We aim at dividing this new ETHOS parameter space into distinct structure formation regions, mapping smoothly between the different possibilities for the small-scale power spectrum (CDM, WDM or DAOs).

We note that previous works have proposed analytical formulae to describe the linear power spectra of different DM models, usually written as a transfer function relative to CDM (e.g. for WDM Bode et al., 2001; Viel et al., 2005; Leo et al., 2018b). More recently, Murgia et al. (2017, 2018) proposed a formula for the transfer function that can seemingly accommodate WDM, fuzzy DM, and also certain ETHOS models. Crucially however, this formula does not describe DAOs since they were deemed not relevant for the properties of interest in Murgia et al. (2017, 2018), namely, for the 1D Lyman$-\alpha$ flux power spectrum, and for the number of observable Milky-Way subhaloes (i.e. those that can host a luminous satellite). As we demonstrate and quantify in this work, DAOs are quite relevant for a range of ETHOS models (see also Bose et al., 2019). Moreover, our approach differs from previous ones since the parametrization we propose goes beyond providing a fit to the power spectrum, with the parameters having a clear physical interpretation.

In this work we study structure formation within ETHOS in the high-redshift Universe down to $z = 5$. This choice was partly done to avoid entering the regime where DM self-interactions (another relevant ingredient in ETHOS) start to have a relevant impact in the centre of DM haloes. We are also only considering the impact of DM physics in structure formation without taking into account the role of baryonic physics, which clearly plays a role in DM clustering, albeit considerably smaller at high-redshift relative to the low-redshift Universe. In this way we can isolate the potential difference between CDM and other DM models, purely due to DM physics, without the influence of baryons; this is in fact needed to disentangle the impact of both effects. Our plan in future work is to extend the spirit of this work, by defining the space of structure formation parameters that are relevant for the physics of galaxies, to include both SIDM and baryonic physics. We also choose $z = 5$ as the lowest redshift we examine since it is roughly the maximum redshift where data from the Lyman-α flux-power spectra have been used to constrain the DM power spectrum at small scales (e.g. Viel et al., 2013; Murgia et al., 2018). We use this both to exemplify how our parametrization can be used to potentially constrain DM models and to define the maximum scale where new DM physics can play a role in galactic-scale structure formation: DM models with a non-linear power spectrum significantly deviating from CDM at $k \lesssim 10\,h\,\mathrm{Mpc}^{-1}$ are not compatible with the data (Iršič et al. 2017; although see Garzilli et al. 2019a). On the other hand, we set the relevant minimum scale to be given by the atomic cooling limit (specifically, the primordial gas in haloes with a virial temperature $\lesssim 10^4$ K cannot cool via atomic transitions; see White & Rees 1978). Galaxy formation is thus suppressed for DM haloes with masses below $\sim 10^8 \mathrm{M}_\odot\, h^{-1}$ (corresponding to non-linear scales of $\sim 500\,h\,\mathrm{Mpc}^{-1}$). In summary, we study non-linear structure formation down to $z = 5$ within the ETHOS framework using DM-only N-body simulations focusing on the non-linear scale range $10\,h\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 500\,h\,\mathrm{Mpc}^{-1}$ (halo virial masses in the range $10^8 \mathrm{M}_\odot \lesssim M_{\mathrm{vir}} \lesssim 10^{10} \mathrm{M}_\odot$).

This paper is organized as follows. In Section 2.2, we describe our simulation setup and the zoom-in method we use to cover the dynamic range of interest. The convergence properties of our simulations is discussed in Appendix 2.6.1. In Section 2.3, the new ETHOS parametrization is constructed and connected to that in Cyr-Racine et al. (2016) (see also Appendix 2.6.2). In Section 2.4, we present our main results on how different structure formation models are classified within the new parametrization based on both the non-linear power spectrum and the halo mass function. Finally, our conclusions are given in Section 2.5.

2.2 Numerical Methodology

The cosmological dark-matter-only $N-$body simulations used in this work were performed with the code AREPO (Springel, 2010). Initial conditions for the simulations were generated using the code MUSIC (Hahn & Abel, 2011) with the cosmological parameters set to $\Omega_{\mathrm{m}} = 0.31069$, $\Omega_\Lambda = 0.68931$, $H_0 = 67.5\,\mathrm{km/s/Mpc}$, $n_{\mathrm{s}} = 0.9653$ and $\sigma_8 = 0.815$, where Ω_{m} and Ω_Λ are the contributions from matter and cosmological constant to the matter-energy density of the Universe today, respectively, H_0 is the Hubble constant today, n_{s} is the spectral index, and σ_8 is the mass variance of linear fluctuations in $8\,\mathrm{Mpc}\,h^{-1}$ spheres at $z = 0$. This choice of cosmological parameters is consistent with a Planck cosmology (Planck Collaboration et al., 2020). The linear power spectrum used as an input for MUSIC for the different DM models we explore is computed with a modified version of CLASS[1] (Archidiacono et al., 2017, 2019)[2].

For this work, we are interested in the matter power spectrum at the *non-linear* scales relevant for dwarf galaxies at high redshift from $k \sim 10$ to $500\,h\,\mathrm{Mpc}^{-1}$. This range roughly corresponds to halo masses $\sim 10^{10}$ to 10^{7} $\mathrm{M}_{\odot}$, which are close to the limits where significant deviations from CDM are possible, and galaxy formation becomes highly inefficient, respectively. Achieving a fair representation of the power spectrum and the halo mass function at such small scales was not feasible with a uniform simulation box due to the following stringent limitation. At a fixed spatial resolution, reducing the size of the cosmological box, reduces the minimal scales probed, but at the cost of missing the power transferred from larger to smaller scales in the non-linear evolution. Thus, the power spectrum at the scales and redshifts of interest would be biased towards lower values. This problem can be alleviated by having a sufficiently large simulation box, but in order to resolve $500\,h\,\mathrm{Mpc}^{-1}$, the amount of particles and thus the computational cost increases dramatically. To achieve our goals we therefore rely instead on cosmological *zoom-in* simulations by using the method described in the following.

2.2.1 Small-scale power spectrum with zoom simulations

In a zoom-in cosmological simulation, the computational resources are focused on a smaller subregion within a large cosmological box. This subregion is simulated at the desired highest resolution, while the volume around contains low resolution elements that still preserve an accurate representation of large scale properties of the density field. The region of interest is usually a halo and its immediate environment, and the procedure to construct the initial conditions for zoom simulations consists of: i) run a low resolution *parent* cosmological simulation within a cosmological box large enough to provide a fair representation of the clustering properties at the scales of the box in the range of redshifts of interest (we found that a box size of $40\,\mathrm{Mpc}\,h^{-1}$per side satisfies this for $z \geq 5$); ii) select a volume within the parent simulation encompassing the region of interest at the redshift of interest; iii) trace back the particles within this region to the starting redshift of the resimulation; this represents the target *Lagrangian* volume for the zoom simulation; (iv) finally, an initial conditions code like MUSIC is used to generate a multi-layered resolution coverage of the resimulation specifying the volume that covers this Lagrangian region with the highest resolution required. For more details of the general procedure see Oñorbe et al. (2014).

In our case we followed the previous standard zoom-in procedure but not focusing on a particular halo in the parent simulation but rather on a smaller subregion with the main requirement for it to have a *similar* power spectrum compared to the larger parent box at the resolved scales. In the following we describe how we find the optimal sub-region according to this requirement.

---

[1] Blas et al. (2011) (class-code.net)

[2] Note that in the first ETHOS paper (Cyr-Racine et al., 2016), this implementation is done with CAMB (Lewis & Bridle, 2002).

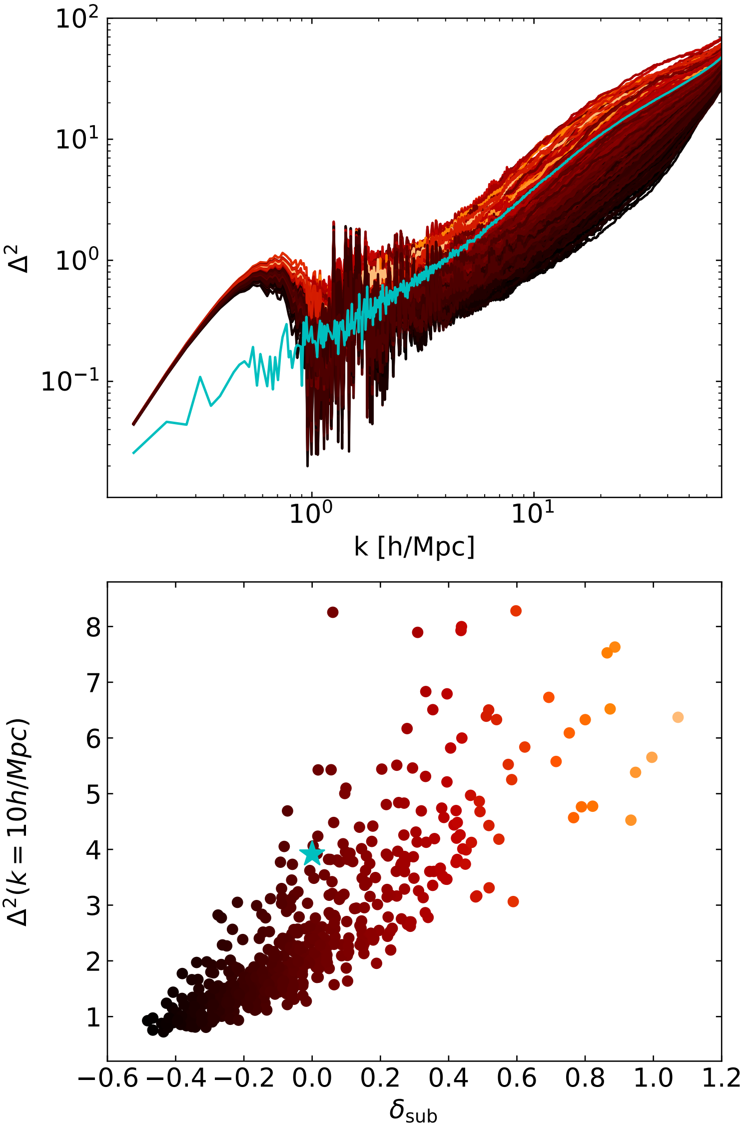

Figure 2.2.2. Top panel: Dimensionless power spectra Δ^2 *at* $z = 5$ *of a large* $40\,Mpc\,h^{-1}$*cubic uniform simulation (cyan) and of sub-regions inside this simulation (*$\sim 6.25\,Mpc\,h^{-1}$*) coloured from under-dense (black) to over-dense (orange) relative to the larger box. The excess of power at small* $k \lesssim 1\,h\,Mpc^{-1}$*in the small box simulations is due to the finite size of the sub-regions, while at large k the models start to converge artificially due to Poisson noise, which starts to dominate the signal (visible at around* $k \gtrsim 40\,h\,Mpc^{-1}$*). Bottom panel: Correlation of the dimensionless power spectrum at* $10\,h\,Mpc^{-1}$*and the overdensity of the sub-region* δ_{sub} *(the colour scale is as in the top). The cyan star corresponds to the value for the whole simulation box.*

We find that at the scales of interest the power spectrum has roughly a similar shape, but with an amplitude that correlates strongly with the overdensity $\delta_{\rm sub}$ of the subregion, as can be seen in Fig. 2.2.2. This type of *cosmic variance* is a well known effect that has been studied in the past, particularly in the linear regime where $\delta_{\rm sub} \ll 1$. In this regime, it is possible to correct for this bias by e.g. using the separate universe approach, where each subregion is treated as a separate universe with a different cosmology, in this case a universe with different background density (see e.g. Chiang et al., 2014; Li et al., 2014). As is clear from the bottom panel of Fig. 2.2.2, the distribution of $\delta_{\rm sub}$ is broad, covering values that are clearly non-linear anymore. This is because we are looking at smaller scales where the impact of non-linear effects is stronger and the variance of Δ^2 for regions with the same overdensity can be quite large, weakening the correlation between the amplitude of the power spectrum and the overdensity. Instead of trying to generalize the separate universe approach into the non-linear case, we decided to carefully select our high resolution sub-region so that it has a power spectrum that is as similar as possible to the one of the larger lower resolution region, at the scales that both can resolve. In this way a correction becomes unnecessary. This is sufficient for our purposes since we are only interested in an average measure of the power spectrum down to small (galactic) scales, rather than in its variance. Nevertheless, Fig. 2.2.2 gives an impression of the (cosmic) variance to be expected in the power spectrum for survey volumes that are small $\lesssim 10$ Mpc.

2.2.2 Performance of the zoom-in simulation technique

As a benchmark test for the reconstruction ability and resource advantage of the method described above, we performed four CDM simulations in a $(40\,{\rm Mpc}\,h^{-1})^3$ volume down to $z = 5$. The baseline is a uniform simulation with 1024^3 particles, while the other three are zoom simulations where the low-resolution region corresponds to 512^3 particles. The first of these has a $(12.5\,{\rm Mpc}\,h^{-1})^3$ zoom region with an effective resolution of 1024^3 particles, the second one has a $(6.25\,{\rm Mpc}\,h^{-1})^3$ zoom region with an effective resolution of 2048^3 particles, and the third one has a $(6.25\,{\rm Mpc}\,h^{-1})^3$ zoom region with an effective resolution of 4096^3 particles.

In Fig. 2.2.3, we can see that all three zoom simulations give a good reconstruction of the baseline power spectrum at large scales. As expected, the one with the same effective resolution as the baseline (1024^3 particles; red line) shows almost the same power spectrum as the uniform one at all scales, with nearly the same level of Poisson noise. The other two zoom simulations can resolve the power spectrum at smaller scales by factors of 4 (green) and 8 (blue) relative to the uniform simulation. This test shows that we can measure the power spectrum across a large dynamical range using the zoom simulation technique described in Section 2.2.1. More importantly, it is possible to achieve this with a reduced computational cost as we show in Table 2.2.1. For instance, the zoom simulation with same effective resolution as the uniform one (red and black lines in Fig. 2.2.3) uses only a small fraction ($\lesssim 1/7$) of the core hours and less than half the memory of the uniform simulation. Even our highest resolution

	core-h	memory	resolved k
uniform (1024^3)	14.9k	594 GB	$\sim 100\,h\,\mathrm{Mpc}^{-1}$
12.5 Mpc (1024^3)	2k	247 GB	$\sim 100\,h\,\mathrm{Mpc}^{-1}$
6.25 Mpc (2048^3)	2.9k	259 GB	$\sim 400\,h\,\mathrm{Mpc}^{-1}$
6.25 Mpc (4096^3)	15.1k	529 GB	$\sim 800\,h\,\mathrm{Mpc}^{-1}$

Table 2.2.1. Computing resources needed to reach $z = 5$ for a uniform simulation with 1024^3 particles and three different zoom simulations.

zoom simulation uses about the same core-hours and memory compared to the uniform run, while improving the scales that can be probed by a factor of 8, making it a very affordable approach to probing the power spectrum at small scales.

Based on this test, we use the following setting for the results presented in this paper (unless stated otherwise): a cosmological box with $40\,\mathrm{Mpc}\,h^{-1}$on a side with a high resolution zoom region (effective resolution of 4096^3 particles with a particle mass of $8 \times 10^4\,\mathrm{M}_{\odot} h^{-1}$) covering a $\sim(6.25\,\mathrm{Mpc}\,h^{-1})^3$ Lagrangian volume, surrounded by a low resolution region (effective resolution of 512^3 particles), and intermediate resolution levels as a buffer zone between them. With this setting, we find that the power spectra of all DM models presented in this work is converged to better than 5% at $500\,h\,\mathrm{Mpc}^{-1}$, while the halo mass function is converged to better than 5% down to $10^8\mathrm{M}_{\odot}/h$. In Appendix 2.6.1 we explicitly show the convergence tests we performed.

2.3 Parametrization of the linear power spectrum

Our goal in this Section is to present a new parametrization of the linear power spectrum for DM models that have a primordial power spectrum cutoff with or without DAOs within the ETHOS framework. This parametrization is purely phenomenological but it is constructed with two objectives in mind: (i) although it parametrizes the linear power spectrum, its parameters should be sufficient to describe with good precision the non-linear power spectrum and (ii) the parameters should have a clear physical interpretation. To accomplish this, our starting point is the work of Murgia et al. (2017) who suggested the following parametrization for the cutoff of non-CDM (nCDM) models in terms of the linear transfer function $T_{\mathrm{L}}^2(k) \equiv P_{\mathrm{nCDM}}(k)/P_{\mathrm{CDM}}(k)$:

$$T_{\mathrm{L}}(k) = [1 + (\alpha k)^{\beta}]^{\gamma}, \tag{2.1}$$

where α is a measure of the cutoff scale length, and β and γ the shape of the cutoff. This is a generalization of the fitting formula for WDM, where $\beta = 2\nu$ and $\gamma = -5/\nu$ with $\nu = 1.12$, and allows for much higher variety in the shape of the cutoff. However,

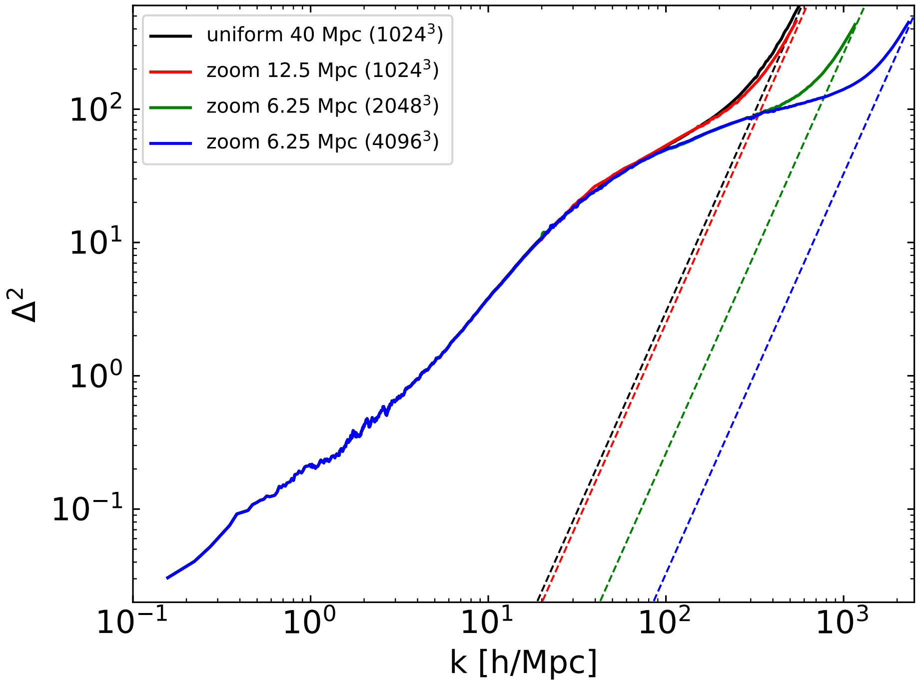

Figure 2.2.3. Dimensionless power spectra Δ^2 *at* $z = 5$ *for a uniform simulation (black) and three zoom simulations with different zoom volumes and resolution levels. Note that the power spectrum from the zoom simulations can reach the same larger scales as the uniform simulation by including the low resolution particles. Thus, all lines completely overlap for* $k \lesssim 30\,h\,Mpc^{-1}$. *The dashed lines show the expected Poisson (shot) noise* ($\Delta^2_{\rm shot} = k^3V/(2\pi^2N)$), *where N is the number of particles and V the volume they occupy) for the corresponding simulation.*

it only describes a single cutoff in the power spectrum, while we want to include models with DAOs as well. We remark however that the transfer of power from large to smaller scales in the non-linear evolution tends to erase the DAOs (e.g. Buckley et al., 2014). Since one of the goals of our parametrization is to reproduce with good precision the non-linear evolution of the power spectrum down to $z = 5$, we thus start by looking at the accuracy to which Eq. (2.1) can be expected to account for the non-linear regime. This can be seen in Fig. 2.3.4, where the red line corresponds to Eq. (2.1). Comparing the results from our simulations using this parametrization and the power spectrum with several DAOs (black lines) as initial conditions, we find that it is not sufficient to capture with precision the amplitude and features of the non-linear power spectrum at small scales for models which have strong DAO features (i.e. where the first oscillations are near the CDM amplitude). As can be seen in Fig. 2.3.4 (red line), for this particular strong DAO model, this parametrization underestimates the power at $k \gtrsim 100\,h\,\mathrm{Mpc}^{-1}$. For instance, by up to 48% and 24% at $k = 500\,h\,\mathrm{Mpc}^{-1}$for $z = 8$ and 5 respectively. In Section 2.4.1 we quantify in detail the impact of DAOs in the non-linear power spectrum

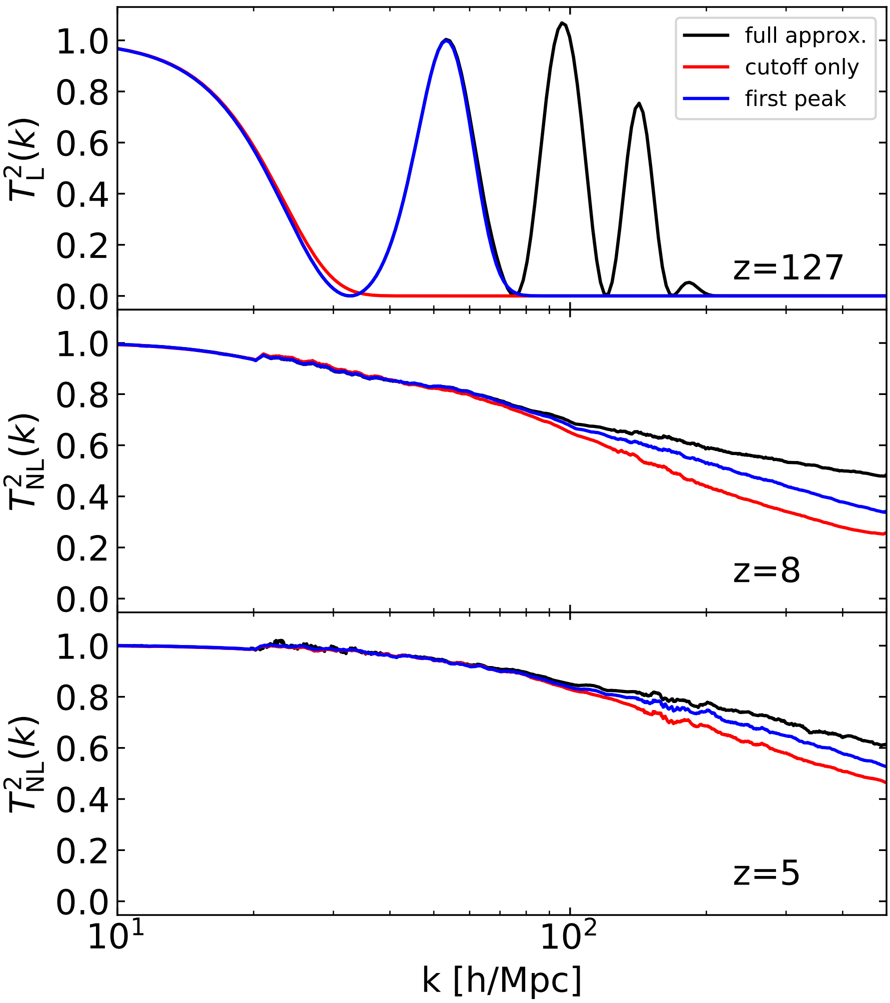

Figure 2.3.4. Top panel: Initial transfer function $T_{\mathrm{L}}^2(k)$ *of a sDAO model computed with Eq.* (2.3) *(black line) and two approximations: considering only the initial cutoff (i.e. with a WDM-like parametrization; see Eq. 2.1), and adding as well the first oscillation (blue line; see Eq. 2.2). Middle and bottom panels: Comparison of the non-linear dimensionless power spectra* Δ^2 *relative to CDM at* $z = 8$ *and* $z = 5$ *for the models shown in the top panel. All these models used the following parameters:* $h_{\mathrm{peak}} = 1$, $k_{\mathrm{peak}} = 53.3$, $h_2 = 1.08$, $\tau = 0.67$, $\sigma = 0.2$, $\beta = 4.05$, $\gamma = -20$, $d = 2.5$, *and* α *according to Eq.* (2.4). *This corresponds to a model in the DAO region of Fig. 2.3.10.*

for a broad range of scales and amplitudes of the DAOs.

In a first attempt to improve Eq. (2.1) to account for DAO models, we add a term that includes the first DAO peak by modelling it with a Gaussian:

$$T_{\mathrm{L}}(k) = [1 + (\alpha k)^{\beta}]^{\gamma} - \sqrt{h_{\mathrm{peak}}} \exp\left(-\frac{1}{2} \left(\frac{k - k_{\mathrm{peak}}}{\sigma k_{\mathrm{peak}}} \right)^2 \right), \tag{2.2}$$

where h_{peak} and k_{peak} give the amplitude (relative to CDM) and position of the first peak (these two will be the most important parameters throughout this work), and σ controls how narrow the Gaussian is. Eq. (2.2) improves the agreement with the full non-linear power spectrum as can be seen in Fig. 2.3.4, but is still not good enough to reconstruct the full power at the smallest scales, where it underestimates the power by 30% and 14% for $z = 8$ and 5 respectively. Therefore, the power provided by the secondary peaks in sDAO models remains relevant down to $z = 5$.

To gain precision in our parametrization for models that have DAOs, we extend Eq. (2.2) by adding terms that model the secondary peaks of the DAOs. These peaks can be described by two features, their envelope and oscillations. The oscillations of the higher order peaks are very regular and can be fitted with a cosine function, whose frequency is determined by k_{peak}. The envelope can be parametrized with the amplitude of the second peak h_2 and two error functions, one giving the steep rise on the left side (similar to the Gaussian describing the first peak) and the other controlling the damping on the right (the oscillations are not fully symmetrical and thus a Gaussian is not enough to describe their shape). The full fitting function is then given by:

$$\begin{aligned} T_{\mathrm{L}}(k) = {} & [1 + (\alpha k)^{\beta}]^{\gamma} - \sqrt{h_{\mathrm{peak}}} \exp\left(-\frac{1}{2} \left(\frac{k - k_{\mathrm{peak}}}{\sigma k_{\mathrm{peak}}} \right)^2 \right) \\ & + \frac{\sqrt{h_2}}{4} \operatorname{erfc}\left(\frac{k - 1.805 k_{\mathrm{peak}}}{\tau k_{\mathrm{peak}}} - 2 \right) \\ & \times \operatorname{erfc}\left(-\frac{k - 1.805 k_{\mathrm{peak}}}{\sigma k_{\mathrm{peak}}} - 2 \right) \cos\left(\frac{1.1083 \pi k}{k_{\mathrm{peak}}} \right), \end{aligned} \tag{2.3}$$

where $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ is the complementary error function. This full parametrization would have 8 parameters, but they are not all independent and can be simplified for ETHOS models, where they are fixed by h_{peak} and k_{peak} (see Section 2.3.1 below). The parameter α can be determined by the scale at which the transfer function dropped to $1/2$ ($k_{1/2}$), which is connected to k_{peak}:

$$\alpha = \frac{d}{k_{\mathrm{peak}}} \left[\left(\frac{1}{\sqrt{2}} \right)^{1/\gamma} - 1 \right]^{1/\beta}, \tag{2.4}$$

where d controls the ratio between k_{peak} and $k_{1/2}$, which is in the range $2.4 - 3$.

Fitting the power spectrum cutoff of ETHOS models with Eq. (2.1) leads in all cases to large negative values for γ, whose precise value makes almost no difference in the reconstruction of the cutoff; thus, we have fixed $\gamma = -20$. From the remaining parameters, h_{peak} and k_{peak} are the most relevant parameters since the former determines the position of the first DAO peak as well as describing the position of the cutoff (see Eq. 2.4), and the latter the amplitude of the first DAO. These parameters are responsible for the leading order effects on the non-linear power spectrum and in fact the only

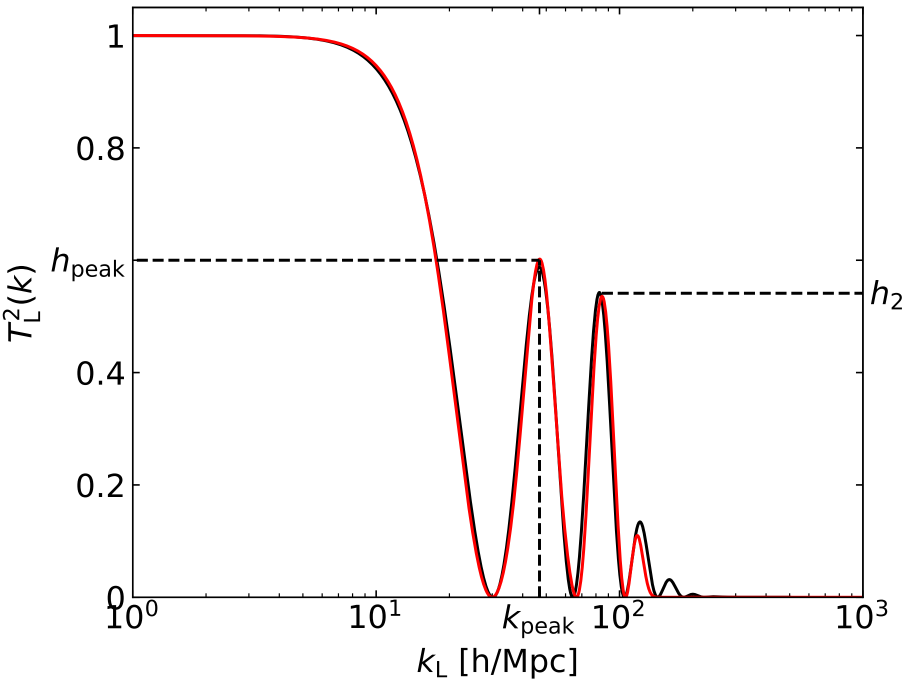

Figure 2.3.5. The transfer function $T^2_{\mathrm{L}}(k)$ of a DAO model computed with a Boltzmann solver (black) and fitted according to Eq. (2.3) *(red). The role of the most relevant parameters h_{peak}, k_{peak} and h_2 is also shown.*

free parameters within the models we study in this work. Regarding the remaining parameters: β is responsible for the cutoff shape, τ controls the damping of the DAOs, and σ gives the width of the first peak. Physically, k_{peak} is connected to the DM sound horizon and τ to the Silk damping scale (the physical interpretation of the key parameters is described in Section 2.3.1 and Appendix 2.6.2). The effects of the parameters and the quality of our final parametrization in the linear transfer function can be seen in Fig. 2.3.5.

We emphasize that the parametrization given by Eq. (2.3) can accurately describe the entire range of DM models in the ETHOS framework that display DAOs in their linear transfer function, including both weak and strong oscillations with only two free parameters ($h_{\mathrm{peak}}, k_{\mathrm{peak}}$). Furthermore, it also naturally encompasses WDM ($h_{\mathrm{peak}} \to 0$) and CDM ($k_{\mathrm{peak}} \to \infty$), allowing us to explore a very broad range of possible DM physics.

2.3.1 Connection with the ETHOS framework and physical interpretation of the parameters: $h_{\rm peak}$, $k_{\rm peak}$ and h_2

Having accomplished the goal of providing a parametrization for DAO models that is simpler than a fully general parametrization of the linear power spectrum (as provided in Cyr-Racine et al., 2016) but still precise enough to describe their non-linear evolution, we proceed now to establish the connection between these phenomenological parameters and the physical parameters of the ETHOS framework (Cyr-Racine et al., 2016) in regards to the effects of the DM-DR interactions in generating the power spectrum cutoff and the DAOs. We recall that with such a connection, it is then possible to have a complete mapping between the particle physics parameters of the models explored in Cyr-Racine et al. (2016) and the parameters relevant for non-linear structure formation.

The physics of the DAOs in the linear power spectrum is captured within the modelling presented in Cyr-Racine et al. (2016)[3] by the parameters n and coefficient a_n that control the redshift scaling of the DM drag opacity $\dot{\kappa}_\chi \propto a_n(1+z)^{n+1}$, plus a set of coefficients α_l that parametrize the angular dependence of the DM-DR scattering cross section[4]. For this work, we will refer only to models that have single values for n and a_n, and a set of constant $\alpha_{l\geq 2}$ values. A specific particle physics scenario contained within these constraints is that of a massive fermionic DM particle interacting with a massless fermion via a massive vector mediator as in van den Aarssen et al. (2012), which corresponds to the case $n=4$, $\alpha_{l\geq 2}=3/2$ with different values of a_4 providing cutoff scales for the power spectrum. This specific model has been studied with simulations in the past (Vogelsberger et al., 2016), particularly the benchmark model referred to as ETHOS-4 in table 1 of Vogelsberger et al. (2016).

Although the parameters n, a_n and the set $\{\alpha_l\}$ are sufficient to characterize the linear power spectrum within the ETHOS framework, they obscure somewhat the physical mechanism behind the DAOs, and they also lack the simple phenomenological interpretation of the parameters described above $\{k_{\rm peak}, h_{\rm peak}, h_2\}$. Because of this, we first attempt to approximate the results of the full calculation of the linear power spectrum based on a Boltzmann code (modified version of CLASS; Archidiacono et al. 2017, 2019) with a simple physical model based on the tight coupling limit approximation (between DM and DR) in analogy with the photon-baryon plasma (see e.g Hu & Sugiyama 1996). This attempt is described in Appendix 2.6.2. Although we find that this approximation is not accurate enough, particularly in describing the damping envelope of the DAOs, it does provide relevant insights into the relevance of the sound horizon scale and the DM decoupling epoch as the physical quantities behind the DAO features. Therefore, we decided to try a phenomenological approach based on these quantities. To test this approach we explore a set of 84 ETHOS models as described above with the set of values: $\{n=(3-15,20), \log_{10}(a_n)=(0,1,2,3,5,7)\}$, and fixing $\alpha_{l\geq 2}=3/2$.

The sound horizon scale $r_{\rm DAO} \approx c_{\rm s}\eta_\chi$, where $c_{\rm s}$ is the DM sound speed and η_χ is

---

[3] We refer specifically to the case where DR-DR interactions are irrelevant.

[4] More specifically, α_l is the ratio between the opacity of the $l^{\rm th}$-moment to that of the dipole moment of the DR multipole hierarchy given by the angular dependence of DM-DR scattering.

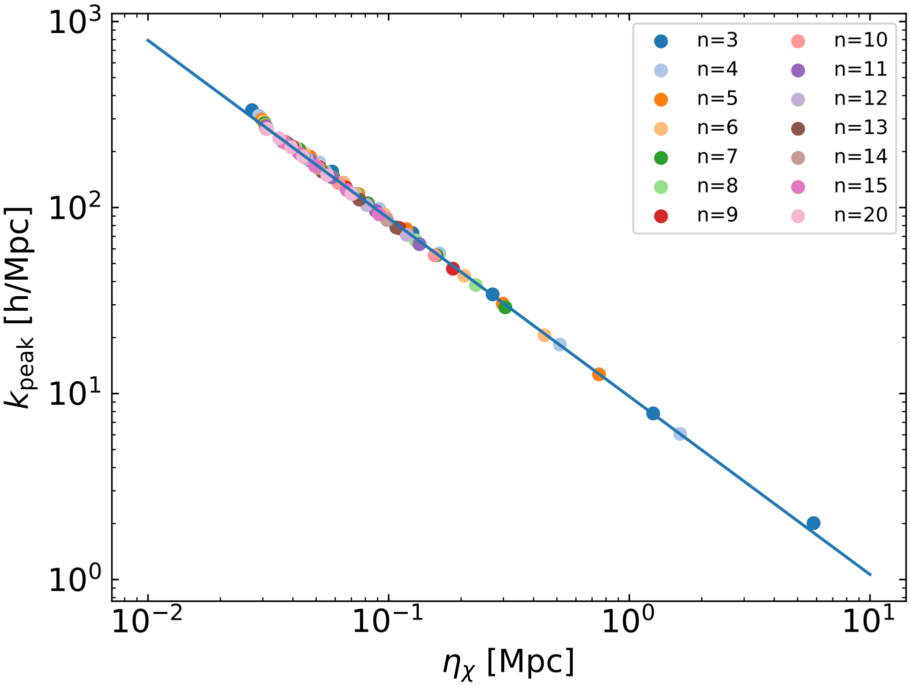

Figure 2.3.6. The position of the first DAO peak $k_{\rm peak}$ *correlates strongly with the time of DM decoupling* η_χ *defined by* $\int_{\eta_\chi}^{\eta_0} -\dot{\kappa}_\chi d\eta = 1$*. Each symbol correspond to a different ETHOS model within a grid of* $\{n, a_n\}$ *values and fixing* $\alpha_{l\geq 2} = 3/2$*. Models with a fixed n but different* a_n *are represented with the same colour as given in the legend. The blue line is a power law fit to the correlation,* $k_{\rm peak} = 9.37\,(\eta_\chi/{\rm Mpc})^{-0.97}\,h\,Mpc^{-1}$.

the conformal time of DM decoupling defined by

$$\int_{\eta_\chi}^{\eta_0} -\dot{\kappa}_\chi d\eta = 1 \tag{2.5}$$

with $\dot{\kappa}_\chi$ being the DM drag opacity due to the DM-DR interactions, should give the largest scale affected by acoustic oscillations and thus should be connected to $k_{\rm peak}$. We found this to be almost accurate, with only a slight deviation from a linear relation (see Fig. 2.3.6):

$$k_{\rm peak} = 9.37 \left(\frac{\eta_\chi}{\rm Mpc}\right)^{-0.97} h\,{\rm Mpc}^{-1}, \tag{2.6}$$

i.e., $k_{\rm peak}$ is given by the sound horizon scale at the time of kinetic decoupling, with just a minor modification. Notice that since a_n is connected to the decoupling time (see Eq. 2.19 and 2.20), then this relation implies a direct connection between $k_{\rm peak}$ and a_n.

On the other hand, we find that the damping of the first DAO, and therefore the parameter $h_{\rm peak}$, is mostly controlled by the DM mean free path (due to the DM-DR

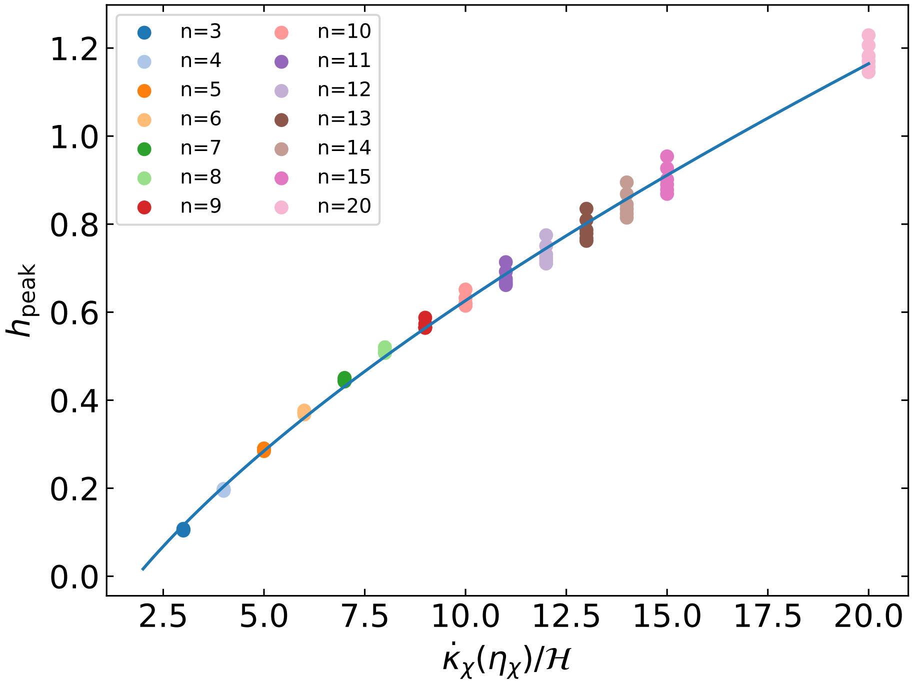

Figure 2.3.7. The amplitude of the first DAO peak h_{peak} scales with the ratio of the DM drag opacity to the Hubble rate at the time of DM decoupling $\dot{\kappa}_\chi(\eta_\chi)/\mathscr{H}(\eta_\chi) = n$. The colours are the same as in Fig. 2.3.6. The blue line is given by $h_{\text{peak}} = 0.21\left(\frac{\dot{\kappa}_\chi(\eta_\chi)}{\mathscr{H}(\eta_\chi)}\right)^{0.66} - 0.31$.

interactions) at DM decoupling, which is given by the inverse of $\dot{\kappa}_\chi(\eta_\chi)$ (see Fig. 2.3.7):

$$h_{\text{peak}} = 0.21\left(\frac{\dot{\kappa}_\chi(\eta_\chi)}{\mathscr{H}(\eta_\chi)}\right)^{0.66} - 0.31, \tag{2.7}$$

where the relevant quantity is actually the ratio $\dot{\kappa}_\chi(\eta_\chi)/\mathscr{H}(\eta_\chi)$, with $\mathscr{H}$ being the Hubble rate (relative to the conformal time). This ratio is actually equal to the ETHOS parameter n (see Appendix 2.6.2). Thus, for $n \gg 1$, the DM drag visibility function $\dot{\kappa}e^{-\kappa_\chi}$ is narrower, which implies a faster decoupling time scale; indeed, the DM-DR plasma is clearly in the tightly coupled regime at η_χ, and thus the damping by DR diffusion is only significant for the smallest scales. On the contrary when $n \gtrsim 1$, the DM drag visibility function is broader so that the timescales for decoupling (which occurs mostly in the weakly coupled regime) are larger. Thus, the DM mean free path at decoupling is relatively large and DM can diffuse substantially, lowering the value of h_{peak}. While for $n \lesssim 9$, n is the only factor in determining h_{peak}, that is not true anymore for $n \geq 10$. In the latter case, h_{peak} spreads around the fit in Fig. 2.3.7 depending on the value of a_n, the larger a_n the larger h_{peak}.

Finally, the ratio of the first two DAO peaks h_{peak}/h_2, is connected to the ratio of

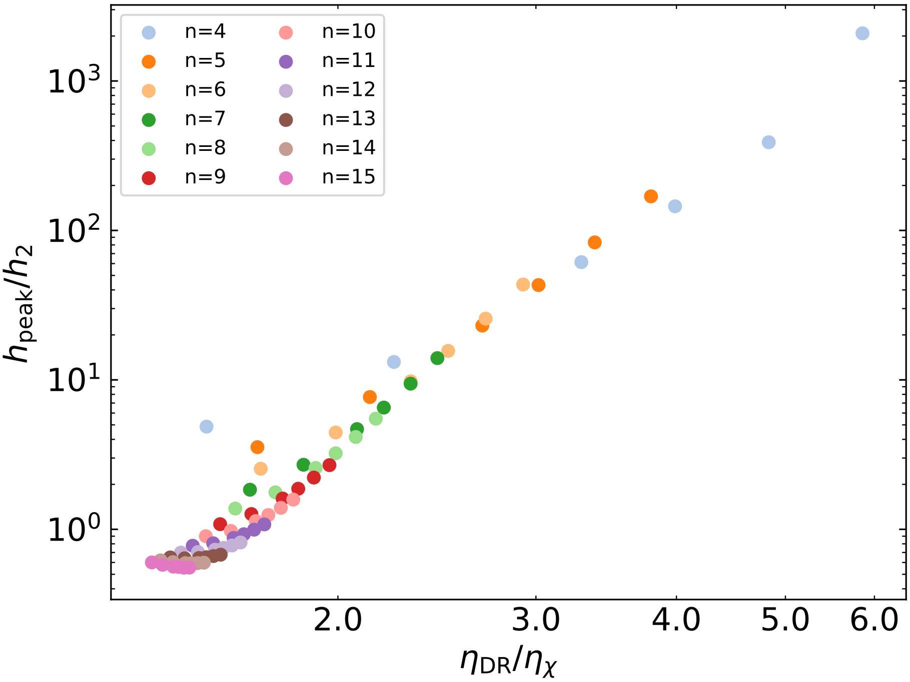

Figure 2.3.8. The ratio of the first two DAO peaks scales with the ratio of the DR and DM decoupling times. The relation is nearly a parabola, with the parameters depending on the specific value of n. The colours are the same as in Fig. 2.3.6.

the DR to DM decoupling times $\eta_{\mathrm{DR}}/\eta_\chi$ (see Fig. 2.3.8), where η_{DR} is the conformal time of DR decoupling defined by

$$\int_{\eta_{\mathrm{DR}}}^{\eta_0} -\dot{\kappa}_{\mathrm{DR}} d\eta = 1 \tag{2.8}$$

with $\dot{\kappa}_{\mathrm{DR}}$ being the DR opacity to DM scattering. This relation can be approximated by a parabola, but with different parameters for different values of n.

There are two additional features in Fig. 2.3.7 that we highlight: (i) for models with very large $n \gtrsim 15$ it is possible to have $h_{\mathrm{peak}} > 1$ (see also Kamada & Takahashi, 2018; Ando et al., 2019); (ii) for models with $n \gtrsim 10$, the value of h_{peak} depends not only on $n = \dot{\kappa}_\chi(\eta_\chi)/\mathscr{H}(\eta_\chi)$, but also on the specific value of a_n (the larger n, the stronger the dependence). For the latter models it is also true that the second DAO peak is significantly larger than the first (see Fig. 2.3.8). We decide to exclude these models, i.e. those with $n \gtrsim 10$, from our analysis for two reasons: (i) they would have a power spectrum that exceeds that of CDM at certain scales, and (ii) the parameter h_2 would no longer be a secondary parameter in determining the non-linear power spectrum. We notice however, that there is potentially interesting phenomenology in these models, which we leave for a future work. With this exclusion and using the strong correlations seen in Figs. 2.3.6-2.3.8, we have accomplished our goals at least for the regime of

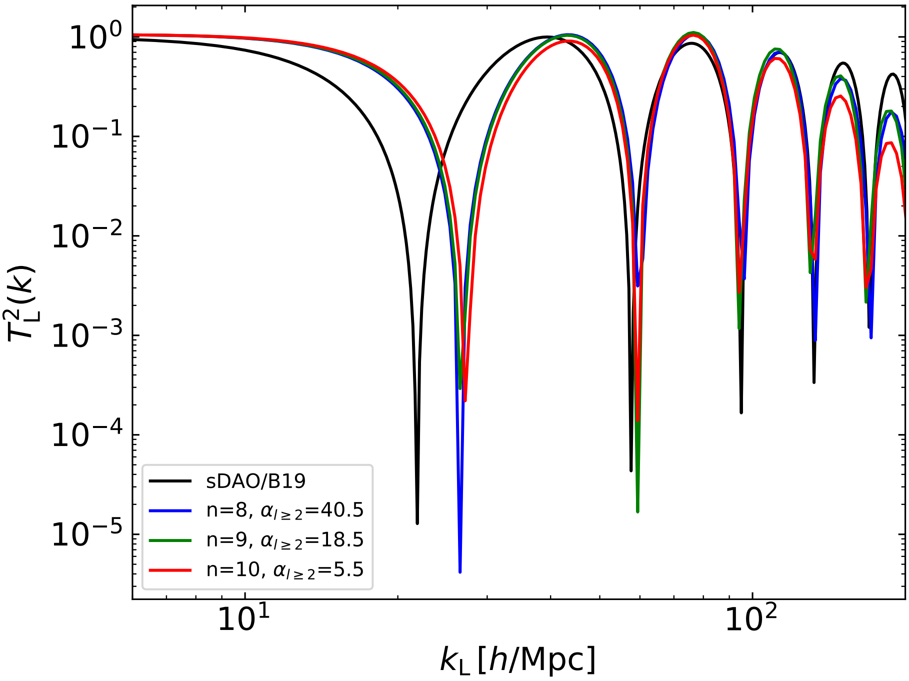

Figure 2.3.9. Power spectra of strong DAO models where the first few DAO peaks have roughly the same height, reaching the CDM amplitude. The black line is the sDAO model from Bose et al. (2019), while the lines with different colours are for ETHOS models. It is possible to accomplish this behaviour within the ETHOS framework for a fixed value of n by systematically changing the value of $\alpha_{l\geq 2}$ for $n = 8, 9, 10$.

weak to moderately strong ($h_{\rm peak} \sim 0.6$) DAO models, i.e., we have found a way to connect the parameters $h_{\rm peak}$ and $k_{\rm peak}$ in our parametrization to n an a_n, respectively, in the original ETHOS framework, as well as to connect $h_{\rm peak}$, $k_{\rm peak}$ and h_2 to physical quantities that are responsible for the DM-DR decoupling.

Within our parametrization, it is possible to include models with stronger DAO features ($h_{\rm peak} \sim 1$), but that do not exceed greatly the CDM power spectrum, such as the benchmark sDAO model analysed in Bose et al. (2019) to show the distinct features this type of models leave in the Ly-α forest 1D flux spectrum. To do so, we need to change the value of $\alpha_{l\geq 2}$ from $3/2$ to a value of $\mathcal{O}(10)$ for $n \sim 9$. In this way, we can create power spectra that have strong DAO features but without having $h_{\rm peak}/h_2 < 1$ (see Fig. 2.3.9). This modification breaks the relations for the peak heights in Fig. 2.3.7 and 2.3.8, while the one for $k_{\rm peak}$ stays unchanged. We used these models with increased $\alpha_{l\geq 2}$ for our strong DAO cases.

h_{peak}	h_2	τ	σ	β	d
0.0	0.0	0.0	0.0	2.24	3.0
0.2	(0.067,-0.086,0.0011)	0.34	0.23	3.1	2.93
0.4	(0.221,-0.025,0.0129)	0.34	0.22	3.61	2.61
0.6	(0.572,-0.008,0.1490)	0.38	0.2	3.91	2.44
0.8	0.88	0.55	0.2	4.0	2.46
1.0	1.08	0.67	0.2	4.05	2.5

Table 2.3.2. Parameters used to construct the linear power spectra used in our simulations (Eq. 2.3) as a function of the amplitude h_{peak} of the first DAO peak relative to CDM. For h_{peak} in the range $[0.2, 0.6]$, the amplitude of the second DAO peak, h_2 is given by $h_2 = A\exp(Bk_{\text{peak}}) + C$, where the (A, B, C) values are given in the column.

2.3.2 Final parameter space

Given all previous considerations, we work with a 2D parameter space with a set $\{k_{\text{peak}}, h_{\text{peak}}\}$. We explore simulations within a range of k_{peak} between 35-300 $h\,\text{Mpc}^{-1}$and h_{peak} from 0 to 1. Notice that $h_{\text{peak}} = 0$ corresponds to thermal WDM models. This parameter space is covered with a grid of 52 simulations, spaced by 0.2 intervals in h_{peak} and equidistant in $\log(k_{\text{peak}})$ on two separate intervals: [35,100] $h\,\text{Mpc}^{-1}$and [100,300] $h\,\text{Mpc}^{-1}$. The parameters of our full parametrization (Eq. (2.3); Table 2.3.2) have been calibrated to the linear power spectra of the corresponding ETHOS model obtained with the Boltzmann solver CLASS. We find that for all models, the parameters τ, σ, β and d only depend on h_{peak}. Following the results in Fig. 2.3.7, we can relate $n = 4$ for $h_{\text{peak}} = 0.2$, $n = 6$ for $h_{\text{peak}} = 0.4$, and $n = 9$ for $h_{\text{peak}} = 0.6$; all of these with constant values $\alpha_{l\geq 2} = 3/2$. The height of the second DAO peak, h_2, depends on k_{peak} (through the correlation seen in Fig. 2.3.8) for all these models, but we find that can be modeled with a simple exponential function $h_2 = A\exp(Bk_{\text{peak}}) + C$ (see Table 2.3.2 for the values of A, B, and C). For $h_{\text{peak}} = 0.8, 1.0$, n was fixed to 9, but $\alpha_{l\geq 2}$ had to be increased to a value in between $\sim 10-30(30-100)$ for $h_{\text{peak}} = 0.8(1.0)$, depending on the value of k_{peak}, in order to reach the desired value of h_{peak} without having a very dominant second DAO peak, which in this case is independent of k_{peak}. The final parameters used in our simulations for a given h_{peak} are given in Table 2.3.2.

2.4 Results

To characterise the differences between DM models and find out which features survive the non-linear evolution, we look at the matter power spectrum and the halo mass function at high redshift $z > 5$.

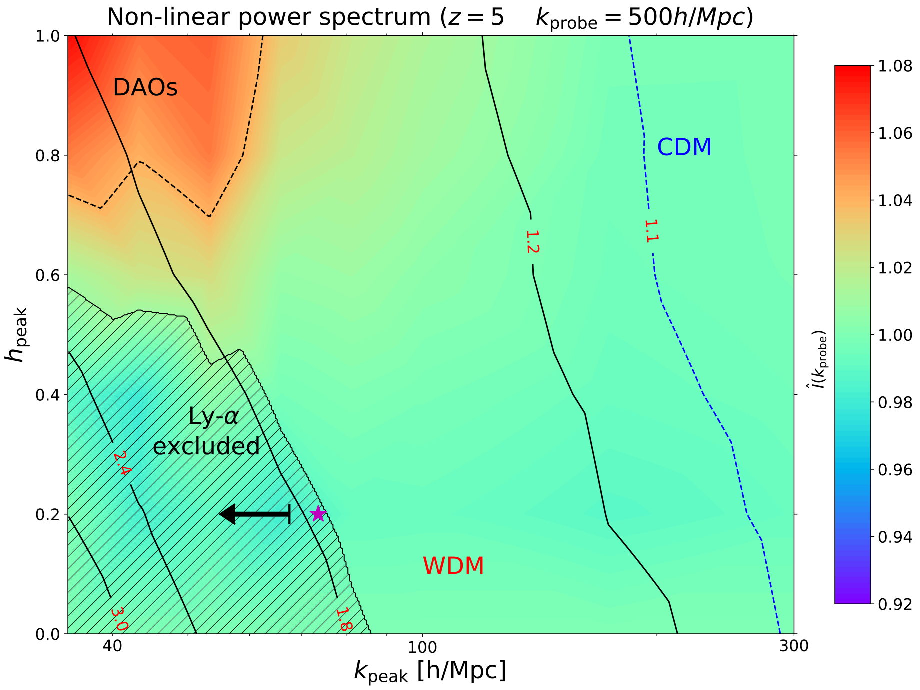

Figure 2.3.10. Division of structure formation models in the effective parameter space $h_{\rm peak}$ and $k_{\rm peak}$ (see Eq. 2.3) according to their power spectra at $z = 5$ for $k_{\rm probe} = 500\,h\,Mpc^{-1}$. The black contour lines correspond to the ratio $R(k_{\rm probe})$ of the CDM power spectrum relative to a given model (see Eq. 2.9). The colour scale shows the re-normalized values of the integrated quantity $\hat{I}(k_{\rm probe})$ (see Eqs. 2.10–2.11), where a value of 1 corresponds to areas that are degenerate with WDM: models with $\hat{I} = 1$ have the same power spectrum at $z = 5$ as the WDM model at the same contour value of R. The black dashed line ($\hat{I}(k_{\rm probe}) > 1.04$) on the upper left encompass the area where DAO features survive until $z = 5$. The region to the right of the blue dashed line, $R(k_{\rm probe}) = 1.1$, can be considered as nearly indistinguishable from CDM up to $k_{\rm probe}$. The hashed region on the lower left encompasses the area of models that are degenerate with a thermal WDM of mass <3.6 keV, which has been ruled out by Lyman-α forest data (Murgia et al., 2018), and the arrow indicates the upper bound from the Lyman-α analysis of $h_{\rm peak} = 0.2$ ($n = 4$) models in Archidiacono et al. (2019). The ETHOS-4 model used as a benchmark in Vogelsberger et al. (2016) is indicated by the purple star.

2.4.1 Matter power spectrum

We evaluate the power spectrum of the simulations at different scales $k_{\rm probe}$ between $10\,h\,{\rm Mpc}^{-1}$and $500\,h\,{\rm Mpc}^{-1}$. This range roughly covers the relevant range where new DM physics can play a role in the physics of galaxies: the larger scales are bounded by current constraints over deviations from CDM (for instance from Ly-α forest measurements e.g. Iršič et al. 2017), while the smaller scales are bounded by the minimum scales at which galaxies can form, where we use the atomic cooling limit as a reference. At $z = 5$, the virial mass of a halo corresponding to a non-linear scale of $k = 500\,h\,\mathrm{Mpc}^{-1}$is $\sim 5 \times 10^7\,\mathrm{M}_\odot\,h^{-1}$, which is just below the atomic cooling limit at this redshift, $\sim 10^8\,\mathrm{M}_\odot\,h^{-1}$.

To quantify the difference of a given model with respect to CDM, we define two different diagnostics: (i) we compute the ratio $R(k_\mathrm{probe})$ of the power spectra with respect to CDM at k_probe:

$$R(k_\mathrm{probe}) = \frac{\Delta^2_\mathrm{CDM}(k_\mathrm{probe})}{\Delta^2(k_\mathrm{probe})}, \tag{2.9}$$

and consider $R(k_\mathrm{probe}) \leq 1.1$ as essentially indistinguishable from CDM (since we set our convergence goal to 5%; see Appendix 2.6.1); (ii) to distinguish different non-cold DM models, $R(k_\mathrm{probe})$ is not sufficient enough, as two models with the same ratio at a given k can have different behaviour on larger scales. To capture this with a single number, we define the following dimensionless integrated quantity[5]

$$I(k_\mathrm{probe}) = \frac{\displaystyle\int_{k_\mathrm{min}}^{k_\mathrm{probe}} \left(\frac{\Delta^2_\mathrm{CDM}(k)}{\Delta^2(k)}\right)^2 d\ln k}{\ln\left(k_\mathrm{probe}/k_\mathrm{min}\right)}, \tag{2.10}$$

where we choose $k_\mathrm{min} = 10\,h\,\mathrm{Mpc}^{-1}$since, as we mentioned above, the models we are interested in have the same power at this scale. By construction, a larger value of $R(k_\mathrm{probe})$ also results in a larger value of $I(k_\mathrm{probe})$, thus we need to normalize it to a reference case in order to define a comparative quantity across the different structure formation models. We choose the WDM case as the reference and normalize Eq. (2.10) for a given model by the value of $I(k_\mathrm{probe})$ of a WDM model with the same value of $R(k_\mathrm{probe})$:

$$\hat{I}(k_\mathrm{probe}) = \left(\frac{I(k_\mathrm{probe})}{I_\mathrm{WDM}(k_\mathrm{probe})}\right)_{R(k_\mathrm{probe})} \tag{2.11}$$

Defined in this way, for a fixed $R(k_\mathrm{probe})$, all models with $\hat{I}(k_\mathrm{probe}) = 1$ have a non-linear power spectrum at the given redshift which is essentially indistinguishable from a WDM of the same $R(k_\mathrm{probe})$, regardless of how different the linear power spectrum of these models is relative to WDM.

Figure 2.3.10 shows the results for our simulations for $z = 5$ and $k_\mathrm{probe} = 500\,h\,\mathrm{Mpc}^{-1}$in the leading order space of parameters k_peak and h_peak (Eq. 2.3). We recall that we ran 50 simulations for models within this parameter space, which are then used to bilinearly interpolate the values of $R(k_\mathrm{probe})$ and $\hat{I}(k_\mathrm{probe})$ between the simulated models (the grid described in Section 2.3.2) to fill in Fig. 2.3.10. The line contours show $R(k_\mathrm{probe})$, which increases from right to left with the models to the right of the blue dashed line ($R = 1.1$) being virtually indistinguishable from CDM, while those to the left become ever more divergent from CDM. The colour scale shows the value of $\hat{I}(k_\mathrm{probe})$ and therefore quantifies how different a model is compared with a WDM model that has

---

[5] Note that we use a quadratic dependence on the ratio $1/T^2_\mathrm{NL} = \Delta^2_\mathrm{CDM}/\Delta^2$ in the integrand in Eq. (2.10) instead of a linear one in order to enhance the difference between models.

the same value of R. Note that the WDM models in this plot lie at the bottom $k_{\rm peak}$ axis ($h_{\rm peak} = 0$). Using the connection between α and $m_{\rm WDM}$ from Viel et al. (2005)

$$\alpha = 0.049 \left(\frac{m_{\rm WDM}}{1\,{\rm keV}}\right)^{-1.11} \left(\frac{\Omega_\chi}{0.25}\right)^{0.11} \left(\frac{h}{0.7}\right)^{1.22} h^{-1}\,{\rm Mpc}, \tag{2.12}$$

we can compute $m_{\rm WDM}$ from $k_{\rm peak}$:

$$\frac{m_{\rm WDM}}{1\,{\rm keV}} = \left[0.050 \left(\frac{k_{\rm peak}}{h\,{\rm Mpc}^{-1}}\right) \left(\frac{\Omega_\chi}{0.25}\right)^{0.11} \left(\frac{h}{0.7}\right)^{1.22}\right]^{\frac{1}{1.11}} \tag{2.13}$$

For a value of $\hat{I}$ close to one (green colour), the model's power spectrum (up to $k_{\rm probe} = 500\,h\,{\rm Mpc}^{-1}$and at $z = 5$) will be indistinguishable from WDM, while larger values mean that the power spectrum shape for $k < k_{\rm probe}$ is truly distinct from WDM regardless of the value of $R(k_{\rm probe})$. The region in the top left where $\hat{I}(k_{\rm probe})$ has the largest values corresponds to models with strong DAO features ($h_{\rm peak} \gtrsim 0.7$ and $k_{\rm peak} \lesssim 65\,h\,{\rm Mpc}^{-1}$; labeled as DAOs) where the impact of the DAO features has not been erased by the non-linear evolution down to $z = 5$, and thus still leaves a signature in the power spectrum up to $k_{\rm probe}$. As is apparent most of the parameter space outside of the latter DAO region has values of $\hat{I}$ close to 1, and thus any models here are essentially degenerate (up to $k_{\rm probe} = 500\,h\,{\rm Mpc}^{-1}$and at $z = 5$) with a WDM model with the same value of R. This degeneracy is either caused by the non-linear evolution erasing the DAO features, especially for the weak DAO models in the lower part of the plot ($h_{\rm peak} \lesssim 0.6$), or because the DAO features appear at smaller scales $k > k_{\rm probe}$ than we are interested in (for models with $k_{\rm peak} \gtrsim 100\,h\,{\rm Mpc}^{-1}$). We remark that comparisons between WDM, wDAO and sDAO models have been done in the past using N-body simulations (see e.g. Buckley et al., 2014; Vogelsberger et al., 2016; Schewtschenko et al., 2015). In particular, Murgia et al. (2018) have shown that the presence of weak oscillations does not affect the scales probed by Lyman-α forest observations. However, we are showing with Fig. 2.3.10 that the degeneracies between DAO and WDM models extend to much smaller scales (including strong oscillations), and crucially, we introduce a quantity $\hat{I}(k_{\rm peak})$ that is a measure of the degree of degeneracy.

The hashed region on the lower left was constructed taking as a reference the constraints on the thermal WDM particle mass from current Lyman-α forest data from Murgia et al. (2018) using MIKE/HIRES data: $m_{\rm WDM} < 3.6$ keV (2σ C.L.). To do this, we follow the contour line corresponding to this WDM model, up to the value of $h_{\rm peak}$ that remains degenerate with this WDM model ($h_{\rm peak} \sim 0.47$) using a value of $\hat{I} = 1.01$ as the dividing threshold. We then continue the Lyman-α constraint line towards larger values of R along this $\hat{I}$ threshold. We remain within this threshold because we expect that beyond, the Lyman-α analysis based on WDM model would no longer be valid due to the impact of the DAO features. In this way, the hashed region on the lower left is our expectation for the exclusion region from Lyman-α data. Drawing this region more precisely would require a full analysis of the predictions of the 1D flux power spectrum within our framework. For the $h = 0.2$ ($n = 4$) case, this was done in Archidiacono et al. (2019) and we indicate their upper limit of $a_n\xi^4 < 30{\rm Mpc}^{-1}$ as an upper bound on $k_{\rm peak}$ with the arrow in Fig. 2.3.10. We notice that their direct constraint on the wDAO model

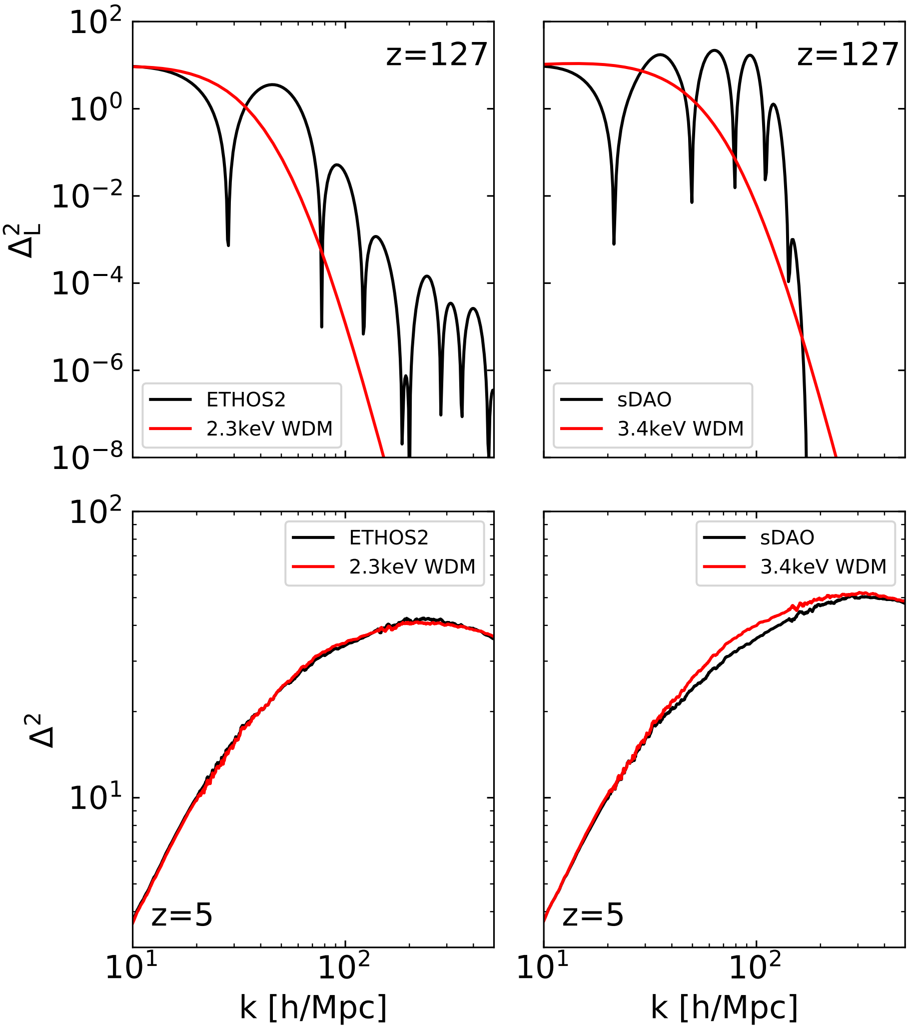

Figure 2.4.11. Comparison of the power spectra from a couple of DAO models and the comparable WDM models with $m = 2.3\,\mathrm{keV}$ *(left) and* $m = 3.4\,\mathrm{keV}$ *(right) at the initial conditions for our simulations (*$z = 127$ *top panels) and at the end of the simulations (*$z = 5$*; bottom panels). The left panels are for a wDAO model (ETHOS-2 in Vogelsberger et al. 2016), while the right panels are for an sDAO model. On the left, the different models have clearly different power spectra at the initial conditions, but become completely degenerate at* $z = 5$*. On the right panels on the other hand, the models remain different even at* $z = 5$*, despite having the same power at* $k = 500\,h\,Mpc^{-1}$*. This shows that only strong DAO models have truly distinct power spectra features relative to the WDM model.*

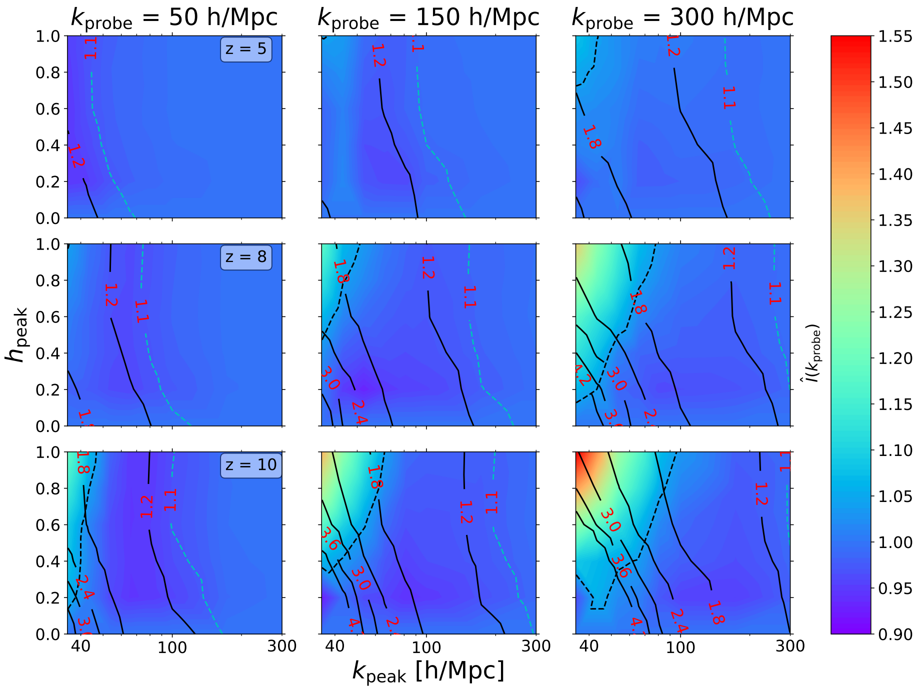

Figure 2.4.12. Division of structure formation models in the parameter space $h_{\rm peak}$ and $k_{\rm peak}$ of Eq. (2.3) *based on the power spectra at different scales $k_{\rm probe}$ and redshifts. See Fig. 2.3.10 for a description of the different elements in this plot.*

is close to our expectation from models that are degenerate with the WDM constraints (envelope of the hashed region in Fig. 2.3.10). However, there are a few factors that are likely responsible of the mismatch. Most notably, our estimate is based on the 3D matter power spectrum, while the constraint from Archidiacono et al. (2019) is derived from the 1D flux power spectrum. Finally, we have indicated with a purple star symbol the location of the benchmark ETHOS-4 model defined in Vogelsberger et al. (2016), which is seemingly barely allowed within the Lyman-α exclusion region. This is a point that was noticed in Vogelsberger et al. (2016) where the linear power spectrum of this model was considered to have a WDM equivalent (in terms of the cutoff) with a thermal particle mass of $m_{\rm WDM} = 3.66$ keV. Our results confirm this correspondence and indicate that the ETHOS-4 model has a power spectrum that is nearly indistinguishable at $z = 5$ from a WDM model with $m_{\rm WDM} \sim 3.5$ keV.

From the slopes of the contour lines in Fig. 2.3.10, we can see that the degeneracies are not between weak DAO and WDM models with the same power spectrum cutoff (i.e. $k_{\rm peak}$), but that the additional power coming from the DAOs at scales smaller than $k_{\rm peak}$ still matters and can only be accounted for by WDM models with a cutoff at smaller scales. This is shown more clearly on the left panels of Fig. 2.4.11, which show the power spectrum for a weak DAO model (ETHOS-2 in Vogelsberger et al. (2016); $h_{\rm peak} = 0.2, k_{\rm peak} = 46.5\, h\,{\rm Mpc}^{-1}$) and a WDM model with $m = 2.3\,{\rm keV}$ at $z = 127$

(top) and $z = 5$ (bottom). Despite having distinct linear power spectra, these models are nearly degenerate at $z = 5$; they have both the same value of $R(k_{\rm probe} = 500\,h\,{\rm Mpc}^{-1})$ and $\hat{I}(k_{\rm probe} = 500\,h\,{\rm Mpc}^{-1})$. Moreover, in order to match the weak DAO model, the WDM model needs to sustain more power at larger scales in the linear power spectra to compensate for its steeper cutoff. The right panels of Fig. 2.4.11 show the distinct behaviour of a strong DAO model. In this case the additional power from the secondary DAO peaks have an impact in the non-linear power spectrum down to $z = 5$ that cannot be replicated by a WDM model: in order for the sDAO and the WDM models to have the same power at $k_{\rm probe} = 500\,h\,{\rm Mpc}^{-1}$, the WDM linear cutoff needs to occur at significantly smaller scales than the one for the sDAO model. Thus, the WDM model has more power at intermediate scales by $z = 5$.

Figure 2.4.12 is equivalent to Fig. 2.3.10 but at different redshifts ($z = 5, 8, 10$) and scales $k_{\rm probe}$ ($50, 150, 300\,h\,{\rm Mpc}^{-1}$). Note that we use the same colour scale for $\hat{I}$ for all cases in order to ease the comparison between the different panels. Focusing on the values of $R(k_{\rm probe})$ represented by the contour lines first, we can see that they shift towards larger $k_{\rm peak}$ as the redshift increases (from top to bottom) or as $k_{\rm probe}$ increases (from left to right). The former trend is expected since at higher redshift the clustering properties have departed less from the linear evolution, where the different DM models differ the most from CDM at all scales. The latter trend is simply due to the damping envelope in the different DM models, which produces an effective cutoff towards smaller scales, and thus R will naturally be larger towards larger values of $k_{\rm probe}$. Looking at the colour contours, it is apparent that at lower redshifts and/or small $k_{\rm probe}$, none of the DAO models explored are clearly distinguishable from WDM models (i.e. the value of $\hat{I}$ is too close to 1) with the DAO region we highlighted in Fig. 2.3.10 (black dashed line) essentially disappearing in the top left panels. We emphasize that this is independent of the strength of the DAO features in the linear power spectrum. The opposite happens as $k_{\rm probe}$ and/or the redshift increases, the DAO region increases to cover a larger region of the parameter space. This is because at higher redshift the WDM and DAO models become ever more divergent since there is less time to erase the DAO features and equalize the power at all scales. At smaller scales, the initial difference in power was larger and needs more time to get erased, and additionally for large $k_{\rm probe}$, the signal over a wide range of k-modes is accumulated.

2.4.2 Halo mass function

The halo mass function provides another relevant measure to characterise structure formation models. It is also more sensitive to the differences across DM models in the linear regime than the non-linear power spectrum since it preserves a stronger memory of the history into collapsed haloes across time (for alternative DM models, including those with DAOs, this was studied e.g. in Leo et al., 2018b).

Figure 2.4.13 shows an example of a comparison of the halo mass function at $z = 5$ between the CDM model, and a WDM, weak DAO and strong DAO model with the same cutoff scale in the initial power spectrum. It can be seen that even though the three models are designed to start deviating from CDM at roughly the same mass, the slope

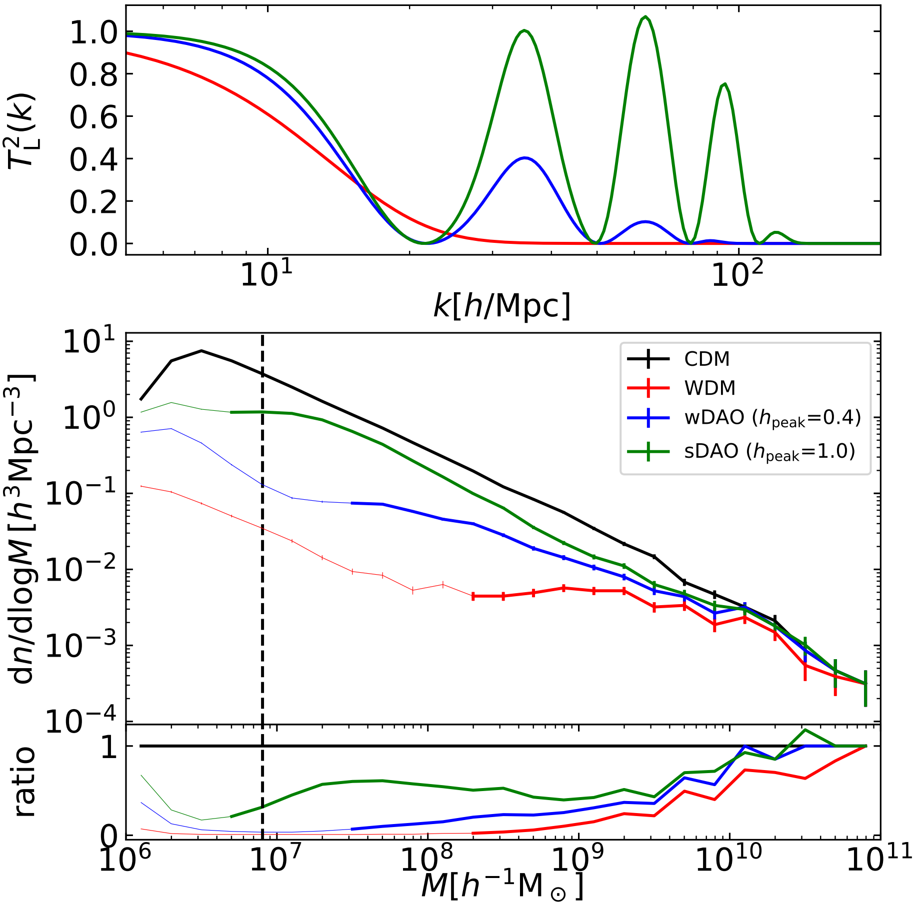

Figure 2.4.13. Top panel: Initial transfer function $T_L^2(k)$ for examples of the WDM (red), weak DAO (blue) and strong DAO (green) models; all of which have the same value of $k_{\mathrm{peak}} = 35\,h\,\mathrm{Mpc}^{-1}$. Bottom panel: Halo mass function at $z = 5$ for the models above and CDM (black). The error bars denote Poisson counting errors. Masses below the limiting mass (see Wang & White 2007) are indicated by thin lines. To the left of the vertical dashed line haloes have less than 100 simulation particles.

of the halo mass function at smaller scales is very different. The halo mass function for the WDM model stays roughly flat towards the left of the cutoff mass until the slope rises again artificially due to the presence of spurious haloes caused by well-known discreteness effects in models with a primordial power spectrum cutoff. The limiting mass below which one can no longer trust the halo mass function is well described by a formula that depends on the cutoff scale of the model and the spatial resolution of the simulation (see Wang & White, 2007). As can be seen in Fig. 2.4.13, this formula

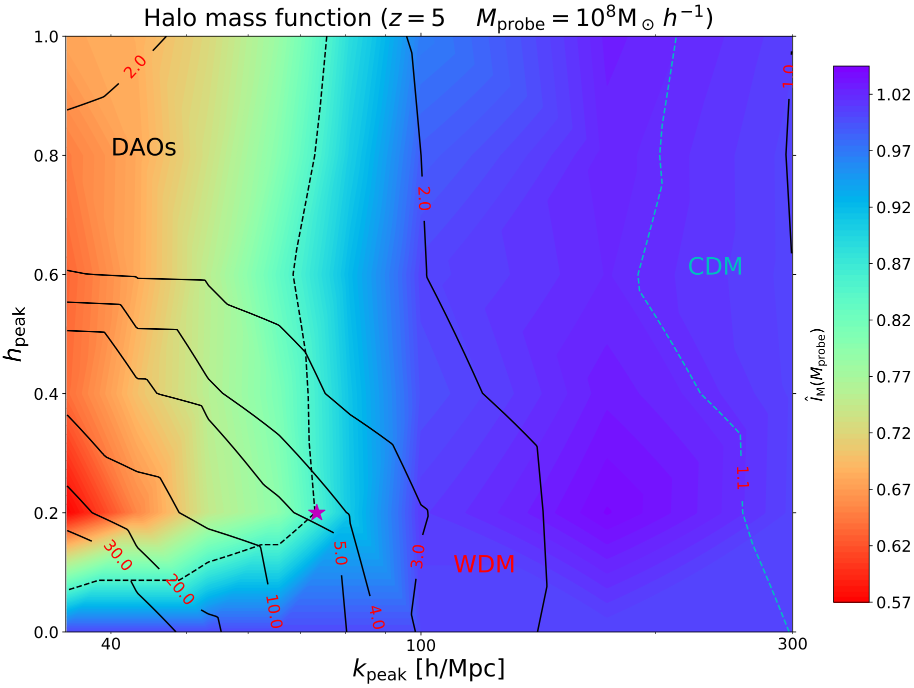

Figure 2.4.14. Division of structure formation models in the effective parameter space $h_{\rm peak}$ *and* $k_{\rm peak}$ *(see Eq. 2.3) according to their halo mass functions at* $z = 5$ *for* $M_{\rm probe} = 10^8 {\rm M}_\odot/h$*. The contour lines correspond to the ratio* $R_{\rm M}(M_{\rm probe})$ *of the CDM halo mass function relative to a given model at* $M_{\rm probe}$ *(*$\hat{I}_{\rm M}(M_{\rm probe})$*; see Eq. 2.14). The colour scale shows the number of haloes with* $M > M_{\rm probe}$ *normalized to a WDM model with the same number of haloes at* $M = M_{\rm probe}$ *(see Eq. 2.15), where a value of 1 corresponds to areas that are degenerate with WDM. The black dashed line (*$\hat{I}_{\rm M}(M_{\rm probe}) < 0.85$*) on the left encompasses the area where DAO features survive until* $z = 5$*. The region to the right of the blue dashed line,* $R_{\rm M}(M_{\rm probe}) = 1.1$*, can be considered as nearly indistinguishable from CDM down to* $M_{\rm probe}$*. The ETHOS-4 model used as a benchmark in Vogelsberger et al. (2016) is indicated by the purple star.*

describes reasonably well the scale at which spurious haloes start to dominate not only for the WDM model, but also for the weak and strong DAO models. For the CDM model, the mass function can be trusted to even lower masses until the simulation particle number is too low to resolve haloes (typically $\lesssim$100). On the other hand, the additional small scale power of the DAO models keeps the slope of the halo mass function steeper at small masses, relative to the WDM case, and for strong DAOs, the halo mass function is even parallel to the CDM case, albeit with a reduced normalization.

Figure 2.4.14 shows the structure formation models in the parameter space ($h_{\rm peak}$,$k_{\rm peak}$) as characterised by the halo mass function of our simulations at $z = 5$ and at a halo mass of $M_{\rm probe} = 10^8 {\rm M}_\odot/h$. This figure is analogous to Fig. 2.3.10, with the contours showing in this case the ratio between the CDM halo mass function and that of a given model at $M_{\rm probe} = 10^8 {\rm M}_\odot/h$:

$$R_{\rm M}(M_{\rm probe}) = \left(\frac{MF_{\rm CDM}}{MF}\right)_{M_{\rm probe}} \tag{2.14}$$

where $MF = dn/d{\rm log}M$ is the differential halo mass function. On the other hand, the colour contours represent the number of haloes with $M > M_{\rm probe}$ normalized to the number of haloes of a WDM model with the same ratio at $M_{\rm probe}$:

$$\hat{I}_{\rm M}(M_{\rm probe}) = \left(\frac{I_{\rm M}(M_{\rm probe})}{I_{\rm M,WDM}(M_{\rm probe})}\right)_{R_{\rm M}(M_{\rm probe})} \tag{2.15}$$

where

$$I_{\rm M}(M_{\rm probe}) = \int_{M_{\rm probe}}^{M_{\rm max}} MF d{\rm log}M \tag{2.16}$$

with $M_{\rm max} = 10^{11}{\rm M}_\odot/h$ being the maximum mass for which we can measure the halo mass function. We can see that the models with $k_{\rm peak} > 100\,h\,{\rm Mpc}^{-1}$are nearly indistinguishable from the corresponding WDM model (that are lying at the same contour line), since in this case the cutoff in the halo mass function is so close to $M_{\rm probe}$ that the different models (irrespective of the value of $h_{\rm peak}$) do not have very different slopes for their halo mass functions yet and thus, they all look alike. On the contrary, for $k_{\rm peak} < 100\,h\,{\rm Mpc}^{-1}$, the DAO models have halo mass functions with slopes that are clearly steeper (and thus distinguishable) than that of the WDM model below the cutoff mass. Therefore, in order for the corresponding WDM model to lie on the same contour line (i.e. to have the same halo mass function at $M_{\rm probe}$), it needs to have a cutoff scale at a relatively smaller mass (larger $k_{\rm peak}$), and thus will necessary have more haloes with $M > M_{\rm probe}$ than the DAO model (see Eq. 2.15). We observe that for $k_{\rm peak} \lesssim 100\,h\,{\rm Mpc}^{-1}$, the slope in the halo mass function towards smaller masses is related to $h_{\rm peak}$, which can be seen by looking at how the contour lines bend ever more sharply towards lower values of $k_{\rm peak}$ as $h_{\rm peak}$ increases, eventually becoming nearly flat for $h_{\rm peak} \sim 0.6$ at $k_{\rm peak} \sim 30-60\,h\,{\rm Mpc}^{-1}$. This implies that for these models, the actual mass cutoff (given by $k_{\rm peak}$) does not matter any longer since they all have the same mass function at $10^8{\rm M}_\odot h^{-1}$ haloes. Naturally, these models are still distinguishable since they have different halo abundances at larger masses.

In contrast to the division of structure formation models based on the non-linear power spectrum (Fig. 2.3.10), in this case represented by the halo mass function, the distinctive DAO region (black dashed line in Fig. 2.4.14) occupies a larger region of the parameter space, reaching into the regime of the weak DAO models. For instance, the red star in Fig. 2.4.14 corresponds to the ETHOS-4 model used in Vogelsberger et al. (2016) and it appears at the border of our definition of the DAO structure formation region. It is thus clear that even though weak DAO models are degenerate with WDM models in their non-linear power spectrum, this degeneracy is broken for the halo mass function.

At different redshifts (Fig. 2.4.15), the DAO region (black dashed line) remains almost unchanged (shrinking slightly at high redshift); the same is true for the CDM-like region (blue dashed line). It is only the contour lines of constant $R_{\rm M}$ that change across redshift, with the ratio of the halo mass function at $10^8 \mathrm{M}_\odot/h$ becoming larger at higher redshift for all the region below $k_{\rm peak} \lesssim 100\,h\,\mathrm{Mpc}^{-1}$. Therefore, the halo mass function provides a diagnostic to classify structure formation models that is less susceptible to being erased by the non-linear evolution than the power spectrum. A more detailed analysis of the halo mass function for the DAO models studied here will be presented in the future.

2.5 Conclusions

There are multiple ways in which non-standard DM physics can introduce a cutoff in the linear matter power spectrum whose shape can range from an exponential featureless free-streaming collisionless damping (as in thermal WDM models) to a shallower collisional damping driven by DM-dark radiation interactions with strong DAOs. The effective theory of structure formation (ETHOS), introduced in Cyr-Racine et al. (2016) aims at connecting the particle physics parameters of a variety of DM models into effective parameters that characterise the linear power spectrum. In this way, DM particle models can be classified in terms of a set of parameters that fully describe the linear power spectrum, particularly the characteristics of the small-scale cutoff and DAOs. It is however, not trivial to characterise the signature that these different departures from the linear CDM power spectrum leave in the non-linear regime of structure formation. It may indeed be possible that the gravitational coupling between different scales erases features like the DAOs making all models essentially indistinguishable from the standard WDM cutoff at the scales that are relevant for galaxy formation and evolution[6] ($10\,h\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 500\,h\,\mathrm{Mpc}^{-1}$). In this work we address this question by performing a large number of cosmological simulations within the ETHOS framework. Our goal is to define a reduced set of simple yet physically motivated parameters that allow to distinguish DM models based on differences in how structure formation proceeds (at the scales relevant for the physics of galaxies). In this first work, we have concentrated on DM-only simulations at high redshift ($z \geq 5$). Our analysis and main results can be summarised as follows:

- We have implemented a zoom-in simulation technique to efficiently cover a wide range of scales ($0.2 - 500\,h\,\mathrm{Mpc}^{-1}$) and accurately reconstruct the (average) matter power spectrum in this range (see Fig. 2.2.3). The computational cost of this method is significantly less than a uniform simulation with equivalent range (see Table 2.2.1).
- We introduced a new analytic formula (Eq. (2.3), Fig. 2.3.5) to describe the linear transfer function (relative to CDM) of models with a primordial cutoff,

---

[6] Relevant departures from CDM are bounded at large scales by current constraints based on e.g. observations of the Lyman-α 1D flux spectrum, and at small scales by the suppression of galaxy formation below the atomic cooling limit.

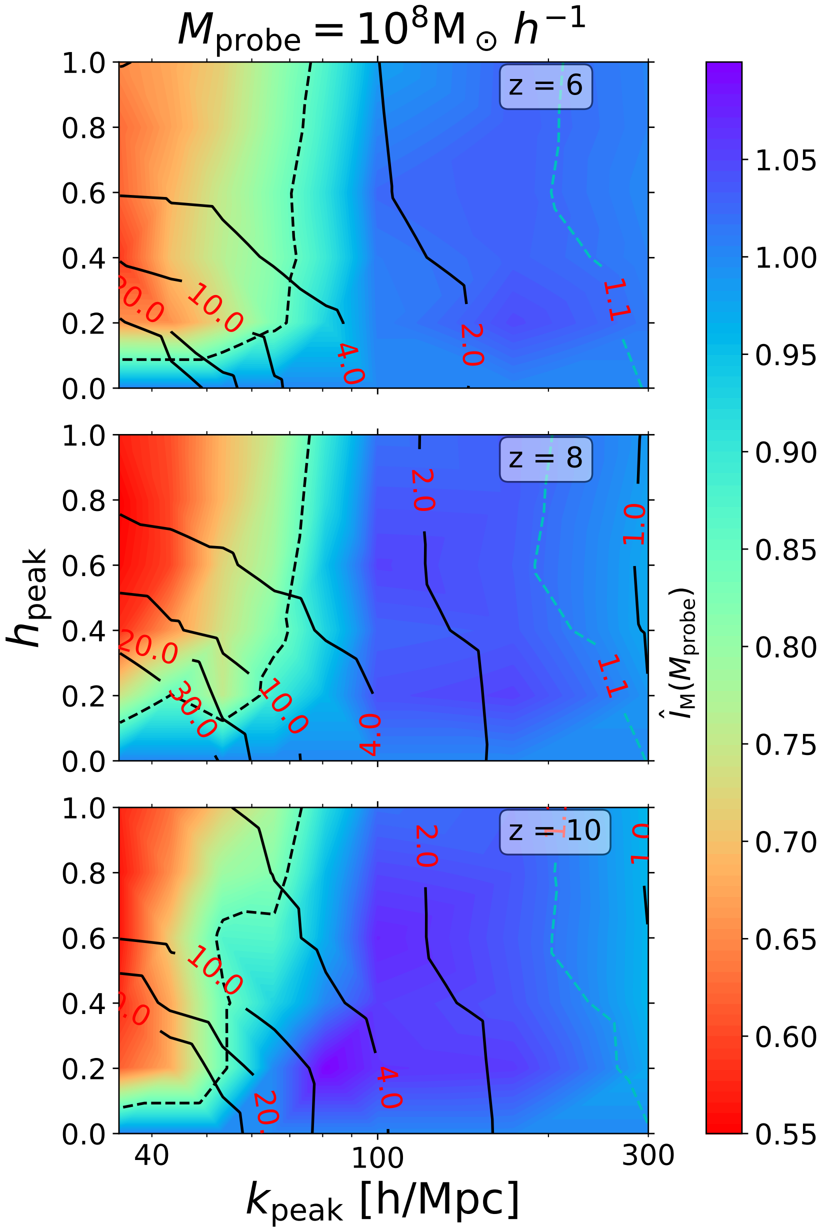

Figure 2.4.15. Division of structure formation models in the parameter space $h_{\rm peak}$ and $k_{\rm peak}$ of Eq. (2.3) based on the halo mass functions at $M_{\rm probe} = 10^8 {\rm M}_\odot/h$ and different redshifts. See Fig. 2.4.14 for a description of the different elements in this plot.

which accommodates both WDM and models with DAOs. This formula is accurate enough to reproduce the non-linear power spectrum for the scales of interest compared to the full calculation with a Boltzmann code (see Fig. 2.3.4). Crucially, only two free parameters in this formula, $\{h_{\rm peak}, k_{\rm peak}\}$ the amplitude and scale of the first DAO peak, are sufficient to characterise non-linear structure formation at high-redshift for WDM[7] and a large class of ETHOS models with DAOs. Moreover, we found a simple physical interpretation for these two main parameters where $k_{\rm peak}$ is connected to η_χ the time of DM decoupling from the DR (see Fig. 2.3.6) and $h_{\rm peak}$ is determined by the ratio of the DM drag opacity to the Hubble rate at the time of DM decoupling $\dot{\kappa}(\eta_\chi)/\mathscr{H}$ (see Fig. 2.3.7).

- Using 50 simulations down to $z = 5$ within the $\{h_{\rm peak}, k_{\rm peak}\}$ parameter space (a new effective space in ETHOS), we have been able to classify DM models into regions with distinct non-linear structure formation at galactic scales (CDM, WDM, DAOs), quantified by the non-linear power spectrum and the halo mass function at high redshift.

- As far as the non-linear matter power spectrum is concerned, we find that only a small region within this effective parameter space, corresponding to relatively small values of $k_{\rm peak}$ and large values of $h_{\rm peak}$ (strong DAO models), still preserves a signature of the DAOs at $z = 5$ at the galactic scales corresponding to the smallest galaxy-forming haloes $500\,h\,\mathrm{Mpc}^{-1}$(Fig. 2.3.10). The rest of the relevant parameter space including weak and strong DAOs is either degenerate with WDM, which we quantify with $\hat{I}(k_{\rm probe})$ (Eq. 2.11), or indistinguishable from CDM. This distinct DAO region expands at higher redshifts and contracts at smaller scales (see Fig. 2.4.12).

- We find that it is possible to break (to a certain extent) the degeneracies between weak DAO models (small values of $h_{\rm peak}$) and WDM models seen in the non-linear power spectrum by characterising structure formation models using the halo mass function instead. This is because the halo mass function retains a memory of the linear power spectrum, having a slope that is very sensitive to the value of $h_{\rm peak}$. In this way, the distinct DAO region covers a much larger region of the parameter space and changes only slightly with redshift (see Figs. 2.4.14−2.4.15).

Using our results, it is possible to use the new analytic formula we propose (Eq. 2.3) to fit the linear power spectrum of a broad class of ETHOS models with DAOs, and use the values of $\{h_{\rm peak}, k_{\rm peak}\}$ to determine to which structure formation region they belong to in the non-linear high-redshift regime (CDM-like, WDM-like or DAO), without performing additional N-body simulations. Notice that this is valid for any DM particle model with a primordial power spectrum with DAOs that can be fitted accurately with our formula (up to the second DAO peak). In other words, given the values of $\{h_{\rm peak}, k_{\rm peak}\}$, our method allows to infer the value of the non-linear power spectrum and halo mass function at any relevant scale/mass at high-redshift $z \sim 10-5$.

Furthermore, The effective parameters for structure formation we propose here represent a potentially powerful way to constrain the parameter space of a variety of particle physics models by using observations in the high-redshift Universe, such

---

[7] In the case of WDM, $k_{\rm peak}$ is connected to $k_{1/2}$, the scale at which the transfer function squared is equal to $1/2$; see Eq. (2.4).

as the Lyman-α 1D flux power spectrum. We should remark however that in order to accurately exploit this avenue, we need to incorporate the baryonic physics that is relevant for the intergalactic medium into the ETHOS parameter space, which is something we plan to do in the future. However, the results for the $h_{\rm peak} = 0.2$ case presented in Archidiacono et al. (2019) and the expectation based on WDM constraints (Murgia et al., 2018) indicate that the lower left region in Fig. 2.3.10 is likely in tension with current Lyman-α data. Another avenue we will explore is to extend our results towards lower redshifts, where DM self-interactions have a significant impact in the centre of DM haloes, and thus need to be incorporated as an additional parameter to classify structure formation regimes.

Acknowledgements: SB and JZ acknowledge support by a Grant of Excellence from the Icelandic Research Fund (grant number 173929). CP acknowledges support by the European Research Council under ERC-CoG grant CRAGSMAN-646955. The simulations were performed on resources provided by the Icelandic High Performance Computing Centre at the University of Iceland, and the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

2.6 Appendix

2.6.1 Convergence tests

To determine the minimum scale at which we can trust our measurements of the power spectrum and the halo mass function in our simulations, we performed convergence tests for a few DM models using three resolution levels for each. These models cover representative regions of the parameter space we explore and, based on our analysis, they bracket the possible range of convergence variations. The convergence reported here is thus a fair representation of the convergence for all the parameter space explored in this paper. The three resolution levels were done within a $\sim(6.25\,{\rm Mpc}\,h^{-1})^3$ Lagrangian zoom region at $z = 5$ with 1024^3 (LR), 2048^3 (MR), and 4096^3 (HR) effective particle resolution. We set the goal to determine for the two lower resolution levels, the wavenumber $k_{\rm conv}$ at which the power spectrum differs by 5% with respect to the highest resolution. The upper panel of Fig. 2.6.16 shows this convergence test for the models highlighted with large circles within the parameter space shown in the inset. Notice that since

we subtract the shot-noise from the power spectra, the power falls off at the smallest scales. This behaviour is responsible for a smaller convergence scale (i.e. larger $k_{\rm conv}$) for the WDM models (green and red lines) compared to the sDAO models; this also applies in general to all wDAO models. In models with a steep *linear* power spectrum cutoff, the non-linear true power is expected to be highly suppressed at sufficiently small scales and thus, there is not much power left at the unresolved scales. However, the sDAO case is similar to CDM, there is still significant power left at the unresolved scales and therefore, the 5% convergence level is reached at larger scales (smaller $k_{\rm conv}$). Despite this difference across different DM models, Fig. 2.6.16 shows that there is at least a factor of 3 improvement in $k_{\rm conv}$ between the LR and MR simulations. The power spectra of the latter are converged to $k_{\rm conv} > 250\,h\,{\rm Mpc}^{-1}$and thus, assuming at least another factor of 2 improvement for the HR simulations, all models are converged to better than 5% at $k \sim 500\,h\,{\rm Mpc}^{-1}$.

The bottom panel of Fig. 2.6.16 is equivalent to the top panel but for the halo mass function. It shows the ratio of the halo mass functions of the two lower resolution levels to that of of the high resolution. It can be seen that the LR simulations drop below 5% convergence at $\sim 2 \times 10^9 {\rm M}_\odot/h$, while for the MR simulation this threshold occurs at $\sim 3 \times 10^8 {\rm M}_\odot/h$. We highlight however, that for the 2.5 keV model discreteness effects cause the well-known effect of spurious haloes (Wang & White, 2007), which appear in this case at $M < 10^8 {\rm M}_\odot/h$, dominating the signal. Although all models with a primordial power spectrum cutoff suffer from spurious haloes we find that in all cases, for our highest resolution, we are free from this effect at a halo mass of $10^8 {\rm M}_\odot/h$. We therefore set this mass as our lower mass limit for all cases and report a convergence of the halo mass function to better 5% for this and larger masses. While the resolution based convergence discussed above affects the small mass end of the halo mass function, the high mass end is affected by the limited volume of the zoom-in region, which can only encompass a few of the most massive haloes, leading to large Poisson (counting) errors. However, this is not relevant for our purposes as our models converge at large masses anyway and the differences we are interested in appear at smaller halo masses.

2.6.2 The DM linear power spectrum in the tight DM-DR coupling limit

The goal in Section 2.3.1 is to connect two different parametrizations for the linear power spectrum. On the one hand, the parameters we have defined in this work $\{h_{\rm peak}, k_{\rm peak}, h_2\}$, and on the other the parameters used in Cyr-Racine et al. (2016), essentially $\{a_n, n, \alpha_{l\geq 2}\}$. As we show in Section 2.4, the former set can be used directly to quantify the differences between different structure formation models in the non-linear regime, while the latter can be connected directly to the particle physics parameters of a given model. In this Appendix we explore the connection between these two sets of parameters, which allows to obtain a physical interpretation of the final parametrization we have used in this work.

We start by recalling that in Cyr-Racine et al. (2016) the DR and DM-drag opacities for a given ETHOS model are respectively written as:

$$\begin{aligned}\dot{\kappa}_{\rm DR}(z) &= -(\Omega_\chi h^2) a_n \left(\frac{1+z}{1+z_{\rm D}}\right)^n \\ \dot{\kappa}_\chi(z) &= -\frac{4}{3}(\Omega_{\rm DR} h^2) a_n \frac{(1+z)^{n+1}}{(1+z_{\rm D})^n},\end{aligned} \tag{2.17}$$

if we just consider a single pure power law, where Ω_χ and $\Omega_{\rm DR}$ are the DM and DR

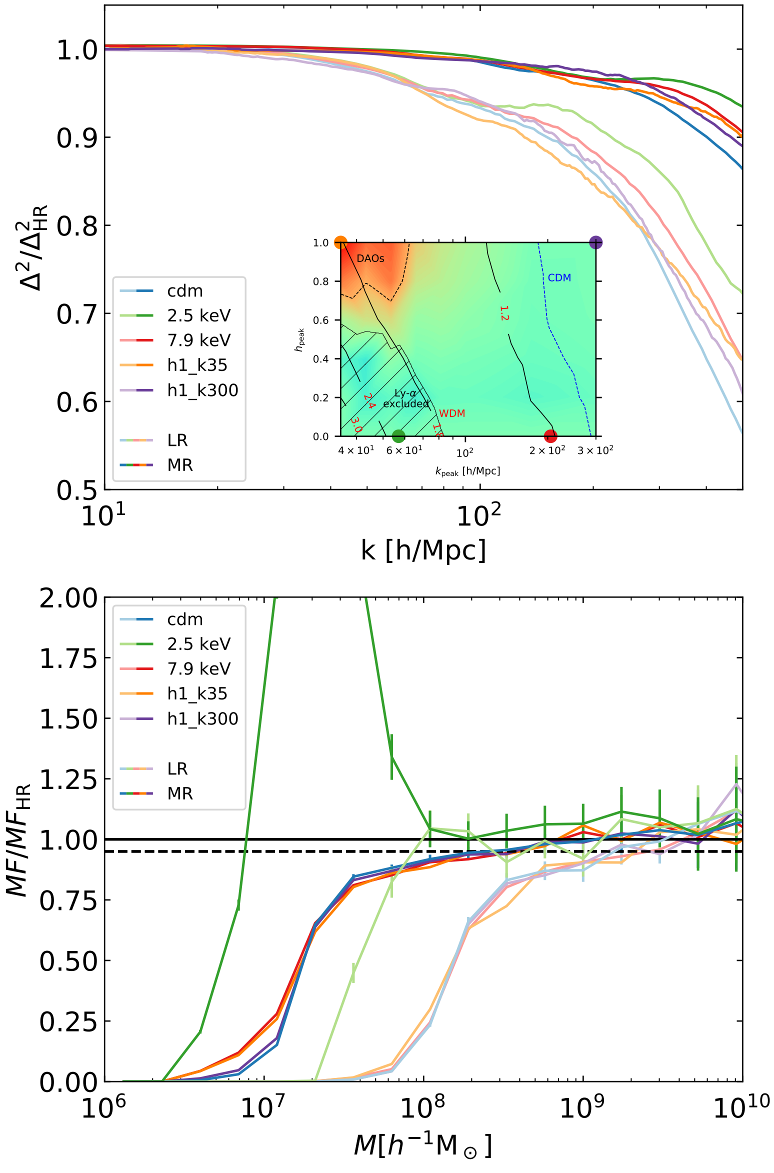

Figure 2.6.16. Top panel: Convergence of the power spectrum for 5 models: CDM (blue) and four models marked with solid circles in inset, which is a reproduction of the parameter space plot Fig. 2.3.10. The vertical axis is the ratio of the dimensionless power spectrum of the low-resolution (faded lines) and medium resolution (solid lines), relative to that of the highest resolution run at $z = 5$. The horizontal shaded line marks a convergence level of 5%. Bottom panel: The same as the top panel but for the halo mass function.

densities in units of the critical density, h is the dimensionless Hubble constant, a_n and n are given by the particle physics model, and $z_{\rm D}$ is an arbitrary normalization factor[8] chosen to be $z_{\rm D} = 10^7$. Eqs. (2.17) is an approximation that assumes that a single power law is sufficient for a given value of n, which is a good approximation for instance in the case of DM interacting with massless radiation via a massive mediator (models of this type were simulated in Vogelsberger et al. 2016). Assuming that DM and DR decouple well within the radiation dominated era (where the Hubble rate, relative to the conformal time goes as $\mathscr{H} = \eta^{-1}$; note that we use the conformal time in units of Mpc), Eqs. (2.17) can be rewritten in terms of the conformal time, using $1+z \propto \eta^{-1}$,

$$\begin{aligned} \dot{\kappa}_{\rm DR}(\eta) &= -(n-1)\frac{\eta_{\rm DR}^{n-1}}{\eta^n} \\ \dot{\kappa}_\chi(\eta) &= \frac{\dot{\kappa}_{\rm DR}}{\mathscr{R}} = -(n-1)\eta_1\frac{\eta_{\rm DR}^{n-1}}{\eta^{n+1}}, \end{aligned} \tag{2.18}$$

where $\eta_{\rm DR}$ is the conformal time of DR decoupling defined by $\int_{\eta_{\rm DR}}^{\eta_0} -\dot{\kappa}_{\rm DR} d\eta = 1$, and η_1 is the conformal time when $\mathscr{R} = \frac{3}{4}\rho_{\rm DM}/\rho_{\rm DR} = \eta/\eta_1 = 1$. Note that this definition of the decoupling time takes the weakly coupled regime into account, i.e., a broad visibility function will lead to a later decoupling time. The DR decoupling time $\eta_{\rm DR}$ can be related directly with a_n and n:

$$\eta_{\rm DR} = \left[\frac{\Omega_\chi h^2 a_n}{n-1}\left(\frac{1}{H_0\Omega_{\rm r}^{1/2}(1+z_{\rm D})}\right)^n\right]^{\frac{1}{n-1}}, \tag{2.19}$$

where $\Omega_{\rm r}$ is the radiation density in units of the critical density, and H_0 is the Hubble constant today. In a similar way, we can define the conformal time for DM decoupling by $\int_{\eta_\chi}^{\eta_0} -\dot{\kappa}_\chi d\eta = 1$, and thus η_χ can be written as:

$$\eta_\chi = \left(\frac{n-1}{n}\eta_1\eta_{\rm DR}^{n-1}\right)^{\frac{1}{n}}. \tag{2.20}$$

We can use the tight coupling approximation (e.g. Hu & Sugiyama 1996) for the DM-DR plasma to roughly capture the acoustic oscillations in the DM fluid as well as the effect of DR diffusion damping. We then propose that the DM density fluctuations have the following $k-$dependence:

$$\delta_\chi(k) \propto \cos(k r_{\rm DAO}) e^{-k^2/k_{\rm D}^2}, \tag{2.21}$$

where the sound horizon scale is given by:

$$r_{\rm DAO} = \int_0^{\eta_\chi} c_{\rm s} d\eta \approx \frac{\eta_\chi}{\sqrt{3}}, \tag{2.22}$$

where $c_{\rm s}$ is the dimensionless DM sound speed. The parameter $k_{\rm D}$ in Eq. (2.21), which controls the damping scale, is given by the tight coupling dispersion relation (Dodelson,

---

[8] This is the numerical value that was chosen in Cyr-Racine et al. (2016) to be the redshift when the DM opacity becomes equal to the conformal Hubble rate.

2003):

$$\frac{1}{k_{\rm D}^2} = \int_0^\eta \frac{d\eta'}{6(1+\mathscr{R})\dot{\kappa}_{\rm DR}} \left[\frac{4}{5\alpha_2} + \frac{\mathscr{R}^2}{1+\mathscr{R}}\right], \tag{2.23}$$

where α_2 is the ratio between the opacity of the quadrupole and dipole moment of the DR multipole hierarchy given by the angular dependence of DM-DR scattering.

For the limit when the visibility function is given as a delta function, the exponential term in Eq. (2.21) gives the damping envelope of the DAOs, but due to a finite width of the visibility function, the damping envelope should be weighted by the DR visibility function $g_{\rm DR} = -\dot{\kappa}_{\rm DR}\exp\left(-\kappa_{\rm DR}\right)$:

$$D_n(k) = \int_0^{\eta_0} d\eta\, g_{\rm DR}(\eta) e^{-k^2/k_{\rm D}^2}, \tag{2.24}$$

Therefore, under these approximations the linear transfer function is given by:

$$T_{\rm L}^2(k) \approx \cos^2\left(\frac{k\eta_\chi}{\sqrt{3}}\right) D_n^2(k), \tag{2.25}$$

where the input parameters $(\eta_\chi, \eta_{\rm DR}, k_{\rm D})$ are ultimately given by the cosmological parameters assumed (Ω_χ, $\Omega_{\rm r}$, H_0) and the set of values $\{\Omega_{\rm DR}, n, a_n, \alpha_2\}$ for a given ETHOS model.

Eq. (2.25) results in damped DAOs that resemble the behaviour of the linear power spectrum of ETHOS models generated by the Boltzmann solver (see Fig. 2.6.17). This approach also predicts correctly that the amplitude of the DAOs scales with the value n. Unfortunately, the agreement is only qualitative, neither the position of the peaks nor the damping envelope agree with the full calculation. An example of this can be seen by comparing the red solid and dashed lines in Fig. 2.6.17, the latter of which is the ETHOS-2 model in Vogelsberger et al. (2016) (with $n = 4$, $a_4 = 1784.05\,h\,{\rm Mpc}^{-1}$, $\alpha_{l\geq 2} = 3/2$; see their table 1).

The previous approach ignores the effect of the velocity perturbations on the density perturbation δ_χ, the so-called velocity overshoot. We improved the modelling by including this effect which is part of a more rigorous treatment (see Hu & Sugiyama 1996). Modelling the velocity perturbation as $\theta_\chi \sim \sin(kr_{\rm DAO})D_n(k)$ moves the first DAO peak to better agree with the numerical results, but the other peaks and damping envelope were still not in agreement. Exchanging the sine and cosine functions for the full analytic solutions, in which case the potentials are given by Bessel functions, gives the evolution of δ_χ and θ_χ until DR decoupling with good accuracy. The position of the peaks of the DAOs are captured quite accurately with this modification, but the approach still does not capture the transition from the tightly coupled to the weakly coupled regimes correctly. The main reason for this seems to be that the exponential diffusion damping in Eq. (2.24) is not an accurate representation of the numerical results for the first few DAO peaks at large scales; it is only a good approximation at much smaller scales. The damping of the first DAO peaks deviates strongly from the exponential behaviour because the timescale for DM decoupling is large compared to the oscillation frequency causing the DM to spend a longer time in the weakly coupled regime. Trying to model the DM decoupling timescale by weighting δ_χ and θ_χ with the visibility function, improves slightly the result of this analytical approach, but it remains inaccurate.

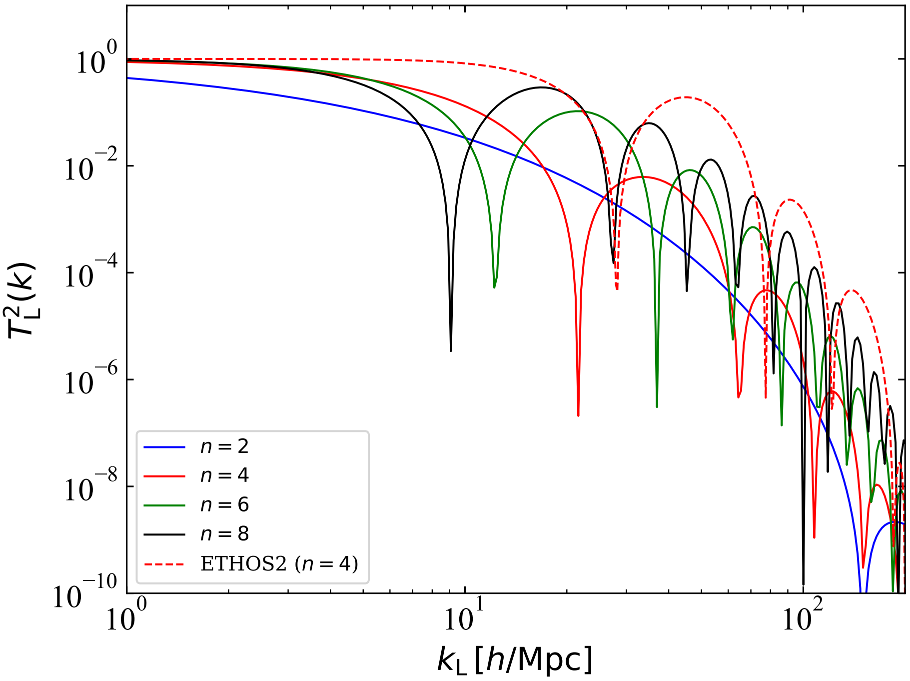

Figure 2.6.17. The solid lines are the linear transfer function for different values of n computed with Eq. (2.25) *using the a fixed value of* η_{DR} *chosen to match that of the ETHOS-2 model in Vogelsberger et al. (2016) (dashed lines) and ETHOS-2. Note that by fixing* η_{DR}*, the value of* a_n *for a given n is given by Eq.* (2.19).

Instead of increasing the complexity of the modelling, which would eventually take us closer and closer to a full approach of the Boltzmann solver, but would defeat the purpose of having a simple physical interpretation, we choose instead to use a phenomenological approach as described in Section 2.3.1. The starting point is to notice that the ratio $\dot{\kappa}_\chi/\mathscr{H} \propto (1+z)^n$ is the relevant ratio of timescales (or length scales) in the DM decoupling process, with $\dot{\kappa}_\chi/\mathscr{H} \gg 1$ corresponding to the tightly coupled regime and $\dot{\kappa}_\chi/\mathscr{H} \ll 1$ to the decoupled regime. We found that the value of this ratio at the DM decoupling time η_χ is actually strongly correlated with the amplitude of the first DAO peak h_{peak} (Fig. 2.3.7) since it is actually equal to n as can be seen through Eqs. (2.18-2.20), which controls how fast the transition is from the tightly coupled to decoupled regimes, and thus how narrow the DM drag visibility function is.

Chapter 3

The properties of ETHOS haloes

This chapter is based on the following article:

The halo mass function and inner structure of ETHOS haloes at high redshift

Published in Monthly Notices of the Royal Astronomical Society, stab1758, Oxford University Press

Authors:
Sebastian Bohr[1], Jesús Zavala[1], Francis-Yan Cyr-Racine[2] and Mark Vogelsberger[3]
[1]Centre for Astrophysics and Cosmology, Science Institute, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland
[2]Department of Physics and Astronomy, University of New Mexico, 210 Yale Blvd NE, Albuquerque, NM 87106, USA
[3]Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

We study the halo mass function and inner halo structure at high redshifts ($z \geq 5$) for a suite of simulations within the structure formation ETHOS framework. Scenarios such as cold dark matter (CDM), thermal warm dark matter (WDM), and dark acoustic oscillations (DAO) of various strengths are contained in ETHOS with just two parameters h_{peak} and k_{peak}, the amplitude and scale of the first DAO peak. The Extended Press-Schechter (EPS) formalism with a smooth-k filter is able to predict the cut-off in the halo mass function created by the suppression of small scale power in ETHOS models (controlled by k_{peak}), as well as the slope at small masses that is dependent on h_{peak}. Interestingly, we find that DAOs introduce a localized feature in the mass distribution of haloes, resulting in a mass function that is distinct in shape compared to either CDM or WDM. We find that the halo density profiles of *all* ETHOS models are well described by the NFW profile, with a concentration that is lower than in the CDM case in a way that is regulated by $k_{\rm peak}$. We show that the concentration-mass relation for DAO models can be well approximated by the mass assembly model based on the extended Press-Schechter theory, which has been proposed for CDM and WDM elsewhere. Our results can be used to perform inexpensive calculations of the halo mass function and concentration-mass relation within the ETHOS parametrization without the need of $N-$body simulations.

3.1 Introduction

A majority of the matter content of the Universe is made up by dark matter (DM), which is therefore a crucial ingredient in cosmological structure formation. A likely explanation for DM is that is made of yet undiscovered particle(s), whose nature remains a mystery. A prominent assumption within the particle hypothesis is that taken by the Cold Dark Matter (CDM) model, which in essence states that the only DM interaction relevant for structure formation is gravity. CDM has been established as the standard paradigm for structure formation since it has been shown to be consistent with the observed structure of the Universe on large scales (e.g. Springel et al., 2005). However, the CDM model remains challenged on smaller (galactic) scales in various ways: (i) the underabundance of low-mass galaxies (either satellites or in the field) (Klypin et al., 1999; Moore et al., 1999; Zavala et al., 2009; Papastergis et al., 2011; Klypin et al., 2015), (ii) the core-cusp problem in low-surface brightness galaxies and possibly in dwarf spheroidals (de Blok & McGaugh, 1997; Walker & Peñarrubia, 2011), (iii) the "too-big-to-fail problem" (Boylan-Kolchin et al., 2011; Papastergis et al., 2015), (iv) the plane of satellites problem (Pawlowski et al., 2013), and (v) the diversity problem of rotation curves in dwarf galaxies (Oman et al., 2015). For recent reviews on the CDM challenges and plausible solutions see Bullock & Boylan-Kolchin (2017) and Zavala & Frenk (2019).

A possible approach to address these potential issues is to invoke additional DM physics, i.e., to consider departures from the CDM hypothesis that change its predictions on small scales while leaving the large scale behaviour intact. A novel framework (ETHOS) has been proposed to incorporate new DM physics into structure formation theory, connecting a broad range of DM particle physics to effective parameters that characterize structure formation in the linear regime (Cyr-Racine et al., 2016; Vogelsberger et al., 2016), and further to effective parameters that capture the behaviour of different DM models in the non-linear regime (Bohr et al., 2020). The new parametrization introduced in Bohr et al. (2020) is based on describing dark acoustic oscillations (DAOs) in the linear power spectrum. The two physically motivated parameters $h_{\rm peak}$

and $k_{\rm peak}$, the amplitude and scale of the first DAO peak, respectively, suffice to describe the linear power spectrum for DM models from WDM ($h_{\rm peak} = 0$) over weak DAOs (wDAO; $h_{\rm peak} \sim 0.2$, like those in Vogelsberger et al. 2016) to strong DAOs (sDAO; $h_{\rm peak} \sim 1$, like those in Bose et al. 2019). In Bohr et al. (2020), it was shown the parameter space of DM ETHOS models in the $(k_{\rm peak}, h_{\rm peak})$ can be divided clearly in distinct structure formation regions (CDM-like, WDM-like and DAO-like). When this division is done according to the non-linear power spectrum at high redshift, only a small region of the parameter space still displays distinct DAO features by $z = 5$. This DAO region can be augmented if the halo mass function is used instead as a measure to classify the models; Bohr et al. (2020) found that the halo mass function is especially sensitive to the presence of DAO features in the linear power spectrum.

In this work, we apply the Extended Press-Schechter (EPS) formalism (Press & Schechter, 1974; Bond et al., 1991; Sheth & Tormen, 1999) to a wide range of ETHOS models, which has not been done before broadly (Sameie et al. 2019 applied this formalism to the small subset of wDAO ETHOS models in Vogelsberger et al. (2016)), and tweak it to accurately represent the simulated halo mass function. The use of this formalism offers a quick way to compute the halo mass function without the need to run dedicated and computationally expensive N-body simulations.

The non-linear power spectrum at small scales depends both on the halo mass function and the inner structure of DM haloes, both of which are affected by the DM nature. In particular cut-offs and additional features in the linear power spectrum due to new DM physics have been shown to affect not only the abundance of DM haloes, but also their inner density profile. For instance, for WDM it has been shown that DM haloes still follow a NFW density profile, but with lower concentration for small haloes relative to CDM (see e.g. Lovell et al., 2014; Ludlow et al., 2016). On the other hand, for DAO models, it has been shown that haloes become overall less centrally dense due to the suppression of power at small scales (see e.g. Buckley et al., 2014; Vogelsberger et al., 2016). However, the inner halo properties of DAO models have not been studied in detail, in terms of their dependence on the scale and amplitude of the DAOs. This is something we pursue in this work by looking at the halo concentration in ETHOS models and attempting to predict its behaviour using the model of Ludlow et al. (2016) coupled with the EPS formalism.

Our work focuses on the high redshift regime ($z \geq 5$) to test the limits of the EPS formalism and the concentration model of Ludlow et al. (2016) for ETHOS models. The high redshift regime has been shown to be a promising one to probe and distinguish different ETHOS models (e.g. see Muñoz et al., 2021, for predictions for the 21-cm signal) and it is therefore important to test the validity of analytical approaches such as EPS. Our work is also motivated by a lack of previous work studying the inner structure of haloes at high redshift for DAO models.

This paper is organized as follows. In Section 3.2, we shortly summarize the setup of the simulations used in this work. Section 3.3 covers the EPS formalism for the halo mass function and its application to our set of ETHOS simulations. In Section 3.4, the inner halo structure is studied by looking at the concentration parameter of DM haloes. Finally, our conclusions are given in Section 3.5.

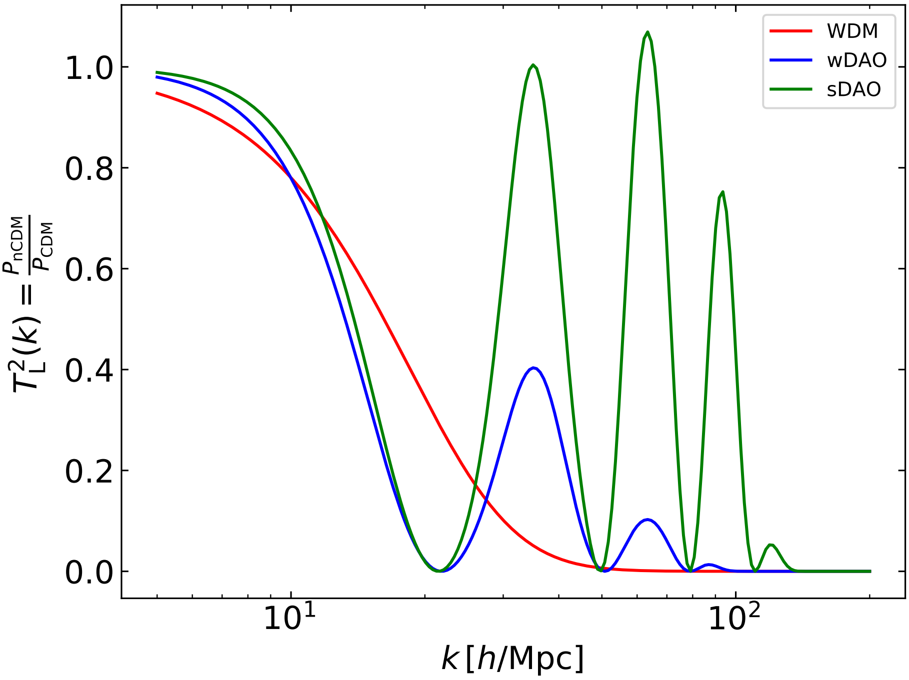

Figure 3.2.1. Initial linear transfer functions $T_{\rm L}^2(k)$ for examples of WDM (red), wDAO (blue) and sDAO (green) models. All three models have identical $k_{\rm peak} = 35\,h\,{\rm Mpc}^{-1}$.

3.2 Simulations

In this work, we use the cosmological DM-only N-body simulations that were described in detail in Bohr et al. (2020); they were performed with the code AREPO Springel 2010 from initial conditions generated with MUSIC Hahn & Abel 2011. All simulations use the cosmological parameters $\Omega_{\rm m} = 0.31069$, $\Omega_\Lambda = 0.68931$, $H_0 = 67.5\,{\rm km/s/Mpc}$, $n_{\rm s} = 0.9653$ and $\sigma_8 = 0.815$, where $\Omega_{\rm m}$ and Ω_Λ are the matter and cosmological constant contributions to the matter-energy density of the Universe today, respectively, H_0 is today's Hubble constant, $n_{\rm s}$ is the spectral index, and σ_8 is the mass variance on $8\,{\rm Mpc}\,h^{-1}$scales. The high resolution region of the simulations has a comoving smoothing length of $\varepsilon = 0.2\,{\rm ckpc/h}$ and a particle mass of $8 \times 10^4\,{\rm M}_\odot/h$.

The suite of simulations covers CDM ($k_{\rm peak} \to \infty$) and WDM-like models ($h_{\rm peak} = 0$) in a wide range of cut-off scales ($k_{\rm peak} = 35 - 300\,h\,{\rm Mpc}^{-1}$; equivalent to WDM masses $m_\chi \approx 1.6 - 11\,{\rm keV}$). The suite covers a range of DAO models from weak DAOs with $h_{\rm peak} = 0.2 - 0.6$ to strong DAOs with $h_{\rm peak} = 0.8 - 1$ (for the effect of sDAO features on the Lyman-α forest, see Bose et al., 2019) and DAO scales of $k_{\rm peak} = 35 - 300\,h\,{\rm Mpc}^{-1}$. We note that some of the WDM models explored here are already in tension with current constraints on the non-linear power spectrum from Lyman-α observations, e.g. the allowed WDM masses $m_{\rm WDM} > 3.6\,{\rm keV}$ (Murgia et al.,

2018) would correspond to $k_{\text{peak}} \gtrsim 85\,h\,\text{Mpc}^{-1}$in the limit $h_{\text{peak}} \to 0$, which is the WDM limit in our parametrization. Since wDAO models show degeneracies with WDM in the matter power spectrum, the wDAO models with $k_{\text{peak}} \lesssim 65\,h\,\text{Mpc}^{-1}$are also ruled out by the same observations in a way that is predicted in Bohr et al. (2020) (see Fig. 10 therein). For sDAO models, there is a single simulation including baryonic physics that explores the impact of sDAO features in the Lyman-α forest 1D flux spectrum (Bose et al., 2019), and given their quite distinct behaviour relative to WDM, a detailed analysis is needed to properly set constraints in sDAO models. Overall, only a few of the models in the simulation suite we use can be considered as being ruled out by current observations, but we nevertheless include them here for illustrative purposes of the extreme behaviour in the wDAO and WDM regimes at low k_{peak} and low h_{peak}.

Figure 3.2.1 shows the linear transfer function of examples of WDM, wDAO and sDAO models with identical k_{peak}. Physically, the DAO models are characterised by the sound horizon scale at the time of DM-DR decoupling, which essentially sets the scale of the first DAO peak, k_{peak}, while the amplitude of this peak is determined by the timescale of the DM-DR decoupling relative to the Hubble rate (see Section 3.1 in Bohr et al. 2020), which is what sets the difference between the wDAO and sDAO regimes. A faster decoupling timescale leads to a fast transition from the tightly coupled regime to the decoupled regime and the DM power spectrum does not get damped significantly (sDAOs). For larger decoupling timescales, there is a slow transition between these regimes, with the extended period of the weakly coupled regime dampening the DAOs significantly (wDAOs).

Finally, we remark that ETHOS models self-consistently contain astrophysically relevant self-interacting cross sections, which can impact the inner structure of DM haloes (see Cyr-Racine et al., 2016; Vogelsberger et al., 2016). However, as in Bohr et al. (2020), in this paper we only consider the effect of the primordial suppression of the matter power spectrum and leave a study of the effect of possible DM self-interactions for future work. We do this for two reasons. First, we want to cleanly separate the effects of the primordial suppression and DM self-interactions. Second, we expect self-interactions to be more relevant at lower redshifts than studied here ($z \gtrsim 5$; see e.g Vogelsberger et al. 2014c). From the simulations, the haloes were constructed using FOF and SUBFIND algorithms included in AREPO with a particle number limit of 32. For more details on the simulations, see Bohr et al. (2020).

3.3 Halo mass function in ETHOS haloes at high redshift

For the effect of different ETHOS models on haloes, we first look at their abundance as measured by the halo mass function. For the halo mass function, we do not include subhaloes, but purely main haloes.

3.3.1 Extended Press-Schechter formalism

The halo mass function can be modelled from the linear power spectrum using variants of the Press-Schechter formalism (Press & Schechter, 1974; Bond et al., 1991; Sheth & Tormen, 1999; Sheth et al., 2001) The following is a brief summary of the key equations in the variant we will use.

Regions with a characteristic size R corresponding to a mean mass scale:

$$M = \frac{4\pi}{3}\bar{\rho}_m R^3, \tag{3.1}$$

where $\bar{\rho}_m = \Omega_m \rho_c$ is the mean matter density (ρ_c is the critical density of the Universe), have a smoothed density field δ_M:

$$\delta_M \equiv \delta(\vec{x};R) = \int \delta(\vec{x}')W_R(\vec{x}-\vec{x}';R)d^3\vec{x} \tag{3.2}$$

where W_R is a window or filter function properly normalised, and $\delta(\vec{x})$ is the matter density contrast. The (linear) mass variance is the most relevant statistical quantity of the smoothed density field in the Press-Schechter formalism, and it is given by:

$$\sigma^2(R) = \frac{1}{2\pi^2}\int_0^\infty dk k^2 P(k)\widetilde{W}_R^2(k) \tag{3.3}$$

where $P(k)$ is the linear power spectrum and $\widetilde{W}_R(k)$ is the Fourier transform of the window function in Eq. (3.2).

In the Extended Press-Schechter (EPS) formalism, it is then argued that the comoving number density $n(M)$ of collapsed haloes of mass M (Eq. 3.1) is given by:

$$\frac{dn}{d\ln M} = -\frac{1}{2}\bar{\rho}_m \frac{f}{\sigma^2}\frac{d\sigma^2}{dM} \tag{3.4}$$

where $f(\nu)$ is the so-called first crossing distribution (or multiplicity function) within the ellipsoidal collapse model (see Sheth et al. 2001):

$$f(\nu) = A\sqrt{\frac{2q\nu}{\pi}}(1+(q\nu)^{-p})\exp\left(-\frac{q\nu}{2}\right) \tag{3.5}$$

where $p = 0.3$, $q = 1$, and we fit A with our simulations, while ν is defined in terms of the (linear) density threshold for collapse in the spherical collapse model:

$$\nu = \frac{\delta_c^2}{D^2(z)\sigma^2} \tag{3.6}$$

where $\delta_c = 1.686$ and $D(z)$ is the growth factor in cosmological linear perturbation theory:

$$D(z) = \frac{H(z)\int_0^{1/(1+z)}\frac{da}{a^3H^3(a)}}{H_0\int_0^1 \frac{da}{a^3H^3(a)}} \tag{3.7}$$

where H is the Hubble parameter.

We note that we need to introduce a correction to the formalism described above since our simulation suite uses a zoom-in technique with a high-resolution volume that is in fact over-dense relative to the mean cosmic volume. Notice that this bias in the mean overdensity over the simulated volume is present even after using the technique described in Bohr et al. (2020) in which the high-resolution region within the larger parent cosmological box is chosen to match as closely as possible the power spectrum of the (lower resolution) parent box in the overlaping scales (see Fig. 2 of Bohr et al. 2020).

Due to this bias, the mass function given by Eq. (3.4) is not directly comparable to the halo mass function extracted from our simulations. It needs to be adjusted for finite volume effects in two ways (see also Sheth & Tormen, 2002): (i) the mass variance has to be corrected for the mass variance of the high-resolution subregion of mass $M_{\rm sub}$

$$\sigma^2(M) \to \sigma^2(M) - \sigma^2(M_{\rm sub}), \tag{3.8}$$

and (ii) the threshold for collapse needs to be shifted by the overdensity of the subregion $\delta_{\rm sub}$:

$$\delta_c \to \delta_c - \delta_{\rm sub}. \tag{3.9}$$

For the window function $\widetilde{W}_R(k)$ in Eq. (3.3), the top-hat filter is the common and successful choice when studying CDM, while a sharp-k filter gives better results for WDM (Schneider et al., 2013), but neither seems to accurately account for DAO features in the linear power spectrum (Schewtschenko et al., 2015). Leo et al. (2018a) proposed a smooth-k space filter, which does not abruptly cut off like the sharp-k filter, but transitions more smoothly according to:

$$\widetilde{W}_R^{\rm smooth}(k) = \frac{1}{1+\left(\frac{kR}{c_{\rm W}}\right)^{\beta}}, \tag{3.10}$$

where the two free parameters β and $c_{\rm W}$ control how sharp the cut off transition is and re-scale the size of the collapsing region ($\widetilde{R} = R/c_W$), respectively. Sameie et al. (2019) used this filter to study the halo mass function of weak DAO models from previous ETHOS simulations ($h_{\rm peak} = 0.2$; based on Vogelsberger et al. 2016) and found a relatively good agreement. In this work, we use this smooth filter to study the suite of ETHOS simulations from Bohr et al. (2020) within the ($h_{\rm peak}$,$k_{\rm peak}$) parameter space.

3.3.2 EPS formalism applied to ETHOS models

With all the previous considerations, we fit the free parameters in the EPS mass function simultaneously to all our ETHOS simulations in the range $h_{\rm peak} = 0-1$ and $k_{\rm peak} = 35-300\,h\,{\rm Mpc}^{-1}$(plus CDM), and across the redshift range $z = 5-12$ by minimizing the χ^2. We fit the halo mass function for $M > 10^7\,{\rm M}_\odot/h$ or $M > M_{\rm lim}$ for models with $M_{\rm lim} > 10^7\,{\rm M}_\odot/h$, where $M_{\rm lim}$ is the limiting mass for spurious haloes as defined in Wang & White (2007). For the mass of our simulated haloes, we use $M_{200} = 4\pi/3r_{200}^3 200\rho_{\rm c}$, where the virial radius r_{200} is defined as the radius at which the enclosed

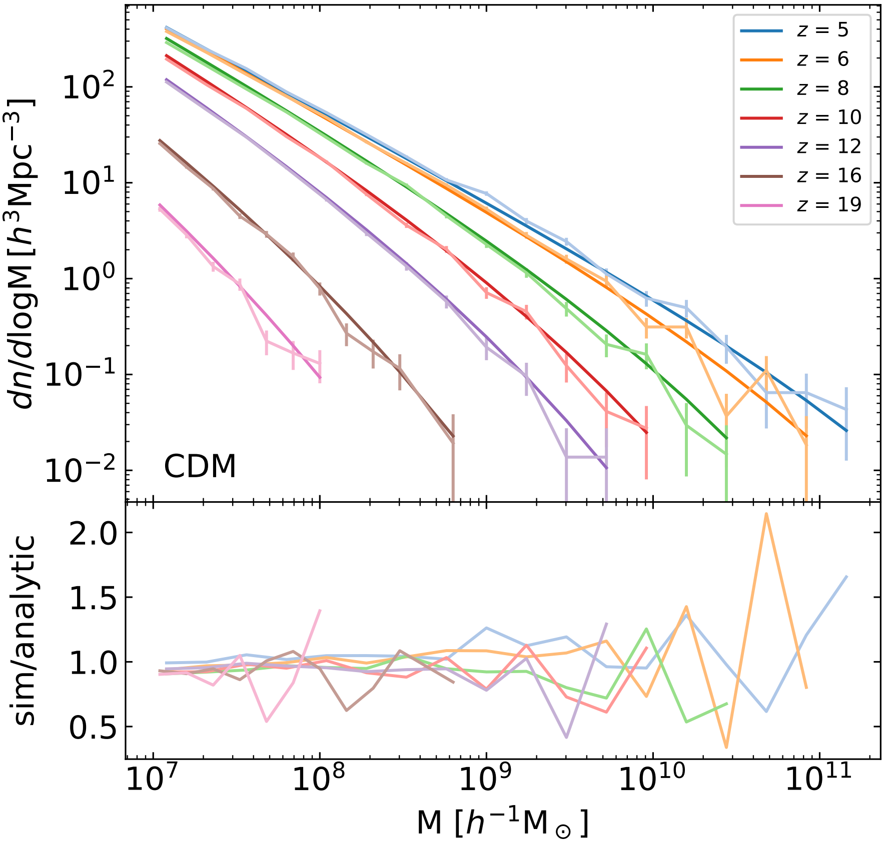

Figure 3.3.2. Halo mass function for the CDM model at different redshifts ($z \geq 5$) according to the colours in the legend. The light coloured lines with error bars are measurements from the CDM simulation in Bohr et al. (2020); the error bars are Poissonian. The dark coloured lines without error bars are computed from the EPS halo mass function (Eq. 3.4) corrected for finite volume effects (Eqs. 3.8–3.9) and with smooth-k space window function (Eq. 3.10). The best-fit parameters of the EPS halo mass function are $A = 0.3658$, $\beta = 3.46$, $c_W = 3.79$. The bottom panel shows the ratio between the simulation and the EPS results at each redshift.

density is 200 times the critical density of the Universe ρ_c. We find the best-fitting parameters to be: $A = 0.3658$, $\beta = 3.46$, $c_W = 3.79$. The agreement between the best-fit parameters of the EPS mass function and the simulation results can be seen in Figs. 3.3.2–3.3.5, where the faded lines with error bars are the result from the simulations and the solid lines are the analytic predictions using the same best-fit parameters as in the CDM case (given in the caption of Fig. 3.3.2).

Fig. 3.3.2 shows the CDM halo mass function and it is clear that our EPS implementation results is in an overall good fit to the simulation data across a wide range of redshifts ($5 \leq z \leq 19$). The scatter at the largest halo masses at a given redshift in the

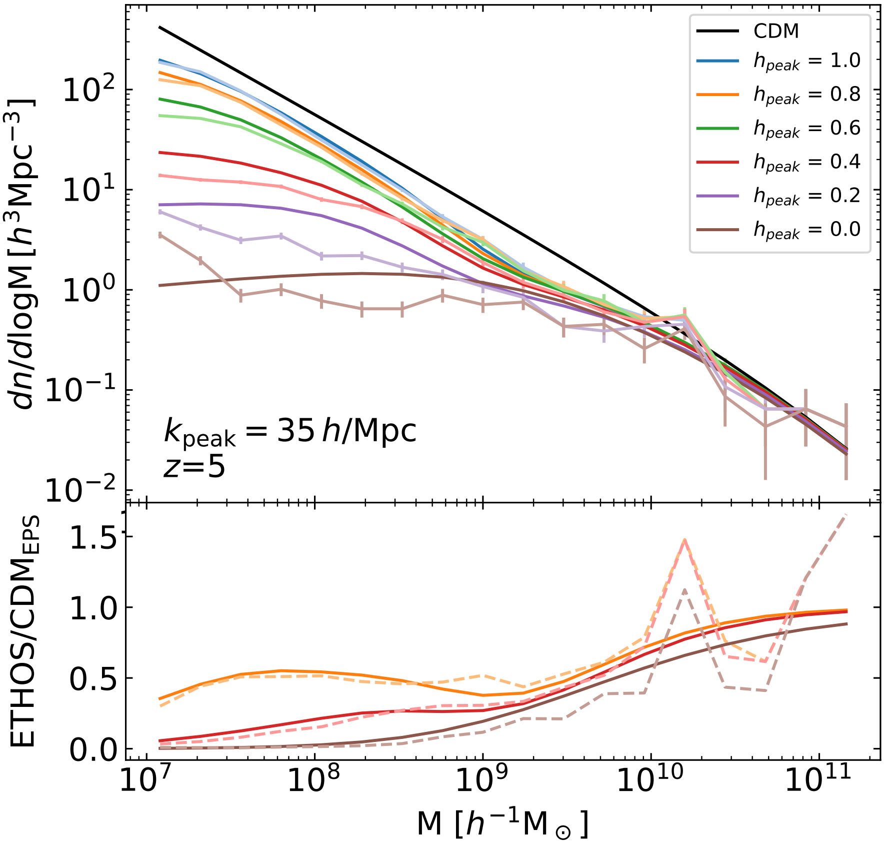

Figure 3.3.3. Halo mass function at $z = 5$ for the ETHOS simulations of Bohr et al. (2020) having the same $k_{\rm peak} = 35\,h\,{\rm Mpc}^{-1}$(i.e. the same linear power spectrum cutoff scale), but with different values of $h_{\rm peak}$ according to the different colours in the legend, from WDM ($h_{\rm peak} = 0$), to models with strong DAOs ($h_{\rm peak} = 1$). The light coloured lines with error bars are the simulation results, while the dark coloured lines are produced with the EPS model as described in the text and the caption of Fig. 3.3.2. For the lowest $h_{\rm peak}$, the onset of spurious haloes (see Wang & White, 2007) is visible just above $10^7\,{\rm M}_\odot/h$. The bottom panel shows the ratio of ETHOS with respect to the CDM EPS result for a selection of models as given by the colours, with the case of the simulations (EPS models) shown with dashed (solid) lines.

simulation results is expected and comes from low-number statistics, given the relatively small volume of our zoom-in simulations. In the mass range where the sampling error is small, the typical mismatch between the EPS modelling and the simulations is $\lesssim 10\%$.

Fig. 3.3.3 shows the models with the smallest value of $k_{\rm peak} = 35\,h\,{\rm Mpc}^{-1}$in our simulations for the full range of $h_{\rm peak} = 0-1$. These models correspond to linear power spectra with the largest cutoff-scale. By looking at the upper panel, it is clear that for sDAO models ($h_{\rm peak} \gtrsim 0.6$) the analytic prediction can accurately reconstruct the

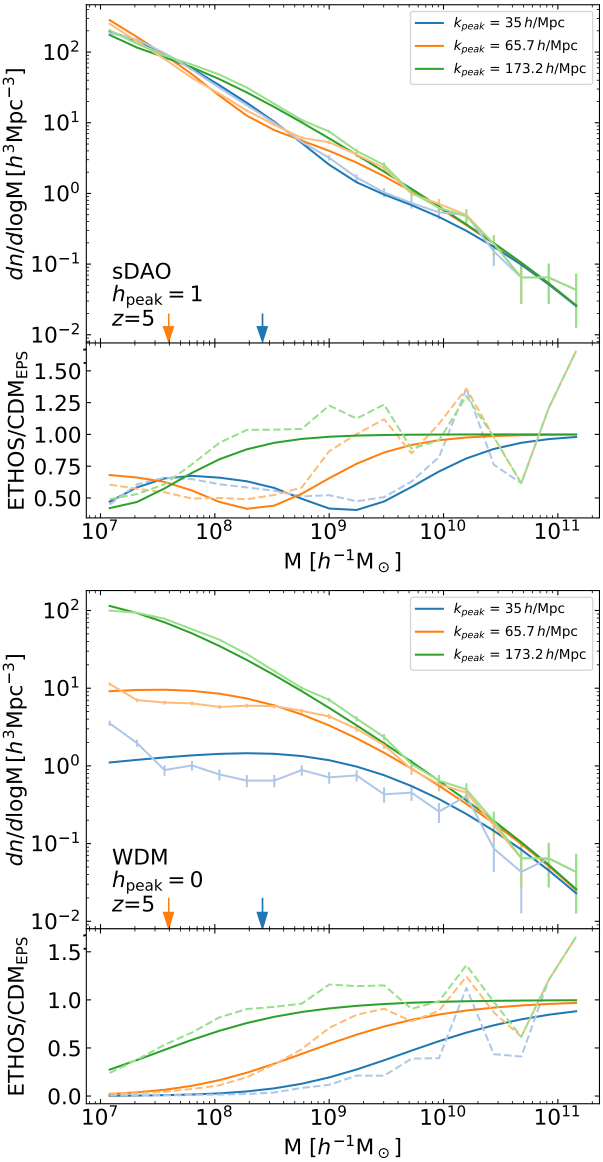

Figure 3.3.4. Halo mass function at $z = 5$ for the ETHOS simulations of Bohr et al. (2020) having the same DAO amplitude ($h_{\rm peak} = 1$, sDAO on the top, and $h_{\rm peak} = 0$, WDM on the bottom), but with different values of $k_{\rm peak}$ (the corresponding mass is indicated with arrows) according to the different colours in the legend. The light coloured lines with error bars are the simulation results, while the dark coloured lines are produced with the EPS model as described in the text and the caption of Fig. 3.3.2. For the WDM models, the onset of spurious haloes (see Wang & White, 2007) is visible just above $10^7\ {\rm M}_\odot/h$. The respective bottom panels show the ratio of ETHOS with respect to the CDM EPS result, with the case of the simulations (EPS models) shown with dashed (solid) lines.

halo mass function across all masses. The small-scale suppression in the linear power spectrum relative to CDM results in a deficit in the abundance of low-mass haloes, which is captured quite well by the EPS formalism, both in the cutoff mass-scale, and even in the subsequent oscillations observed at smaller masses. For the wDAO ($h_{\rm peak} \lesssim 0.6$) and WDM ($h_{\rm peak} = 0$) models on the other hand, only the general cut-off is captured by the analytic prediction, while the amplitude and details at small masses are slightly over-predicted and not captured as well. The bottom panel of Fig. 3.3.3 shows the ratio between the halo mass function of the ETHOS model (simulation in faded lines, and EPS predictions in dark coloured lines) to that of the CDM EPS prediction. If we just compare a given ETHOS model to CDM, the suppression of low mass haloes for WDM and wDAO models far outweighs the differences between the simulation result and the EPS formalism. We note that overall, our results in regards to the mismatch between the EPS formalism and the case $h_{\rm peak} = 0.2$ (belonging to the class of ETHOS models studied in Vogelsberger et al. 2016; see Bohr et al. 2020) is in general agreement with the high redshift results of Sameie et al. (2019) who directly studied the ETHOS simulations of Vogelsberger et al. (2016).

We also notice that the models with $h_{\rm peak} \leq 0.2$ suffer from the presence of spurious haloes due to discreetness effects; a well known artifact in models where the linear power spectrum is well below the unavoidable Poisson noise present in the creation of the initial conditions (see Wang & White, 2007). For these models, the halo mass function starts rising artificially towards the smallest masses just below a few times $10^7\ {\rm M}_\odot/h$. We notice that the mass scale where spurious haloes becomes apparent in the halo mass function of our simulations is roughly in agreement with the limiting mass formula for discreteness effects given by Wang & White (2007)[1]. For instance, for our most extreme WDM model, the limiting mass according to Wang & White (2007) is $1.6 \times 10^8\ {\rm M}_\odot/h$, whereas we see a clear artificial increase in the halo mass function at about half this value. For most of the models we analyse, the limiting mass is significantly lower than that of the extreme WDM model, and since the range of masses we are interested on is above this maximum limiting value, we will not discuss the presence of spurious haloes any further. Notice that in Section 3.4 below we only analyse the inner structure of haloes having a mass at least an order of magnitude larger than the mass where spurious haloes starts to become apparent in the halo mass function.

The behaviour of the halo mass function for a fixed $h_{\rm peak}$ but with different $k_{\rm peak}$ values (i.e. effectively different cutoff scales in the linear power spectrum) is shown in Fig. 3.3.4. The top panel exemplifies the sDAO models ($h_{\rm peak} = 1$) while WDM models ($h_{\rm peak} = 0$) are shown in the bottom panel. The EPS formalism remarkably captures the shift of the cut-off mass for different values of $k_{\rm peak}$; the signature of the DAOs in the halo mass function is also well reproduced by the model. In light of this agreement with the EPS formalism, we can say that this results confirms the expectation that the halo mass at which the cut-off occurs is directly connected to the mass within a radius proportional to the DAO scale $k_{\rm peak}$. On the other hand, for the WDM models (bottom panel of Fig. 3.3.4) it is especially noticeable that the agreement between the EPS model and the simulation becomes progressively better with increasing $k_{\rm peak}$. That behaviour is expected as the models approach CDM with increasing $k_{\rm peak}$.

---

[1] $M_{\rm lim} = 10.1 \times \bar{\rho} d k_{\rm p}^{-2}$, where $\bar{\rho}$ is the mean density of the Universe, d is the mean interparticle separation, and $k_{\rm p}$ is the wavenumber at which the initial dimensionless power spectrum reaches its maximum.

Finally, Fig. 3.3.5 shows that also the redshift evolution of sDAO models is well captured by the EPS formalism in a way that is essentially as good as for CDM (see Fig. 3.3.2). In the ratio relative to CDM (bottom panel of Fig. 3.3.5), shown only for three redshifts, it is also visible that the deficit of low-mass haloes (relative to CDM) is higher at larger redshifts and progressively decreases towards lower redshifts.

Overall, we conclude that the halo mass function predicted by the EPS formalism, corrected by finite volume effects and with the smooth-k space filter works very well for sDAO models ($h_{\rm peak} = 0.6 - 1$) across all probed masses. The formalism however, over-predicts the small mass abundance for WDM and wDAO models although the difference with respect to CDM is still reasonably captured. Finally, we note that we were also able to reconstruct the halo mass function of the wDAO ETHOS models presented in Vogelsberger et al. (2016) in the redshift range studied here and found a reasonable agreement with our EPS modelling, in line with what was described above for wDAO models.

3.3.3 Shape of the Halo Mass Function for ETHOS models

Figures 3.3.3−3.3.5 make clear that the shape of the halo mass function for models displaying DAOs in their linear matter power spectrum differs significantly from either the WDM or CDM case. While WDM ($h_{\rm peak} = 0$) mass functions are characterized by a uniform and monotonic suppression below a given mass scale (usually parameterized by their half-mode mass), DAO models display non-monotonic mass functions for which the initial (higher mass) suppression is followed by a localized feature where the mass functions converge back towards the CDM amplitude before decaying again on even smaller mass scales. This localized feature is clearly visible in the lower panel of Fig. 3.3.3 where we see that it becomes more prominent as $h_{\rm peak}$ increases. The presence of this feature is a direct consequence of the early-universe acoustic waves propagating in the dark sector for these ETHOS models, which later become imprinted in the dark matter density field once the latter decouples from the radiation bath. These frozen density waves then provide a slight enhancement of the dark matter fluctuations field once smoothed over a scale corresponding roughly to the DAO scale, hence leading to an excess of halos as compared to a WDM model with a similar initial suppression.

The upper panel of Fig. 3.3.3 makes clear that the presence of the DAO feature can change the halo mass function by orders of magnitude compared to the simpler WDM case. Indeed, while the sDAO model with $h_{\rm peak} = 1$ and the WDM ($h_{\rm peak} = 0$) model deviate from the CDM case in a very similar fashion near $\mathrm{M} = 10^{10} h^{-1} \mathrm{M}_{\odot}$, the sDAO model then reconverges towards the CDM mass function, resulting in an abundance of $10^7 h^{-1} \mathrm{M}_{\odot}$ haloes that is more than 2 orders of magnitude greater than that of WDM. The peculiar shape of the ETHOS halo mass function means that constraints on dark matter physics based on the abundance of small-scale structure (using, e.g. , lensing or satellites) cannot straightforwardly be applied to these models.

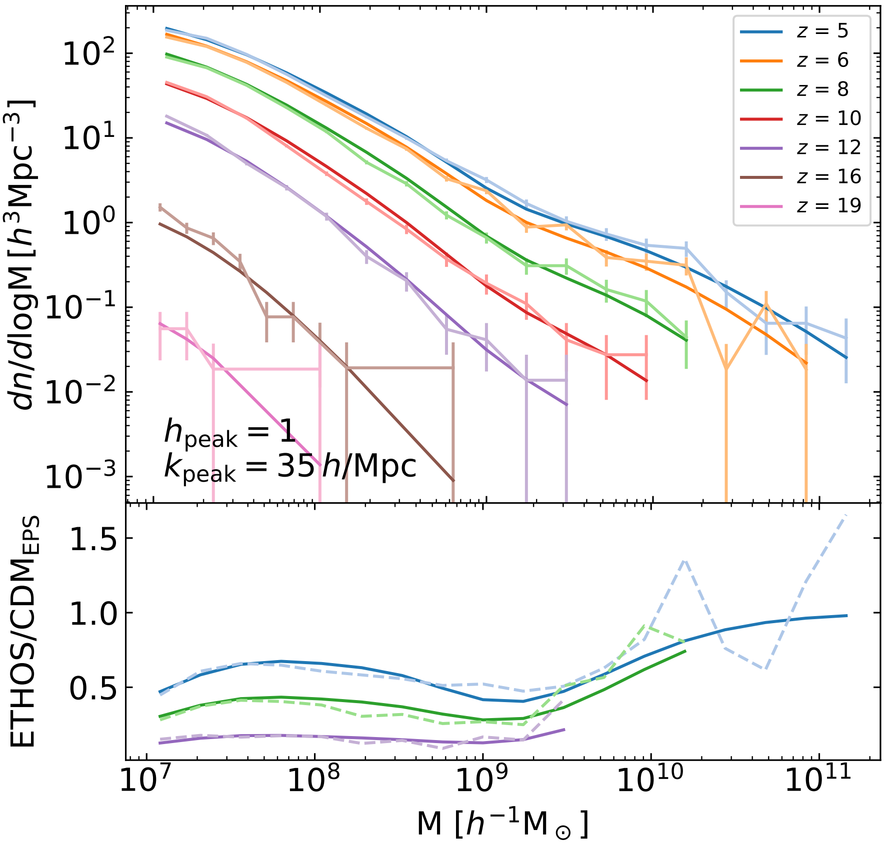

Figure 3.3.5. Halo mass function for a sDAO model ($h_{\rm peak} = 1$, $k_{\rm peak} = 35\,h\,{\rm Mpc}^{-1}$) at different redshifts ($z \geq 5$) according to the different colours in the legend. The light coloured lines with error bars are the simulation results, while the dark coloured lines are produced with the EPS model as described in the text and the caption of Fig. 3.3.2. The bottom panel shows the ratio of the sDAO model with respect to the CDM EPS result for a selection of redshifts as given by the colours, with the case of the simulations (EPS models) shown with dashed (solid) lines.

3.4 The inner structure of ETHOS haloes at high redshift ($z = 5$)

Having described and analysed the abundance of haloes at high-z within ETHOS in the context of the EPS formalism, we now look at the dark matter distribution within these haloes. In particular, we study the spherically-averaged density profile of ETHOS haloes at high-z and focus on the concentration-mass relation in the context of the halo assembly model of Ludlow et al. (2016).

3.4.1 Density profile

The near-universality of CDM haloes has been well established since the seminal papers of Navarro et al. (1996, 1997). The well-known 2-parameter Navarro-Frenk-White (NFW) profile has been shown to be a remarkably well fit to the spherically-averaged radial density profile of CDM haloes. Although more recent, higher resolution simulations show that other profiles such as the Einasto profile provide an even better fit to the structure of simulated haloes (e.g. Springel et al. 2008), the simplicity and accuracy of the NFW profile remains valid. This is particularly relevant when one considers that the NFW profile effectively becomes a function of one free parameter since there is a tight correlation (monotonically decreasing) between the virial mass of the halo and its concentration (e.g. Bullock et al., 2001; Eke et al., 2001; Neto et al., 2007; Prada et al., 2012; Ludlow et al., 2014; Sánchez-Conde & Prada, 2014; Klypin et al., 2016; Ishiyama et al., 2020; Wang et al., 2020). The concentration parameter for NFW haloes is defined as $c = r_{200}/r_{\rm s}$, where r_{200} is the virial radius, and $r_{\rm s}$ is the scale radius, which for the NFW profile coincides with r_{-2}, the radius where the logarithmic slope of the density profile is -2.

The near-universality of the NFW profile extends to the WDM case as well, where it has been shown that WDM haloes are also well described by this profile, albeit with lower concentration than the CDM counterpart at fixed mass (e.g. Lovell et al., 2014; Bose et al., 2016a). We thus begin this section by analysing if the NFW profile provides a good fit to ETHOS haloes in general, that is, we explore if the near-universality of this profile extends as well to models with DAOs. To quantify this we create density profiles for all haloes with at least 5000 particles[2] in a given ETHOS model using concentric shells from the centre of each halo, defined from the minimum of the halo potential. The shells are binned logarithmically in the range $6\varepsilon/r_{200} < r/r_{200} < 3$, where six times the smoothing length ε is the convergence limit of our simulations (see Appendix 3.6.2). We then fit the simulated profiles with the NFW profile by minimizing the following quantity:

$$Q^2 = \frac{1}{N_{\rm bins}} \sum_{i=1}^{N_{\rm bins}} \left(\ln\rho_i - \ln\rho_i^{\rm NFW}\right)^2, \tag{3.11}$$

where $N_{\rm bins} = 50$ (see Navarro et al., 2010). The left panel of Figure 3.4.6 shows the Q^2 distribution for all haloes with more than 5000 particles for the CDM, WDM, wDAO and sDAO models. While a slightly higher fraction of CDM haloes are in the smallest Q^2 bin, the shape and width of the distribution is quite similar across all DM models. If we only look at relaxed haloes (right panel of Fig. 3.4.6; see Appendix 3.6.3 for the relaxation criteria we used), we have a narrower distribution with an even higher fraction of haloes with small Q^2, which are therefore well described by a NFW profile.

3.4.2 Concentration-mass relation

Ludlow et al. (2016) developed an analytic model for the concentration-mass relation

---

[2] This limit is used to obtain a robust sampling of the spatial structure of a halo.

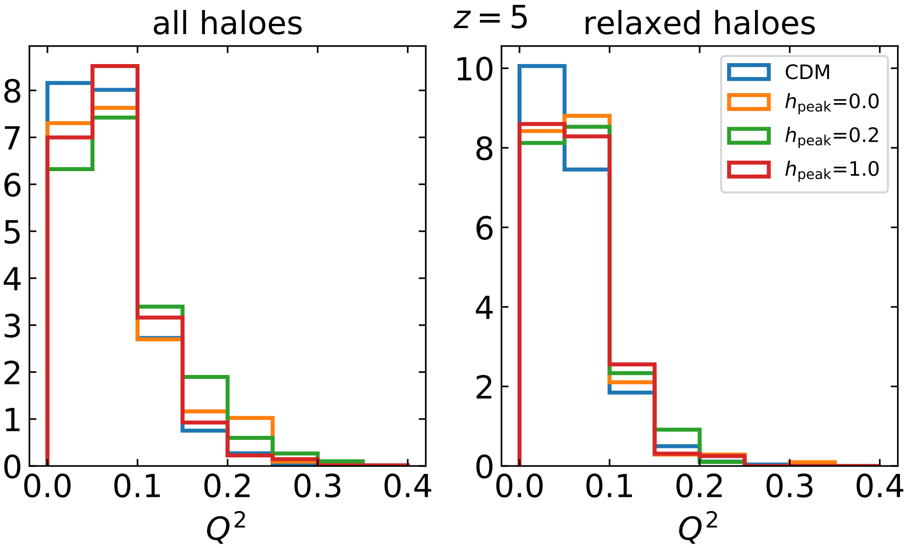

Figure 3.4.6. Goodness of the fit of NFW profiles for CDM, WDM ($h_{\rm peak} = 0$), wDAO and sDAO models (all three non-CDM models with $k_{\rm peak} = 35\,h\,Mpc^{-1}$) measured by Q^2 (Eq. 3.11). The left panel shows all simulated haloes with more than 5000 particles, while the right panel considers in addition only relaxed haloes (according to the criteria in Appendix 3.6.3).

based on EPS theory and applied it to CDM and WDM models (this model is an extension of the one developed earlier in Ludlow et al. 2014). Their model assumes that the mean inner density within the scale radius $\langle\rho_{-2}\rangle$ is proportional to the critical density of the Universe at an assembly redshift z_{-2}:

$$C\left(\frac{H(z_{-2})}{H(z_0)}\right)^2 = \frac{\langle\rho_{-2}\rangle}{\rho_0} = 200c^3\frac{\ln(2)-0.5}{\ln(1+c)-c/(1+c)}, \qquad (3.12)$$

where the second equality is only valid for NFW profiles, C is a free parameter.Secondly, the model assumes that the assembly redshift is defined as the redshift when the enclosed mass within the scale radius M_{-2} of the descendant halo was first assembled into progenitors having a mass larger than a fraction f of the descendant, and is given by

$$\mathrm{erfc}\left(\frac{\delta_c(z_{-2})-\delta_c(z_0)}{\sqrt{2(\sigma^2(f\times M)-\sigma^2(M)}}\right) = \frac{M_{-2}}{M_0} = \frac{\ln(2)-0.5}{\ln(1+c)-c/(1+c)}, \qquad (3.13)$$

where the second equality is only valid for NFW haloes and $\delta_{\rm c}(z) = \delta_{\rm c}/D(z)$ is the redshift dependent critical density for collapse. The left hand side of Eq. 3.13 corresponds to the collapsed mass fraction in EPS theory (Lacey & Cole, 1993). Across the paper, we use $C = 575$ and $f = 0.02$ for the free parameters in Eqs. (3.12)−(3.13).

Figures 3.4.7−3.4.9 show the concentration-mass relation at $z = 5$ for CDM, WDM, wDAO and sDAO models obtained from our simulations and the analytic model de-

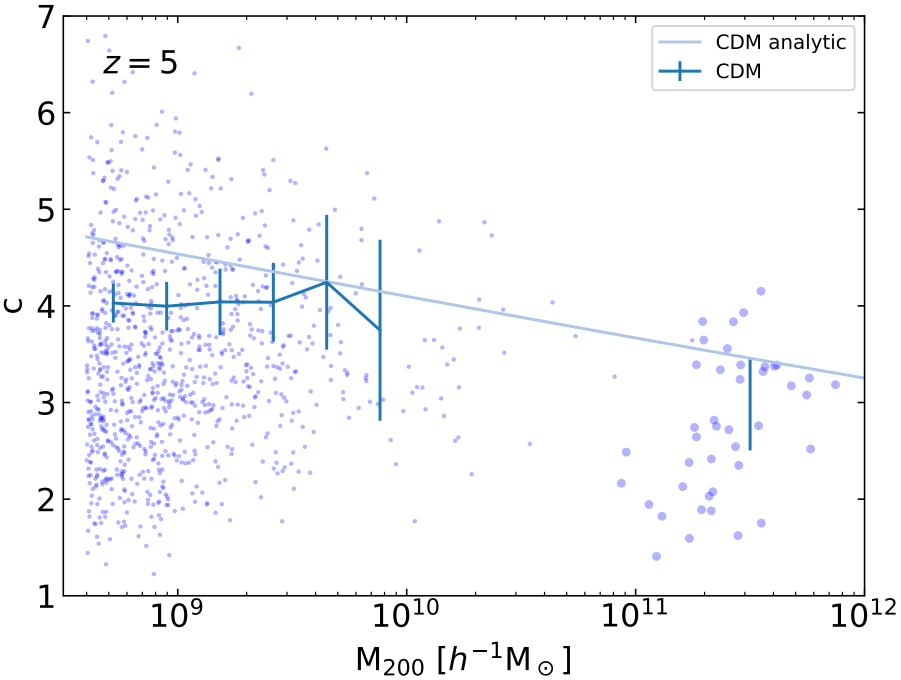

Figure 3.4.7. Concentration mass relation for CDM haloes at $z = 5$. The scatter points represent individual measurements of the NFW concentration for all relaxed CDM haloes while the lines with Poisson error bars correspond to the median concentration at that mass. The median is computed at four equally sized bins for $M = 10^9 - 10^{10}\, M_\odot/h$ from high-res haloes and one bin for $M = 10^{11} - 10^{12}\, M_\odot/h$ from low-res haloes. The lines correspond to the concentration computed using the analytic model of Ludlow et al. 2016 with a smooth-k space filter (see Eq. (3.13)*).*

scribed above using a smooth-k filter for all models. For these plots, we only took relaxed haloes into account and binned the haloes in the high-resolution region (hereafter high-res haloes) into four bins, equally sized in logarithmic mass bins in the range $10^9 - 10^{10}\,\mathrm{M}_\odot/h$. We obtained the concentration for each bin by taking the median density profile of all relaxed haloes within that bin (stacking the profiles by re-scaling the radius to the virial radius and only up to $r/r_{200} = 0.8$) and fitting a NFW profile to it. The irregularities of individual haloes are smoothed out in this way. Additionally, we have used data from the low-resolution regions of the simulations (hereafter low-res haloes), combining the low-res haloes with $\mathrm{M} = 10^{11} - 10^{12}\,\mathrm{M}_\odot/h$ into one mass bin to serve as high-mass anchor point, when comparing the analytic model to our simulations.

For CDM, we use the smooth-k filter when calculating the mass variance σ^2 in Eq. (3.13) and Fig. 3.4.7 shows that, as expected, the analytic model and our simulations are in reasonable agreement within the Poisson sampling errors. However, at low masses we observe that the concentration remains flat in the simulations, while the model predicts a monotonically decreasing concentration.

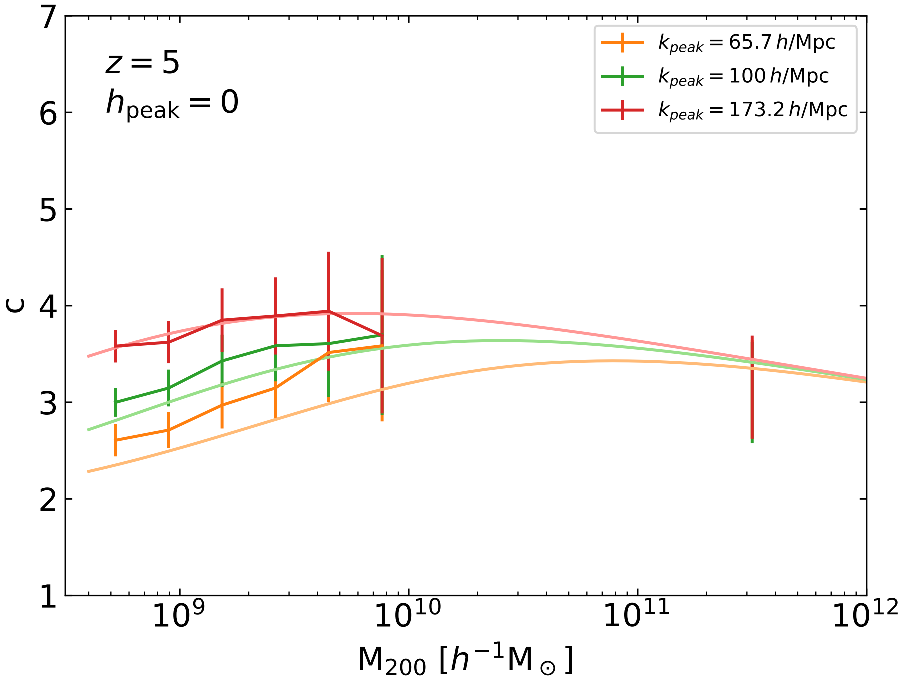

Figure 3.4.8. Concentration mass relation for WDM haloes at $z = 5$. The lines with Poisson error bars correspond to the median concentration of haloes at that mass. The median is computed at four equally sized bins for $M = 10^9 - 10^{10}\,M_\odot/h$ from high-res haloes and one bin for $M = 10^{11} - 10^{12}\,M_\odot/h$ from low-res haloes. The lines correspond to the concentration computed using the analytic model of Ludlow et al. 2016 with a smooth-k space filter (see Eq. 3.13).

For WDM (Fig. 3.4.8), we also use the smooth-k filter and the simulations are captured well in this way. The low-mass behaviour is followed closely by the analytic model and the high-mass anchor point is also in good agreement. Even though the analytic model and the simulation results agree within the Poisson noise, we notice a trend that the model under-predicts the concentration for small $k_{\rm peak}$ and over-predicts for large $k_{\rm peak}$. This suggests that the scaling with $k_{\rm peak}$ is not captured completely accurately by the model. However, the number of haloes in our simulations is not large enough to fully trust this trend. We note that the WDM model with $k_{\rm peak} = 35\,h\,{\rm Mpc}^{-1}$shows very high concentration values for low mass haloes for which we suspect numerical issues, as this is the model with the most extreme suppression of small scale power. Therefore, we have omitted this model in Fig. 3.4.8. We note here, that Ludlow et al. (2016) originally tested their analytic approach for WDM models at lower redshifts than those in our simulations ($z = 0 - 3$). The authors found a good agreement, although their simulation with the smallest WDM particle mass $m_{\rm WDM} = 1.5$ keV suggests some inconsistencies at low masses for $z = 3$ (green line in Fig.4 of Ludlow et al. 2016).

Fig. 3.4.9 shows the concentration-mass relation for wDAO (top) and sDAO models

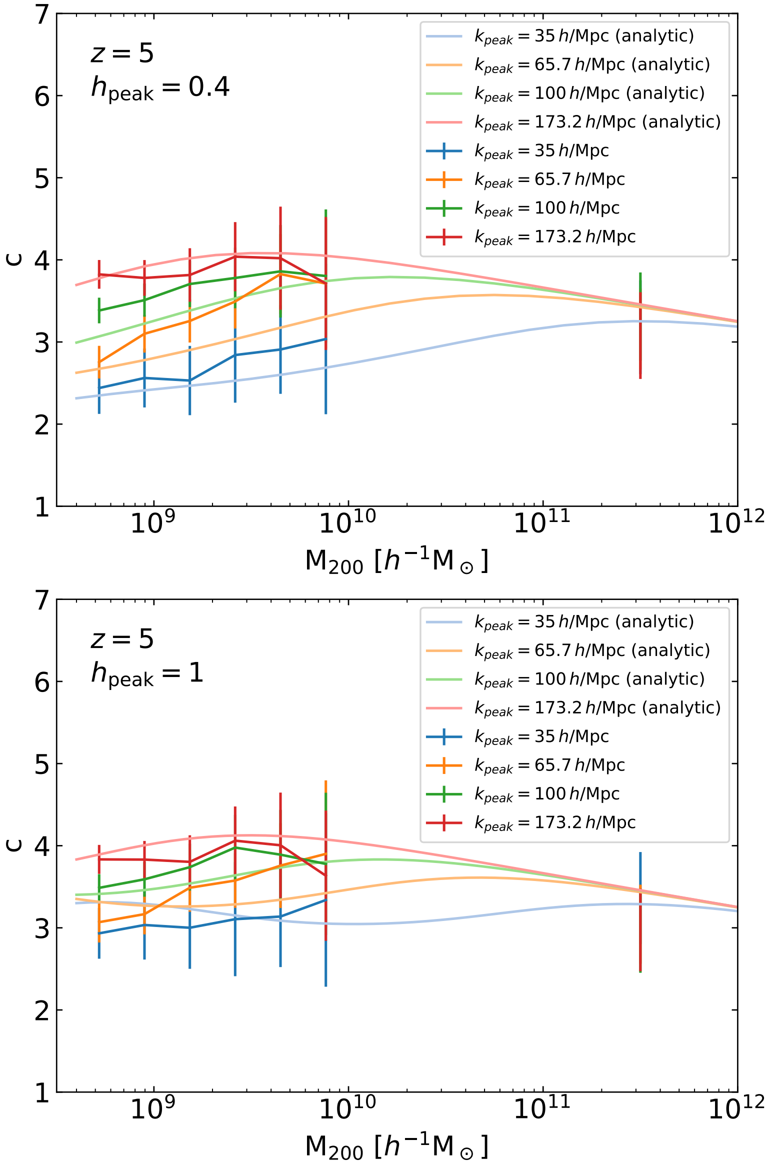

Figure 3.4.9. Concentration mass relation for wDAO (top) and sDAO (bottom) haloes at $z = 5$. The lines with Poisson error bars correspond to the median concentration of haloes at that mass bin. The median is computed at four equally sized bins for $M = 10^9 - 10^{10}\,M_\odot/h$ from high-res haloes and one bin for $M = 10^{11} - 10^{12}\,M_\odot/h$ from low-res haloes. The lines correspond to the concentration computed using the analytic model of Ludlow et al. 2016 with a smooth-k space filter (see Eq. 3.13).

(bottom) using the smooth-k space filter (Eq. 3.10). As with WDM, the analytic prediction for the DAO models is in reasonable agreement with the simulation. In both cases, we see a similar trend to under-predict the concentration for small $k_{\rm peak}$ and over-predict for large $k_{\rm peak}$. For small $k_{\rm peak}$ however, the concentration of sDAO haloes is predicted to increase again towards the smallest masses, which we do not observe in our simulations. Therefore, it seems that the sDAO features of the most extreme models are not correctly captured by the analytic model of Ludlow et al. (2016). We notice however, that across all DAO models, the decrease (soft cutoff) of concentration towards intermediate masses predicted by the analytic model is seen in our simulations, while a predicted increase at lower masses in the sDAO cases is not present in the simulations. We notice that the latter trend (continuous increase of concentraton at low masses), is actually not seen neither in the CDM case (see Fig. 3.4.7). The over-prediction might therefore not be an explicit problem of the sDAO model.

We can conclude that Figs. 3.4.8 and 3.4.9 demonstrate that the reduced small-scale power in ETHOS models reduces the concentration of small haloes depending on the value of $k_{\rm peak}$. Furthermore Eq. (3.13), based on the analytic model of Ludlow et al. 2016 (ultimately based on EPS theory), can capture the concentration-mass relation reasonably well for a wide range of DM models. However, the model struggles to reproduce the small mass behaviour for the most extreme sDAO models and the trend indicates that the scaling with $k_{\rm peak}$ is not captured correctly. We remark that in order to improve the analytic model, simulations with a larger volume but similar resolution are needed to reduce the sampling errors, while also covering higher mass haloes with $\mathrm{M} > 10^{10}\,\mathrm{M}_{\odot}/h$.

3.5 Conclusions

Performing dedicated cosmological N-body simulations to extract basic but precise measurements of the properties of haloes for specific DM models requires access to HPC resources, which can be computationally expensive when a broad exploration of models is desired. Such a broad exploration is essential to cover the range of alternatives to the Cold Dark Matter (CDM) model, which predict a halo population with distinct properties. A relevant category of such alternatives is that of models with a (galactic-scale) primordial cutoff in the linear power spectrum, caused by either the free streaming mechanism (Warm Dark Matter, WDM) or by collisional damping with relativistic species in the early Universe (models with Dark Acoustic Oscillations, DAOs). The difference between these models and CDM increases at low halo masses, which are more affected by the small-scale suppression of power, particularly at high redshift.

For these reasons, in this work we take the simulation suite from Bohr et al. (2020) within the ETHOS framework (Cyr-Racine et al., 2016; Vogelsberger et al., 2016), which covers both WDM and DAO models, to investigate the abundance and inner structure of dark matter haloes at high redshift ($z \geq 5$). Bohr et al. (2020) presented a convenient parametrization of these different structure formation models based only on two parameters $h_{\rm peak}$ and $k_{\rm peak}$, the amplitude and scale of the first DAO peak. CDM and WDM are both included in this parametrization by taking $k_{\rm peak} \rightarrow \infty$ in the former and $h_{\rm peak} = 0$ in the latter. Specifically, our objective is mainly to describe the behaviour of i) the halo mass function and ii) the halo concentration-mass relation across the ETHOS models in the simulation suite, and to interpret the results based on the Extended Press-Shechter (EPS) formalism. The latter objective is particularly relevant since it offers an alternative to quickly compute statistical halo properties, which have so far not been fully tested across the broad range of dark matter models explored in the ETHOS framework.

We have shown that the EPS formalism within the ellipsoidal collapse model (Eqs. 3.4−3.5) using a smooth-k window function (Eq. 3.10) with the fitting parameters $\beta = 3.46$, $c_{\rm W}$=3.79 is able to accurately reproduce the halo mass function (in the redshift range $5 \leq z \lesssim 19$ and mass range $10^7\,{\rm M}_\odot/h \lesssim M_{200} \lesssim 10^{11}\,{\rm M}_\odot/h$) for CDM and ETHOS models with $h_{\rm peak} = 0.6 - 1$ (see Figs. 3.3.2−3.3.5). For models with weaker DAO features ($h_{\rm peak} < 0.6$), the cut-off in the halo mass function is reproduced accurately and the overall behaviour at lower halo masses is well captured, but the accuracy below the cut-off scale is much lower than in models with higher $h_{\rm peak}$ (see Fig. 3.3.4).

Regarding halo structure, we found that the haloes of all ETHOS models at $z = 5$ are well described by an NFW profile (see Fig. 3.4.6). The smaller the value of $k_{\rm peak}$, the lower the halo concentration towards lower halo masses relative to the CDM case. As can be seen in Figs. 3.4.7−3.4.9, the (median) concentration-mass relation at $z = 5$ for most of the ETHOS simulations is well reproduced with the analytic model based on the EPS formalism introduced in Ludlow et al. (2016) (tested there only for CDM and WDM). However, the most extreme DAO models (strong DAOs, where $h_{\rm peak} \sim 1$) have a measured concentration-mass relation that behaves differently than the analytic model towards low halo masses ($M_{200} \lesssim 10^9\,{\rm M}_\odot/h$); albeit our limited sampling of haloes (due to the small volume of our simulations) carries counting errors that remain too large to fully quantify the level of disagreement between the simulations and the analytic model. Simulations within a larger cosmic volume and with a larger mass range coverage are needed to firmly conclude whether an improved analytic model is needed to capture the concentration-mass relation in the full spectrum of ETHOS models.

In this work we have thus shown that it is possible to use analytic models based on the EPS formalism to reproduce the halo mass function essentially in the whole spectrum of relevant ETHOS models, that is, covering CDM, WDM and DAO models that have (allowed) galactic-scale cutoffs. This analytic prescription calibrated to our simulations has already been used in Muñoz et al. (2021) to make predictions for the 21-cm hydrogen line signal during the cosmic dawn ($z \sim 10 - 30$). We have also shown that a similar analytic approach (based on Ludlow et al., 2016) is able to reproduce the halo concentration-mass relation, albeit care is needed at low-masses where the reliability of the model remains unclear.

The difference between the halo mass functions across currently allowed ETHOS models will become increasingly important in the near future, when a detection/constraint in the relevant mass range becomes more feasible with upcoming observing facilities. The James Webb Space Telescope (JWST) will likely be able to probe the halo mass function indirectly through the luminosity function and test the viability of a large range of ETHOS models (see e.g. Lovell et al., 2018a, for a study of a specific wDAO model). The hydrogen epoch-of-reionization array (HERA) will offer another promising approach to (indirectly) distinguish the different halo mass functions of ETHOS models through observation of the 21-cm signal (see e.g. Muñoz et al., 2021, for predictions based directly in the simulation suite and EPS modelling presented in this paper).

Acknowledgements: SB and JZ acknowledge support by a Grant of Excellence from the Icelandic Research Fund (grant number 173929). MV acknowledges support through NASA ATP grants 16-ATP16-0167, 19-ATP19-0019, 19-ATP19-0020, 19-ATP19-0167, and NSF grants AST-1814053, AST-1814259, AST-1909831 and AST-2007355. The simulations were performed on resources provided by the Icelandic High Performance Computing Centre at the University of Iceland, and the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

3.6 Appendix

3.6.1 Halo mass function of WDM models

Figure 3.6.10 shows a comparison between the redshift evolution of the halo mass function of the most extreme WDM model in our simulations and the result of the EPS formalism computed with the smooth-k (eq. 3.10) and the sharp-k window functions. The sharp-k window function is given by

$$\tilde{W}_R^{\mathrm{sharp}-k}(k) = \theta(1 - kR), \tag{3.14}$$

where θ is the Heaviside step function. In the case of the sharp-k window function, the sharp cut-off in the window function leads to a sharp cut-off in the halo mass function, which is clearly not a feature we resolve in our simulations before the appearance of spurious haloes for masses a few times $10^8\,\mathrm{M}_\odot/h$. Given this limitation in resolution, and given the limited sampling of haloes in our simulations for the models with the strongest cutoffs in the power spectrum (low k_{peak}; particularly at high redshift), it is not possible to convincingly establish which of the window function achieves a better modelling of the halo mass function.

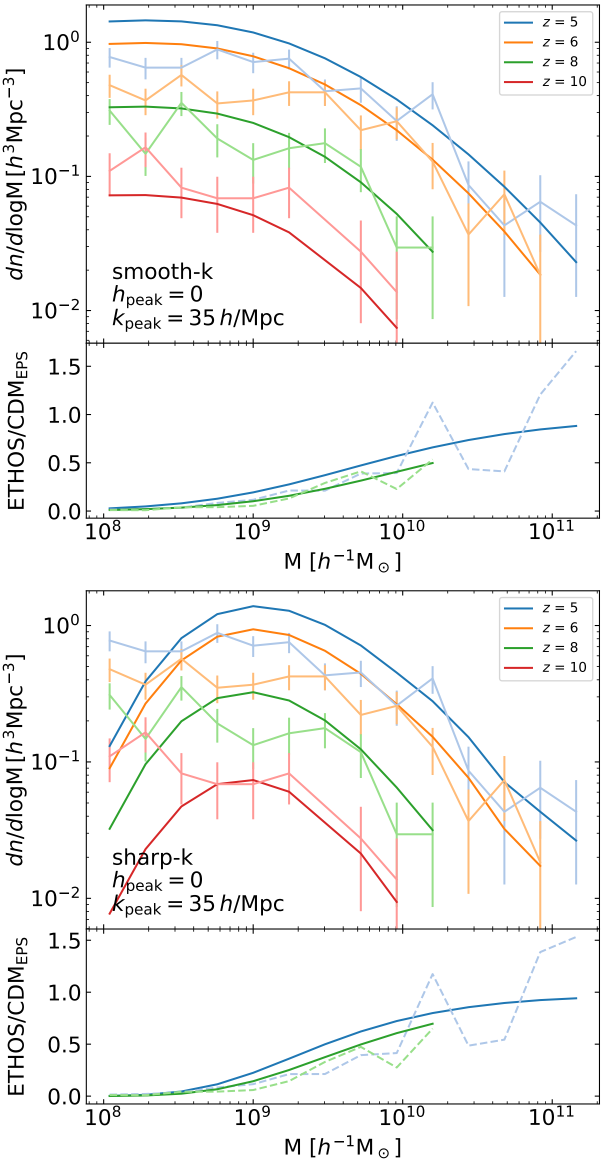

Figure 3.6.10. Halo mass function for our most extreme WDM model ($h_{\rm peak} = 0$, $k_{\rm peak} = 35\,h\,\mathrm{Mpc}^{-1}$) at different redshifts ($z \geq 5$) according to the different colours in the legend. The light coloured lines with error bars are the simulation results, while the dark coloured lines are the EPS model using the smooth-k window function (upper panels) and the sharp-k window function (bottom panels). The small bottom panels for each case show the ratio of the WDM model with respect to the CDM EPS result for a selection of redshifts according to the colour legend.

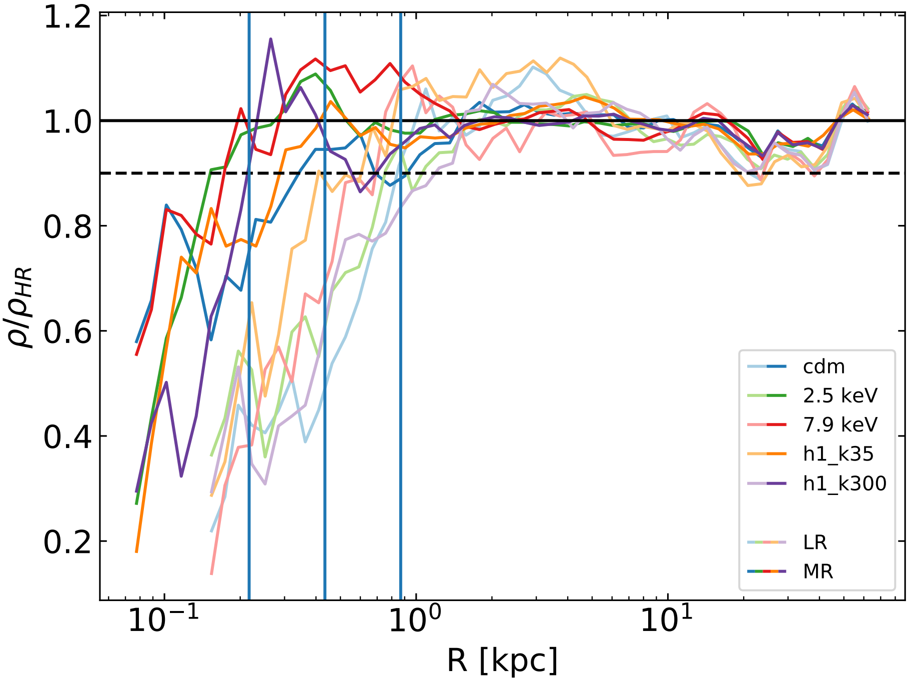

Figure 3.6.11. Convergence of the halo density for 5 models covering our parameter space. The vertical axis is the ratio of the density of the low-resolution (faded lines) and medium resolution (solid lines), relative to that of the highest resolution run at $z = 5$. The dashed line indicates a convergence level of 10%. The blue vertical lines indicate six times the softening length for high-resolution to low-resolution from left to right.

3.6.2 Convergence test for halo density profiles

To determine the smallest radius at which we can trust the density profile measured in our simulations, we compare the density profile of the largest halo for a few DM models using three resolution levels. The chosen models cover representative regions of the parameter space and show the range of possible convergence levels. The three resolution levels were done with smoothing lengths $\varepsilon = 0.87$ (LR), 0.43 (MR), and $0.22\,\mathrm{ckpc}/h$ (HR). As the information at the smallest scales in a halo is absent due to limited resolution, we looked for the radius below which the density in the two lower resolution levels drops by more than 10% with respect to the highest resolution. Figure 3.6.11 shows the ratio of the density profile between the MR and the HR levels (solid lines) and between the LR and HR levels (faded lines). At large radii, this ratio fluctuates only slightly around 1 and then drops substantially below six times the smoothing length, which is indicated by the vertical lines. We have therefore assigned 6ε as the smallest resolvable scale in our high resolution simulations.

3.6.3 Relaxation criteria for haloes

The assembly of haloes is a very dynamic process, but the NFW profile describes a halo in equilibrium. Substantial departures from equilibrium in a halo would result in substantial deviations over the NFW profile. Therefore, we have to clean our halo catalogue by selecting only the haloes that are sufficiently virialised, have a subhalo population that is clearly subdominant by mass, and are not currently in the process of merging with a massive substructure. We adopt the relaxation criteria of Neto et al. (2007) to accomplish these goals:

- The mass fraction in subhaloes must be low $f_{\mathrm{sub}} = \mathrm{M}_{\mathrm{sub}}/\mathrm{M}_{200} < 0.1$, where M_{sub} is the total mass in subhaloes.
- The distance between the minimum of the potential and center of mass of the halo must be small compared to the virial radius $d_{\mathrm{off}} = |r_{\mathrm{pot}} - r_{CM}|/R_{200} < 0.07$
- The virial ratio must be close to virialization $2|T/U|$<1.5, where T and U are the total kinetic and potential energies, respectively.

We note that we have relaxed the threshold for $2|T/U|$ from 1.35 to 1.5 as we are considering a higher redshift than those studies that typically used these criteria (such as Neto et al. 2007). At high redshift, we expect haloes to be less virialized than at low redshift (see e.g. Fig.3 in Zjupa & Springel, 2017).

Figure 3.6.12 shows the distribution of the three relaxation criteria for the haloes of sDAO, wDAO, and WDM models with $k_{\mathrm{peak}} = 35\,h\,\mathrm{Mpc}^{-1}$, as well as CDM. We can see that a smaller number of haloes are virialized in non-CDM models than in CDM. This trend probably arises from the delayed halo formation in models with a galactic-scale cutoff. The suppression of small scale structure in the case of WDM or wDAO is also clearly visible in the distribution of f_{sub}. However, we find that the most restrictive criteria is given by the limit in the d_{off} value; clearly a substantial fraction of haloes at these redshifts are actively merging.

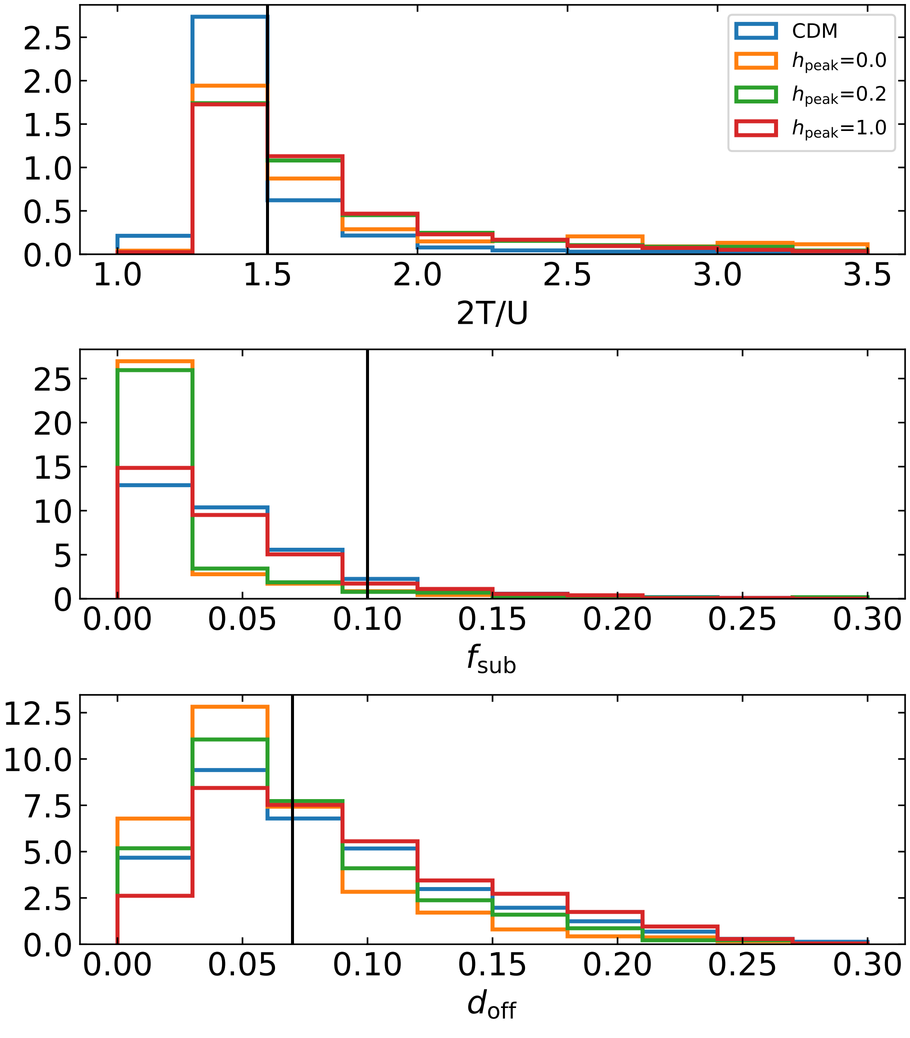

Figure 3.6.12. The distribution of the three relaxation criteria $2|T/U|$, f_{sub} *and* d_{off}. *The black vertical lines indicate our used thresholds for determining relaxed haloes. All ETHOS models have* $k_{\mathrm{peak}} = 35\,h\,Mpc^{-1}$.

Chapter 4

Dark acoustic oscillations during cosmic dawn

This chapter is based on the following article:

> **ETHOS - An Effective Theory of Structure Formation: Impact of Dark Acoustic Oscillations on Cosmic Dawn**

Published in Physical Review D, Volume 103, Issue 4, article id.043512 (2021)

Authors:
Julian B. Muñoz[1,2], Sebastian Bohr[3], Francis-Yan Cyr-Racine[4], Jesús Zavala[3] and Mark Vogelsberger[5]
[1]Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA
[2]Department of Physics, Harvard University, 17 Oxford St., Cambridge, MA 02138, USA
[3]Centre for Astrophysics and Cosmology, Science Institute, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland
[4]Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87106, USA
[5]Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Upcoming data of the 21-cm hydrogen line during cosmic dawn ($z \sim 10-30$) will revolutionize our understanding of the astrophysics of the first galaxies. Here we present a case study on how to exploit those same measurements to learn about the nature of dark matter (DM) at small scales. Focusing on the Effective Theory of Structure Formation (ETHOS) paradigm, we run a suite of simulations covering a broad range of DM microphysics, connecting the output of N-body simulations to dedicated 21-cm simulations to predict the evolution of the 21-cm signal across the entire cosmic dawn. We find that observatories targeting both the global signal and the 21-cm power spectrum

are sensitive to all ETHOS models we study, and can distinguish them from CDM if the suppression wavenumber is smaller than $k \approx 300\,h/\mathrm{Mpc}$, even when accounting for feedback with a phenomenological model. This is an order of magnitude smaller comoving scales than currently constrained by other data sets, including the Lyman-α forest. Moreover, if a prospective 21-cm detection confirmed a deficiency of power at small scales, we show that ETHOS models with strong dark acoustic oscillations can be discriminated from the pure suppression of warm dark matter, showing the power of 21-cm data to understand the behavior of DM at the smallest physical scales.

4.1 Introduction

The majority of matter in our universe is dark, and seemingly collisionless (Davis et al., 1981; Blumenthal et al., 1982, 1984; Davis et al., 1985; Peebles, 1984; Bertone et al., 2005). Decades of observational efforts have provided us with increasingly precise constraints on the nature of dark matter (DM) (Abbott et al., 2018; Planck Collaboration et al., 2020; Murgia et al., 2018; Nadler et al., 2019; Nadler et al., 2021; Vogelsberger et al., 2020), albeit not a solution to its nature yet. An exciting possibility is that a complex dark sector hosts dark matter, as well as other components, which may interact with each other throughout cosmic history (Foot, 2004; Ackerman et al., 2009; Arkani-Hamed et al., 2009; Feng et al., 2009; Kaplan et al., 2010; Behbahani et al., 2011; Kaplan et al., 2011; van den Aarssen et al., 2012; Cline et al., 2012; Hooper et al., 2012; Das & Sigurdson, 2012; Cyr-Racine & Sigurdson, 2013; Diamanti et al., 2013; Baldi, 2013; Fan et al., 2013a,b; McCullough & Randall, 2013; Cline et al., 2014a,b; Bringmann et al., 2014; Chu & Dasgupta, 2014; Archidiacono et al., 2015; Randall & Scholtz, 2015; Buen-Abad et al., 2015; Lesgourgues et al., 2016; Choquette & Cline, 2015).

Searching for these dark-sector interactions between DM and light degrees of freedom, while impossible in the lab, is feasible with astrophysical data sets (see e.g. Tulin & Yu 2018). DM interactions can leave an imprint on the formation of cosmic structure, which can be searched with precision cosmic data sets such as the cosmic microwave background (CMB) and large-scale structure (LSS) of the universe (Cyr-Racine et al., 2014; Archidiacono et al., 2017, 2019). Past analyses have shown these cosmological data sets to be broadly consistent with the standard cold dark matter (CDM) paradigm on large scales. Any significant departure from the “vanilla" CDM behavior thus ought to appear preferentially at smaller scales. In this regime, observations of the Lyman-α forest (Murgia et al., 2018), of the luminosity function of Milky Way satellites (Nadler et al., 2019; Nadler et al., 2021), and of flux-ratio anomalies of multiply imaged strongly lensed quasars (Dalal & Kochanek, 2002; Gilman et al., 2019; Hsueh et al., 2020; Gilman et al., 2020) have shown consistency with CDM on halo mass scale $\gtrsim 10^9 M_\odot$. Pushing this boundary to even smaller scales is a major goal of a current and future multi-pronged effort (see e.g. Drlica-Wagner et al. 2019).

A telltale signature of DM interacting with light degrees of freedom in the early Universe is the presence of dark acoustic oscillations (DAOs) in the linear matter power spectrum. Detailed simulations (Buckley et al., 2014; Schewtschenko et al., 2015; Vogelsberger et al., 2016) of the nonlinear evolution of structure within such models have shown that this key signature gets partially erased as power is regenerated on small scales at late times. Therefore, observations at higher redshifts have the potential to probe DAOs and their effect on structure formation in a more pristine state. One of the earliest probe of nonlinear structure formation in our Universe is the 21-cm signal from cosmic dawn. At that epoch, the ultraviolet (UV) radiation emitted by the first stars recouples the neutral hydrogen spin temperature to that of the cooler gas via the Wouthuysen-Field effect (Wouthuysen, 1952; Field, 1959; Hirata, 2006), leading to a net absorption of 21-cm photons from the Rayleigh-Jeans tail of the CMB. Since early stellar formation depends sensitively on the abundance and properties of small DM halos with mass $M_{\rm h} \sim 10^6 - 10^8 M_{\odot}$, the timing and shape of this absorption feature can be used to search for the presence of DAOs and related damping on those scales.

In general, any model which suppresses or modifies the amplitude of DM fluctuations on small scales could affect the 21-cm cosmic dawn signal (see e.g. Lopez-Honorez et al. 2016; Schneider 2018; Lopez-Honorez et al. 2019; Escudero et al. 2018; Muñoz et al. 2018; Muñoz & Loeb 2018; Yoshiura et al. 2020; Mena et al. 2019; Muñoz et al. 2020). Exploring the 21-cm signal from this broad parameter space of possible DM models can be quite costly since it generally requires detailed simulations. A promising approach is to map the different DM microphysics to effective parameters that govern how structure forms. The effective theory of structure formation (ETHOS; Cyr-Racine et al. 2016; Vogelsberger et al. 2016) provides such a mapping. It naturally interpolates between DM models having sharp transfer function cutoff such as warm DM (WDM) to theories displaying damped or strong acoustic oscillations, and to models looking nearly like CDM. So far, the ETHOS framework has been used to study the satellite galaxies of Milky Way-like hosts (Vogelsberger et al., 2016), the high-redshift UV luminosity function and reionization (Lovell et al., 2018b), and the impact of DAOs on Lyman-α forest signal (Bose et al., 2019).

In this paper, we use the simple but powerful phenomenological ETHOS parametrization introduced in Bohr et al. (2020) to describe deviations from the standard CDM scenario and compute the expected 21-cm signal from cosmic dawn. This two-dimensional parameter space spans a broad range of models ranging from WDM and models with suppressed DAOs, to models displaying strong DAOs and theories that are phenomenologically undistinguishable from CDM. Using this parametrization, we compute both the expected 21-cm global signal and power spectrum and study the distinguishability of different dark matter models in upcoming experiments.

This paper is structured as follows. In Sec. 4.2 we describe the ETHOS parametrization and the N-body simulations we use. We show the effect of the different ETHOS models on the 21-cm global signal in Secs. 4.3, and on the 21-cm fluctuations in 4.4. We conclude in Sec. 4.5.

4.2 The ETHOS framework and simulations

Here we describe the matter power spectrum within the ETHOS framework, and the simulations that we use.

4.2.1 Effective parametrization

The ETHOS paradigm was developed to capture the effects of DM microphysics on the formation of structure in our universe in a few convenient parameters (Cyr-Racine et al., 2016). Throughout this work we will employ the effective ETHOS parametrization introduced in Bohr et al. (2020), which provides a convenient—and accurate—shortcut to the full ETHOS parameter space.

This circumvents modeling the DM interactions, and instead approximates the matter power spectrum through two relevant parameters, which control the height $h_{\rm peak}$ and wavenumber $k_{\rm peak}$ of the first DAO peak, as illustrated in Fig. 4.2.1. In this notation the limit $h_{\rm peak} \to 0$ corresponds to WDM, whereas $h_{\rm peak} \to 1$ are strong DAOs. As an example, an atomic-DM model will have $h_{\rm peak} \to 0$ if diffusion damping occurs at larger scales than the DAOs, and $h_{\rm peak} \to 1$ if dark recombination occurs instantaneously. These two parameters capture the main features of the matter power spectrum for a large variety of ETHOS models (which include more details about the DM microphysics), and it was shown in Bohr et al. (2020) that the high-redshift halo mass function (HMF) is well approximated with only these two degrees of freedom. The acoustic origin of the DAOs determines the heights and locations of the subsequent peaks as a function of the first one for the models we study (see Bohr et al. 2020 for a detailed explanation).

The connection between these phenomenological parameters and particle physics model parameters (masses, couplings, etc.) is provided in Bohr et al. (2020). For instance, the $h_{\rm peak} = 0$ cases are equivalent to a WDM mass

$$\frac{m_{\rm WDM}}{1\,{\rm keV}} = \left[0.050\left(\frac{k_{\rm peak}}{h\,{\rm Mpc}^{-1}}\right)\left(\frac{\Omega_\chi}{0.25}\right)^{0.11}\left(\frac{h}{0.7}\right)^{1.22}\right]^{\frac{1}{1.11}}, \tag{4.1}$$

where Ω_χ is the DM abundance. We use the same models as Bohr et al. (2020) in this work, i.e., 48 simulations with $h_{\rm peak} = 0-1$ in steps of 0.2 and $k_{\rm peak} = 35-300\,h/{\rm Mpc}$ (where h is the reduced Hubble constant) with equidistant steps in $\log(k_{\rm peak})$ on the intervals [35,100] h/Mpc and [100,300] h/Mpc.

4.2.2 N-body Simulations

We run cosmological DM-only N-body simulations with the code AREPO (Springel, 2010) using the zoom-in technique described in Bohr et al. (2020) with a particle mass of $8\times10^4\,{\rm M}_\odot h^{-1}$ in the high-resolution region. The initial conditions are generated by the code MUSIC (Hahn & Abel, 2011) and the cosmological parameters of the simulations are $\Omega_{\rm m} = 0.31069$, $\Omega_\Lambda = 0.68931$, $H_0 = 67.5\,{\rm km/s/Mpc}$, $n_{\rm s} = 0.9653$ and

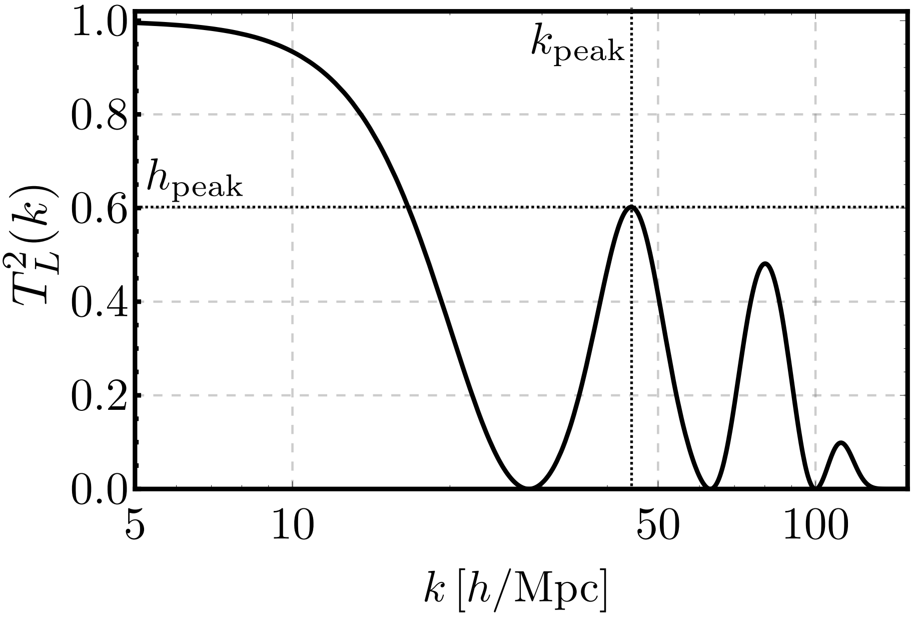

Figure 4.2.1. Diagram of the ETHOS parametrization of the power spectrum. Shown is the linear “transfer" function $T_L^2 = P_m^{\rm ETHOS}/P_m^{\rm CDM}$ as a function of wavenumber k. The two parameters determine the location $k_{\rm peak}$ and height $h_{\rm peak}$ of the first peak, where $h_{\rm peak} = 0$ corresponds to WDM with different masses, and $k_{\rm peak} \to \infty$ to CDM.

$\sigma_8 = 0.815$, where $\Omega_{\rm m}$ and Ω_Λ are the fraction of the matter-energy density of the Universe today, that is provided by matter and cosmological constant, respectively, H_0 is the Hubble constant today, $n_{\rm s}$ is the spectral index, and σ_8 is the mass variance of linear fluctuations in $8\,h^{-1}$ Mpc spheres at $z = 0$.

The output we will use are the HMFs measured at each redshift in the range $z = 10 - 25$ with redshift intervals $\Delta z = 0.3$, which are passed as an input to our modified version of `21cmvFAST`, as we will describe below. We find the HMF through counting the number of haloes identified by the friends-of-friends and Subfind algorithm in AREPO within the high-resolution region of the simulation.

4.2.3 Ingredients for the 21-cm Simulations

Let us now describe how we use the ETHOS results from above to find the evolution of the 21-cm signal across cosmic dawn. In this work we will use semi-numerical 21-cm simulations with a modified version of the public code `21cmvFAST` (Muñoz, 2019b,a)[1],

---

[1] `https://github.com/JulianBMunoz/21cmvFAST`

which itself is based on `21cmFAST` (Mesinger et al., 2011; Greig & Mesinger, 2015)[2]. Here, however, we do not assume the HMF of a CDM model. Instead, we use the HMF from the ETHOS simulations, denoted as dn/dM, to compute the fraction of baryons collapsed into stars as

$$F_{\rm coll} = \int_{\rm M_{cool}}^{\infty} dM \frac{M}{\rho_m} \frac{dn}{dM} \frac{f_g}{f_b} f_*(M), \tag{4.2}$$

f_b and f_g are the baryon and gas fractions, and f_* is the fraction of gas that gets converted onto stars. This integral runs over masses larger than $M_{\rm cool}$, which parametrizes the smallest halo that can form stars efficiently (note that an alternate parametrization exponentially suppresses low-mass haloes, instead of providing a sharp cut-off, providing similar results (Park et al., 2019)). Throughout this work we assume, for simplicity, that only haloes above the atomic-cooling threshold can form stars, i.e., $M_{\rm cool} = M_{\rm atom}(z)$ (Oh & Haiman, 2002). This provides a conservative estimate of the reach of cosmic-dawn data to probe ETHOS models, as smaller (molecular-cooling) minihaloes would be formed out of larger wavenumbers k, which are further affected by deviations from CDM for fixed astrophysical assumptions.

In practice we evaluate Eq. (4.2) by directly adding the mass of haloes above $M_{\rm cool}(z)$, to avoid errors induced by binning the HMF. We show the resulting $F_{\rm coll}$ as a function of redshift for all our ETHOS models, and CDM, in Fig. 4.2.2. As expected, this quantity grows exponentially for all models as the cosmic evolution makes fluctuations grow bigger, and more haloes form. However, models with low $k_{\rm peak}$ take significantly longer to form galaxies, shifting all their lines to lower z. We note, in passing, that for very low values of $F_{\rm coll}$ (corresponding to high redshifts) the Poisson noise is important for all models. This causes the $F_{\rm coll}$ curves of some ETHOS models to overcome the CDM case, albeit only briefly and at very high z.

As we neglect molecular-cooling haloes, the main source of feedback to consider is photo-heating, which can evaporate the gas within haloes (Efstathiou, 1992; Dijkstra et al., 2004). However, atomic-cooling haloes are not expected to be significantly affected by photo-heating feedback until $z \sim 10$ (Sobacchi & Mesinger, 2013; Qin et al., 2020; Qin et al., 2021), where we stop our simulations. To account for any residual feedback (such as due to SNe), we will implement a model where the gas fraction that turns into stars as (Trenti et al., 2010; Sitwell et al., 2014; Mason et al., 2015; Park et al., 2019).

$$f_*(M) = f_*^{(0)} \times \left(\frac{M}{M_0}\right)^{\alpha}, \tag{4.3}$$

where we take $f_*^{(0)} = 0.1$ at a scale $M_0 = 1.6 \times 10^{11} M_\odot$ (note that this power-law behavior is expected to break for higher-masses haloes (Tacchella et al., 2018; Trenti et al., 2010), which however do not significantly affect the 21-cm signal during cosmic dawn). While this simplistic model is not expected to capture all the complexities of feedback in the first galaxies, it will serve to study the impact of feedback on the detectability of our models. We will explore more detailed feedback scenarios in future work.

---

[2] `https://github.com/andreimesinger/21cmFAST`

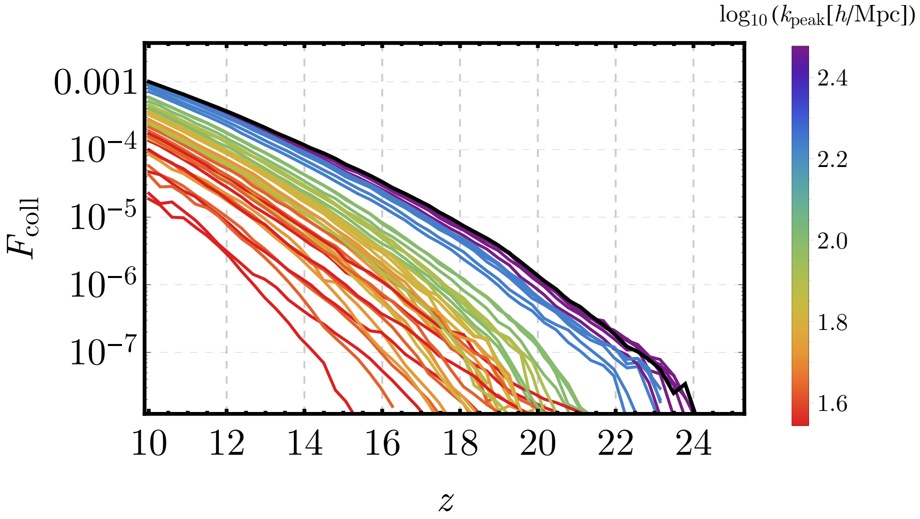

Figure 4.2.2. Collapsed fraction of baryons to star-forming haloes as a function of redshift z, for all our simulations. In all cases we assume that haloes above the atomic-cooling threshold can form stars, and consider no further feedback in this plot. Lines are colored by the wavenumber of their first peak $k_{\rm peak}$*, regardless of the height* $h_{\rm peak}$*, with CDM corresponding to the highest* $k_{\rm peak}$ *shown. The black line corresponds to CDM.*

We will conservatively assume that $\alpha = 0$ for all ETHOS models, as further feedback would only make them deviate more from CDM. For CDM, on the other hand, we will vary α in the range $[0-0.5]$, in order to estimate the impact of feedback, and whether the different ETHOS models can be distinguished from it. We note that our range of values of α is lower than typical of lower-z probes, such as galaxy luminosity functions, where $\alpha \approx 1$ (Gillet et al., 2020; Tacchella et al., 2018; Yung et al., 2018), as we expect feedback to be less important during cosmic dawn.

As our ETHOS HMFs are obtained exclusively from a zoom-in region within a larger simulation box (see Bohr et al. 2020), we need to apply a correction for the possible difference in mean density between the zoom region and the whole cosmological volume. To do so, we use an extended Press-Schechter formalism (Press & Schechter, 1974) in which we rescale the collapsed fractions as

$$F_{\rm coll}(z) \to F_{\rm coll}(z) \frac{\operatorname{erfc}\left[\dfrac{\delta_{\rm crit} - \delta_{\rm zoom}(z)}{\sqrt{2S(z)}}\right]}{\operatorname{erfc}\left[\dfrac{\delta_{\rm crit}}{\sqrt{2S(z)}}\right]} \tag{4.4}$$

given the overdensity $\delta_{\rm zoom}$ in the zoom-in region (measured in the simulations), where $\delta_{\rm crit} = 1.686$ is the critical density for collapse, and

$$S(z) = \sigma^2_{\rm cool}(z) - \sigma^2_{\rm zoom}(z) \tag{4.5}$$

is the variance on the cooling haloes, corrected by that in patches of the zoom-in region, $\sigma^2_{\rm zoom}$. We expand on how we compute the variances below. This procedure is exact for $\alpha = 0$, and we have confirmed that it reproduces the collapsed fraction in zoom-in simulations with average density for CDM. Moreover, the standard procedure used in `21cmFAST` is to modify $F_{\rm coll}$ in over/under-dense pixels via this same formula, so our re-scaling would be equivalent to changing the average density of the overall `21cmFAST` box to be $\delta_{\rm zoom}(z)$ and using the zoom-in overdense $F_{\rm coll}$. We note that this formula is not exact for $\alpha > 0$, which can affect the rescaling of our CDM+feedback results (though not our ETHOS ones, as those always have $\alpha = 0$).

The other ingredient modified in ETHOS models is the matter power spectrum, which changes the variance σ^2 of fluctuations on different scales. Since all ETHOS models we consider follow CDM at large scales the variance on the pixel size is not altered. Nevertheless, the variance on the scale at which atomic-cooling haloes form will change. We calculate it as

$$\sigma^2_{\rm cool}(z) = \int \frac{d^3k}{(2\pi)^3} P_m(k) |\mathscr{W}(kR_{\rm cool})|^2, \tag{4.6}$$

where $R_{\rm cool} = R_{\rm atom}(z)$ is the comoving radius of atomic-cooling haloes at each z, and $\mathscr{W}$ is a window function, which can have different functional forms, such as a (real-space) top-hat. Nevertheless, it has been shown that the HMFs of non-CDM models are better fit when using a smooth window function

$$\mathscr{W}(x) = \frac{1}{1 + (x/c)^\beta}, \tag{4.7}$$

with $c = 3.7$ and $\beta = 3.5$, as calibrated in Sameie et al. (2019) to fit the HMF of models with DAO, such as the ones we study here.

We note that we conservatively do not alter the reionization calculation from `21cmFAST`, as we are interested in the cosmic-dawn era only. We encourage the reader to see Lovell et al. (2018b, 2019); Bose et al. (2019) for the effect of ETHOS models on reionization and the Lyman-α forest.

4.3 Effect on the 21-cm Global Signal

The different histories of early structure formation of each of the ETHOS models will give rise to different 21-cm signals during cosmic dawn. Here we explore this observable, starting with the global signal—the average absorption or emission of 21-cm photons across the entire sky at each frequency or redshift. This signal has been targeted by several experiments (Price et al., 2018; Singh et al., 2018; Philip et al., 2019; Voytek et al., 2014; DiLullo et al., 2020), including a first detection claimed by the EDGES collaboration (Bowman et al., 2018).

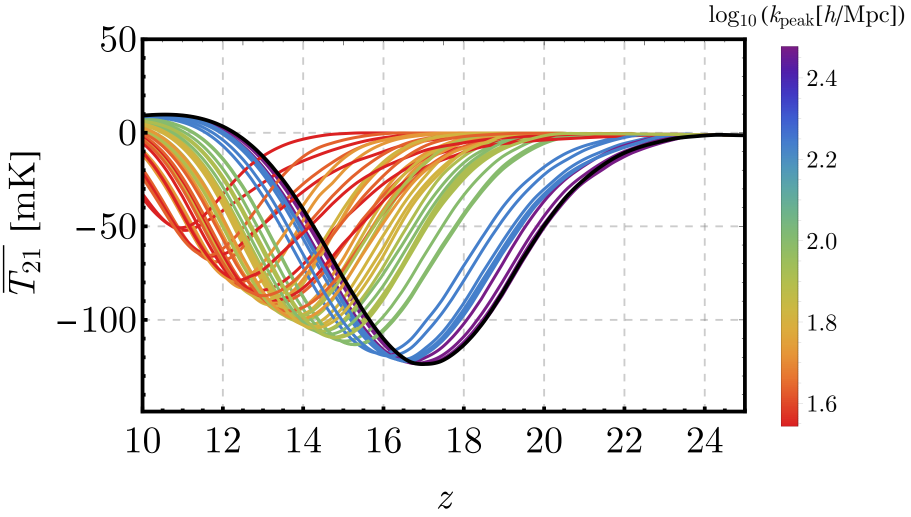

Figure 4.3.3. Global signal as a function of redshift for all our ETHOS simulations. As in Fig. 4.2.2, the color scale indicates the scale $k_{\rm peak}$ of the first peak, and black corresponds to CDM. All models show the same landmarks of evolution, explained in the main text, although the location and depth of the peaks change between models.

4.3.1 The observable

We define the usual 21-cm brightness temperature as,

$$T_{21}(\mathbf{x}) = 38\,\mathrm{mK}\left(1-\frac{T_\gamma}{T_S}\right)\left(\frac{1+z}{20}\right)^{1/2} \times \left(\frac{\partial_r v_r}{H}\right)^{-1} x_{\rm HI}(1+\delta_b), \tag{4.8}$$

where $\partial_r v_r$ is the radial velocity gradient, $H(z)$ is the Hubble expansion rate, δ_b is the baryonic overdensity, and T_γ and T_S are the photon (CMB) and spin temperatures, respectively. During the cosmic-dawn era that we are interested in the hydrogen neutral fraction $x_{\rm HI} \approx 1$. For a thorough review of the physics of the 21-cm line, we refer the reader to Pritchard & Loeb (2012); Furlanetto et al. (2006). The 21-cm temperature will be computed at each point using the `21cmvFAST` simulations outlined above, and the global signal $\overline{T_{21}}$ is obtained by simply averaging the entire box output at each redshift.

Throughout this work we will use a single set of initial conditions for all the simulations, to ease comparison, generated with the *Planck* 2018 best-fit cosmological parameters (Planck Collaboration et al., 2020). Moreover, we will fix the astrophysical parameters to be the same as in Muñoz et al. (2020). Our simulation boxes have 600 Mpc comoving in size, and 3 Mpc resolution, and are ran from $z = 35$ to $z = 10$, to avoid the bulk of reionization.

Under these conditions, we show the output of all of our ETHOS models, and CDM, in Fig. 4.3.3. Their overall evolution can be summarized as follows. First, during

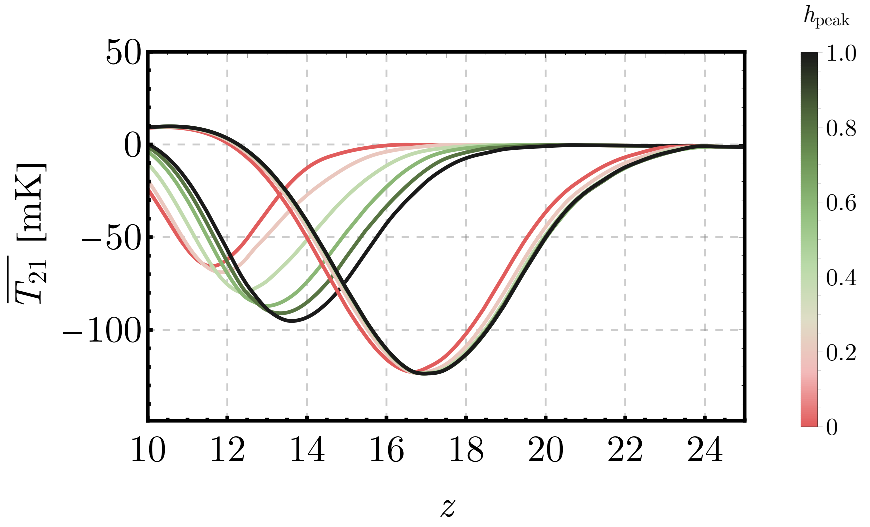

Figure 4.3.4. Same as Fig. 4.3.3 but for only models with $k_{\rm peak} = 43\,h/\mathrm{Mpc}$ (left lines) and 300 h/Mpc (right lines), where the color indicates the value of $h_{\rm peak}$. Models closer to WDM ($h_{\rm peak} \sim 0$) have less structure formation, and thus a delayed 21-cm absorption signal.

the onset of the Lyman-α coupling era (LCE; at $z \sim 22$ for CDM) the GS deviates from zero due to the UV photons from the first stars, which produce Wouthuysen-Field coupling (Wouthuysen, 1952; Field, 1959; Hirata, 2006). Second, the transition from the LCE to the epoch of heating (EoH; at $z \sim 17$ for CDM) the signal starts growing due to the X-ray heating of the neutral hydrogen (Pritchard & Furlanetto, 2007; Pacucci et al., 2014). Finally, the EoH gives way to the epoch of reionization (EoR; at $z \sim 12$ for CDM) where the IGM is fully heated and the signal is reduced as hydrogen slowly becomes ionized (Barkana & Loeb, 2001; Pritchard & Loeb, 2008).

While all the models shown in Fig. 4.3.3 exhibit a similar overall evolution, models with more suppressed power are delayed with respect to CDM. Furthermore, the entire shape of the GS depends on the details of the initial power spectrum, as models with additional power at large k produce a more quickly evolving 21-cm global signal at high z. To illustrate this point, we show in Fig. 4.3.4 the GS for a subset of models with $k_{\rm peak} = 43$ and 300 h/Mpc, for different values of $h_{\rm peak}$. Stronger DAOs (higher $h_{\rm peak}$) produce less suppression in the HMF, and thus an earlier 21-cm evolution. This effect is more apparent for low $k_{\rm peak}$, as the haloes observed probe a broader range of k in the matter power spectrum. As we will explore below, this will allow us to distinguish different ETHOS models from one another.

Finally, we are also interested in distinguishing ETHOS models from CDM in the presence of feedback. As described in the previous section, we will phenomenologically account for further possible sources of feedback by varying the parameter α in Eq. (4.3), which suppresses the amount of star formation in a mass-dependent way for each halo. We show how the 21-cm GS varies with increasing α in Fig. 4.3.5, which trivially delays the evolution of the GS. Note that this delay is relatively smooth, as opposed to the sharper delay shown in Fig. 4.3.4, especially for $h_{\rm peak} = 0$ (WDM) models. This is to be expected, as this power-law-like astrophysical feedback does not cut off all haloes below some scale, whereas the ETHOS models approximately do. This will help us to differentiate ETHOS models from CDM+feedback.

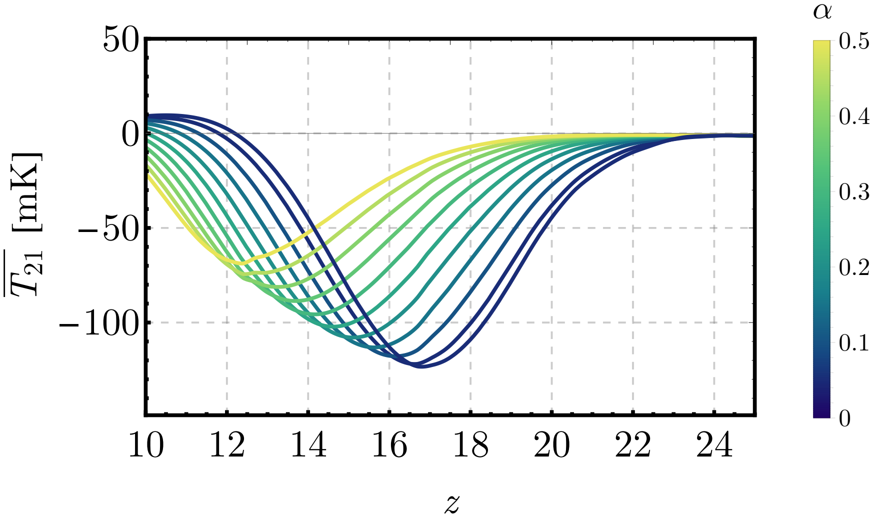

Figure 4.3.5. Same as Fig. 4.3.3 but for CDM only, where we vary the feedback parameter α *from Eq.* (4.3). *Larger* α *corresponds to stronger feedback, and thus to a delayed 21-cm absorption signal.*

4.3.2 Detectability

While we have shown that different ETHOS models show very different 21-cm signals as a function of redshift, given their different amounts of structure formation, we have not shown whether this effect can be mimicked by feedback, and if different ETHOS models can be distinguished from each other (as for instance models with stronger DAOs and a lower $k_{\rm peak}$ can produce similar amounts of suppression as WDM with higher $k_{\rm peak}$, see Fig. 4.3.4). We now perform a simple analysis to find how differentiable ETHOS models are from each other and from CDM, even when including potential feedback.

A realistic analysis should simultaneously fit for the cosmological 21-cm signal as well as the Galactic, extra-Galactic, and atmospheric foregrounds that swamp it. This is costly to perform for all of our simulations, so instead we will define the difference

$$d_{1,2}(z) = \overline{T_{21}}^{(1)}(z) - \overline{T_{21}}^{(2)}(z) \tag{4.9}$$

between two GS models ($T_{21}^{(1)}$ and $T_{21}^{(2)}$, respectively), and simply compute the χ^2

statistic

$$\chi^2 = \sum_{i,j} d_{1,2}(z_i) C_{ij}^{-1} d_{1,2}(z_j), \tag{4.10}$$

as a metric of how different these two models are *in theory*. Here the indices i, j run over redshifts (or frequencies), and C is the covariance matrix, where for our first analysis we can neglect the cosmic-variance component of C (Muñoz & Cyr-Racine, 2021), and take $C_{ij} = \sigma_i^2 \delta_{ij}$, with an instrumental noise

$$\sigma_i = \frac{T_{\rm sky}}{\sqrt{B t_{\rm obs}}}, \tag{4.11}$$

determined by the observation time $t_{\rm obs} = 1$ year, bandwidth $B = 0.4$ MHz, and a sky temperature $T_{\rm sky}(\nu) = 1570 \times (\nu/\nu_0)^{-2.5}$, anchored at $\nu_0 = 72$ MHz, all chosen to closely match EDGES (Bowman et al., 2018). Moreover, in this analysis we will consider the frequency range $\nu = 50 - 110$ MHz, covering $z = 12 - 27$, which covers the entire cosmic-dawn range of interest, and cuts off the beginning of reionization.

Before showing our results, let us emphasize that the χ^2 obtained with Eq. (4.10) should be interpreted with caution. This is for two main reasons. First, we are not including any foreground marginalization, which can make two models appear more similar to each other, as well as diminish the overall significance of a prospective detection. Second, we are keeping all astrophysical parameters fixed, as varying them would dramatically increase the dimensionality of the problem, making it prohibitively expensive. We will vary only one parameter, α, which encapsulates the effect of feedback during cosmic dawn. As a consequence, our reported χ^2 values in this section ought to be interpreted as a theoretical best-case scenario of the difference between models, aimed to guide future detailed studies, whereas the specific values of χ^2 will dampen when other effects are included.

We start by studying the differences between ETHOS and WDM models in the 21-cm GS. In order to perform a meaningful comparison we will find the closest WDM model (with $h_{\rm peak} = 0$ but $k_{\rm peak} < \infty$) to each ETHOS one, and report the χ^2 difference between them. For this, we interpolate the GS from our finite sample of WDM simulations to obtain results for arbitrary values of $k_{\rm peak}$. We show the result of this analysis in Fig. 4.3.6. As expected, low values of $h_{\rm peak}$ are very similar to WDM, and in fact for $h_{\rm peak} \leq 0.2$ the difference between WDM and ETHOS is small. This difference grows for stronger DAOs, showing that the 21-cm signal has the potential to distinguish them from WDM. Note that, at fixed $k_{\rm peak}$, higher values of $h_{\rm peak}$ produce less suppression, and thus the closest WDM model has a larger free-streaming scale (defined as the value of $k_{\rm peak}$ for $h_{\rm peak} = 0$).

We now move to find how different each ETHOS model is from CDM with feedback. The analysis is similar to the WDM case, although now we interpolate between different values of $\alpha = 0 - 0.5$, which parametrizes the feedback strength. We report the value of α that makes CDM closest to each ETHOS simulation, as well as the χ^2 difference between them. The results of this analysis are summarized in Fig. 4.3.7. Larger values of $k_{\rm peak}$ in ETHOS correspond to more CDM-like behavior, and thus lower α. Interestingly, at fixed $k_{\rm peak}$ increasing the height $h_{\rm peak}$ of the DAOs requires lower α, as there is more structure formation (and thus it is more similar to CDM). The value of χ^2 between the two models grows for smaller $k_{\rm peak}$, as warmer DM produces a more marked—and

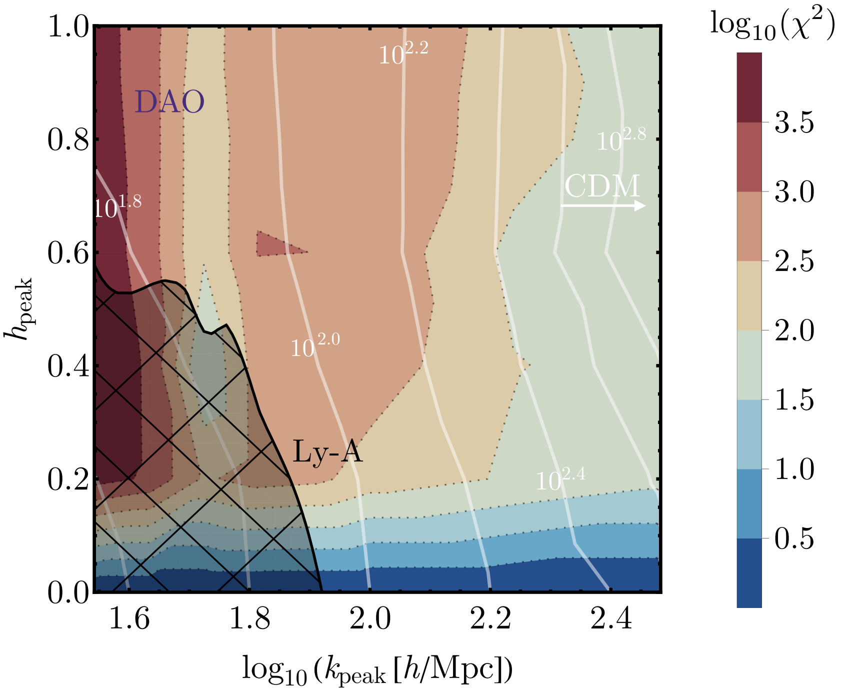

Figure 4.3.6. We show the comparison between each ETHOS model (as a function of its two effective parameters $k_{\rm peak}$ and $h_{\rm peak}$) and the closest WDM model (with different $k_{\rm peak}$ but $h_{\rm peak} = 0$), using the 21-cm global signal. The color scale indicates the value of the χ^2 difference between the two cases (which ought to be interpreted with care, see Eq. (4.10) *and the surrounding discussion), whereas the white lines denote the free-streaming scale for the closest WDM model (in units of* $\log_{10}[k_{\rm peak}/(h/{\rm Mpc})]$*, see Eq.* (4.1) *for a translation to a WDM mass), which grows with $h_{\rm peak}$, as expected. The black shaded region is ruled out by Lyman-α data (Bohr et al., 2020; Murgia et al., 2017; Archidiacono et al., 2019).*

rapid—suppression than the smooth feedback. Note that for $k_{\rm peak} \lesssim 10^{1.8}\,h/{\rm Mpc}$ the closest value of α saturates at 0.5, the maximum value we allow.

While in the comparison between ETHOS and WDM models (Fig. 4.3.6) the χ^2 difference reached small values in part of the parameter space ($\lesssim 10$ for $h_{\rm peak} \leq 0.2$), that is not the case when contrasting ETHOS and CDM+feedback. Even for large values of $k_{\rm peak}$ we find a significant ($\chi^2 \gtrsim 100$) deviation between ETHOS and the closest CDM+feedback model. These χ^2 values would be reduced once foreground and astrophysical-parameter marginalization are included, as argued above. Nevertheless, we expect that the relative size of these differences to hold, and thus that the ETHOS models that we explore are fairly distinguishable from CDM+feedback, due to the cutoff nature of ETHOS suppression versus the smooth power-law suppression of the feedback, given the phenomenological feedback model that we have implemented.

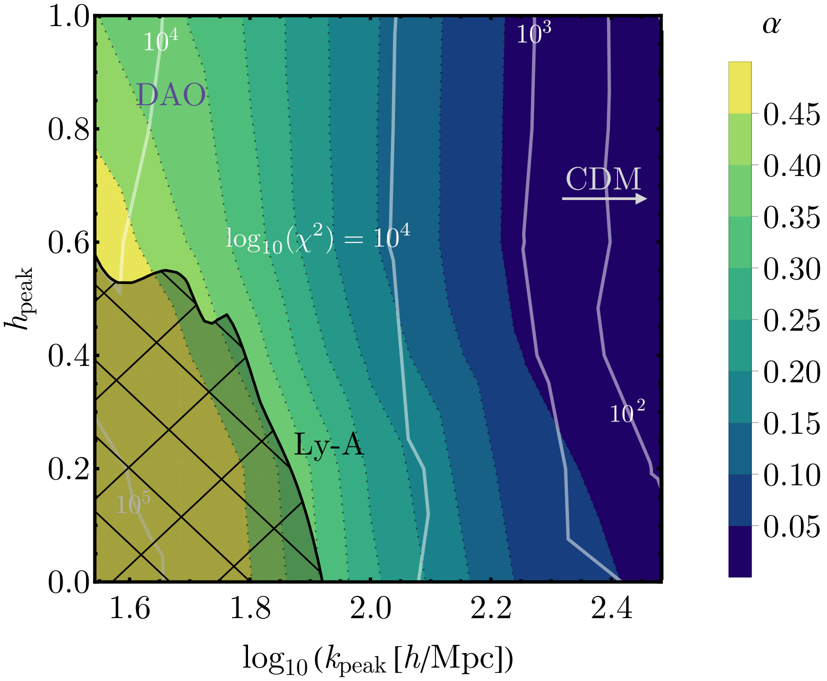

Figure 4.3.7. Same as Fig. 4.3.6 but for the closest CDM+feedback model, parametrized by α. Here the white lines indicate the χ^2 difference, whereas the color scale follows the α of the closest CDM+feedback model. The difference between CDM+feedback and ETHOS models grows with lower $k_{\rm peak}$.

4.4 Effect on the 21-cm Fluctuations

In addition to the 21-cm GS, changing the HMF has a profound impact of the 21-cm fluctuations, which are expected to be measured soon by 21-cm interferometers (van Haarlem et al., 2013; Eastwood et al., 2019; Beardsley et al., 2016; Koopmans et al., 2015; DeBoer et al., 2017). Let us now turn our attention to them.

4.4.1 The observable

We begin describing the 21-cm fluctuations and how we calculate them. We use the same `21cmvFAST` simulation boxes from above, where we decompose the 21-cm temperature at each point as

$$T_{21}(\mathbf{x}) = \overline{T_{21}} + \delta T_{21}(\mathbf{x}), \tag{4.12}$$

and calculate the Fourier-space two-point function of the 21-cm fluctuation δT_{21}. This two-point function is the 21-cm power spectrum P_{21}. For convenience we will employ

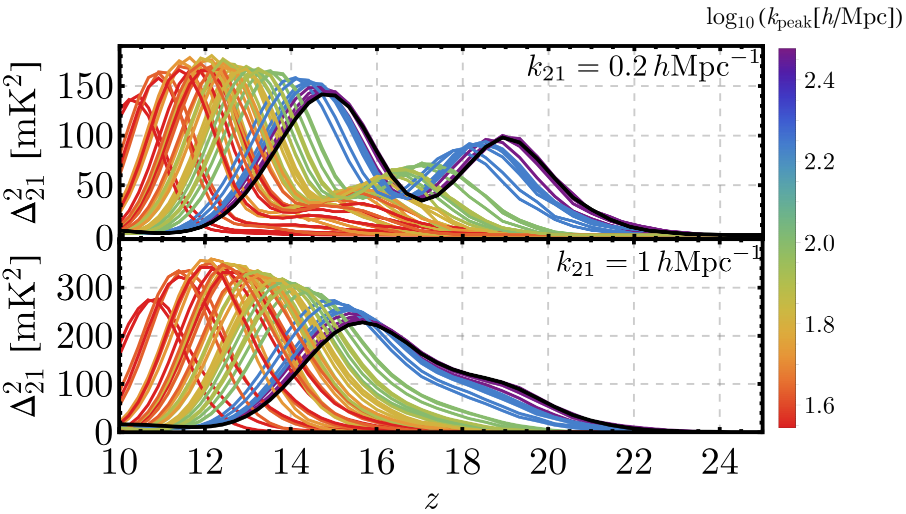

Figure 4.4.8. Amplitude of the 21-cm fluctuations as a function of redshift at two wavenumbers $k_{21} = 0.2h\,{\rm Mpc}^{-1}$ *(top) and* $1h\,{\rm Mpc}^{-1}$ *(bottom). As in previous figures, the color encodes the wavenumber of the first peak* $k_{\rm peak}$.

the amplitude of 21-cm fluctuations, defined as

$$\Delta^2_{21}(k_{21}) = \frac{k^3_{21}}{2\pi^2} P_{21}(k_{21}), \tag{4.13}$$

and refer to it as the 21-cm power spectrum (PS) unless confusion can arise. In order to notationally differentiate the wavenumbers of 21-cm fluctuations from those of matter fluctuations, we refer to the former as k_{21}. Interferometers such as the hydrogen epoch-of-reionization array (HERA) will probe the range $k_{21} \sim 0.1 - 1\,h/{\rm Mpc}$, as for lower wavenumbers foregrounds dominate, whereas for higher ones thermal noise does (DeBoer et al., 2017).

To build intuition, we show in Fig. 4.4.8 the 21-cm PS at two wavenumbers, $k_{21} = 0.2$ and $1\ h/{\rm Mpc}$, for all our ETHOS simulations. These wavenumbers are chosen to represent both large- and small-scale 21-cm fluctuations that are observable by the current generation of experiments. The origin of 21-cm fluctuations is different during each of the eras described above, so let us begin by describing the overall features of these curves.

We begin at early times, during the LCE ($z \sim 17 - 22$ for CDM), where fluctuations grow at all scales k_{21}. That is because the UV photons are emitted from anisotropically distributed galaxies, and as they produce more WF coupling these fluctuations grow. The large-scale (small k_{21}, top panel of Fig. 4.4.8) fluctuations decrease in size during the transition between the LCE and the EoH ($z \sim 17$ for CDM), as the effect of X-ray and UV photons cancel out (Muñoz, 2019b), whereas at small scales (large k_{21}, bottom panel of Fig. 4.4.8) there is no such cancellation. Finally, the 21-cm fluctuations increase again during the EoH, until they nearly vanish by the time the gas is fully heated ($z \sim 12$

for CDM). There will be a third peak at lower redshifts, due to reionization, which we do not consider, as we do not include lower redshifts in our analyses.

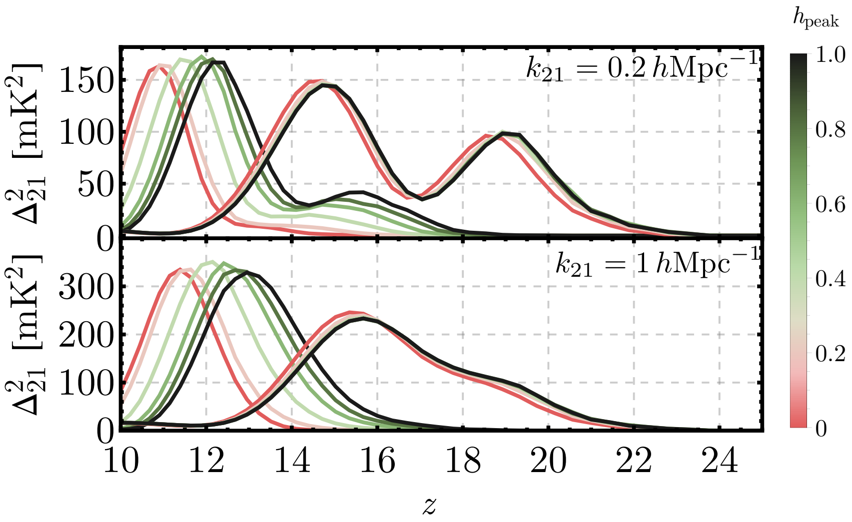

Figure 4.4.9. Same as Fig. 4.4.8, but for the subset of models with $k_{\rm peak} = 43$ *(left lines) and* $300\,h/$*Mpc (right lines), with* $h_{\rm peak}$ *denoted by the line color.*

As was the case for the GS, ETHOS models show delayed structure formation, and thus the 21-cm PS curves shift to lower redshifts. Nevertheless, the 21-cm fluctuations provide us with angular information, in the form of different k_{21}, which will allow us to better differentiate between models. This is apparent, for instance, in Fig. 4.4.9. There we show the 21-cm PS for models with two $k_{\rm peak}$, as a function of $h_{\rm peak}$. The $h_{\rm peak} = 0$ cases tend to form structure later than their higher-$h_{\rm peak}$ counterparts, as argued above. Nevertheless, the shift in the high- and low-k_{21} fluctuations is different. For instance, the WDM ($h_{\rm peak} = 0$) and full-DAO ($h_{\rm peak} = 1$) curves with $k_{\rm peak} = 43\,h/$Mpc in the top panel of Fig. 4.4.9 have very different shapes, showing that the effect of DAOs is not just a shift, and the entire cosmic history of the 21-cm line can be used to differentiate between models.

Finally, as we did before, we include CDM with feedback by varying the parameter α in Eq. (4.3). We show the resulting power spectra in Fig. 4.4.10, where as before larger α (stronger feedback) delay the onset of all the 21-cm transitions. Interestingly, however, the 21-cm power is not just delayed, but its shape as a function of redshift also changes, owing to the impact that haloes of different masses have on the 21-cm line as a function of redshift (Muñoz et al., 2020).

4.4.2 The Noise

There are different ongoing and proposed 21-cm interferometers targeting the cosmic-dawn era (van Haarlem et al., 2013; Eastwood et al., 2019; Beardsley et al., 2016;

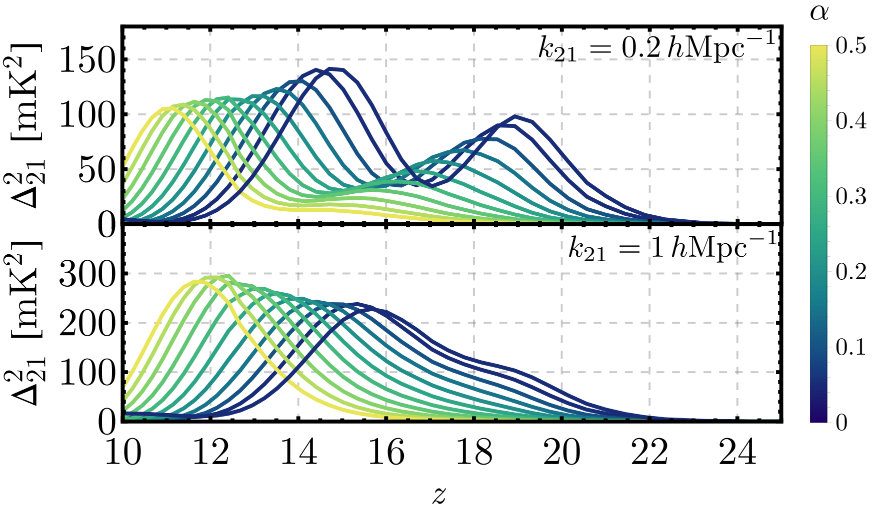

Figure 4.4.10. Same as Fig. 4.4.8, but for CDM with varying amounts of feedback, parametrized through α.

Koopmans et al., 2015; DeBoer et al., 2017). For concreteness, here we will focus on HERA (DeBoer et al., 2017), and study how well it will be able to detect the fluctuations from all these models, as well as to distinguish them from one another and from CDM. We will perform a realistic χ^2 analysis here (as opposed to that in the previous section), using the noise expected of HERA. We assume three years (540 days) of HERA data, and use the standard package `21cmSense`[3] to forecast the noise (Pober et al., 2014; Pober et al., 2013). We discard all wavenumbers within the foreground wedge (Liu et al., 2009; Morales et al., 2012; Datta et al., 2010; Parsons et al., 2012), whose extent we vary from an optimistic case, where the horizon limit is given by the experiment resolution, to a moderate and a pessimistic case, which include a supra-horizon buffer, following Muñoz (2019a) (see Appendix 4.6.1 for more details).

A subtlety that we have to address is that, while the telescope (thermal) noise is the same for all of our simulations, they each have a different cosmic-variance (CV) noise, given their different fiducial power spectra. This CV is important for low wavenumbers ($k_{21} \sim 0.1\,h$Mpc^{-1}), where thermal noise is small. Instead of running `21cmSense` for each of our simulations individually, which is computationally slow, we devise a way of including CV for any arbitrary 21-cm PS quickly but exactly. The full noise of the 21-cm PS can be expressed as a sum of the thermal (th) and CV components, where the former is independent of the 21-cm model assumed, and the latter can be described as $\sigma_{\rm CV}(\Delta^2_{21}) = a_{21} \times \Delta^2_{21}$ for some a_{21} that depends on k_{21} and z, and varies with the experimental setup, but not with Δ^2_{21}. Thus, we calibrate this a_{21} by using `21cmSense`, and find the full error as

$$\sigma_{\rm full}(\Delta^2_{21}) = \sigma_{\rm th} + a_{21}\,\Delta^2_{21}, \tag{4.14}$$

for each 21-cm PS Δ^2_{21}, where we have suppressed the dependence on k_{21} and z of all

---

[3] `https://github.com/jpober/21cmSense`

terms in that equation. We have confirmed that this expression exactly recovers the full noise when using different input 21-cm power spectra in `21cmSense`.

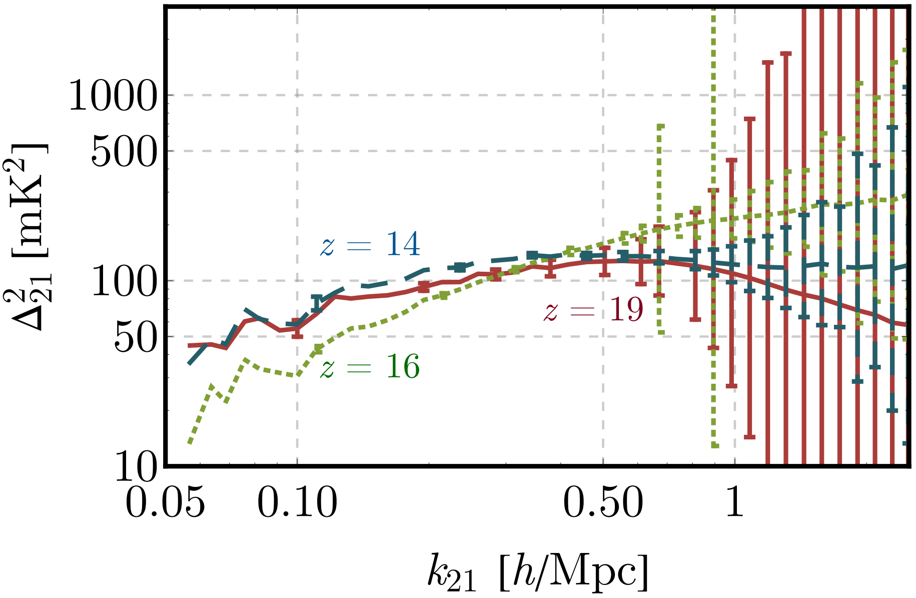

Figure 4.4.11. Amplitude of 21-cm fluctuations as a function of wavenumber for our CDM model (with $\alpha = 0$), as well as the forecasted noise for 540 days of HERA data, assuming moderate foregrounds. We show the results at three redshifts, roughly corresponding to the peak of the LCE ($z = 19$), the transition to the EoH ($z = 16$), and the peak of the EoH ($z = 14$) for this model. Wavenumbers without an errorbar cannot be measured at any precision.

In order to perform our analysis we divide the frequency range $\nu = 50 - 120$ MHz in bins that are 4 MHz in size. These are wider than for the GS, as we ought to average over more cosmic volume to bring the noise per k_{21} mode down at each z. We show the expected noise for our CDM model, under moderate foregrounds, in Fig. 4.4.11. We will analyze wavenumbers in the range $k_{21} = 0.05 - 2.5\,h/$Mpc, though the majority of modes do not have a measurement, as clear in Fig. 4.4.11, due to the foreground wedge. For low k_{21} only a handful of modes can be observed, although they can reach small errors as they are observed many times. For larger k_{21} (smaller scales), however, the situation is reversed, and more modes with $k_{21} \gtrsim 0.5\,h/$Mpc can be observed, while they each have large noise.

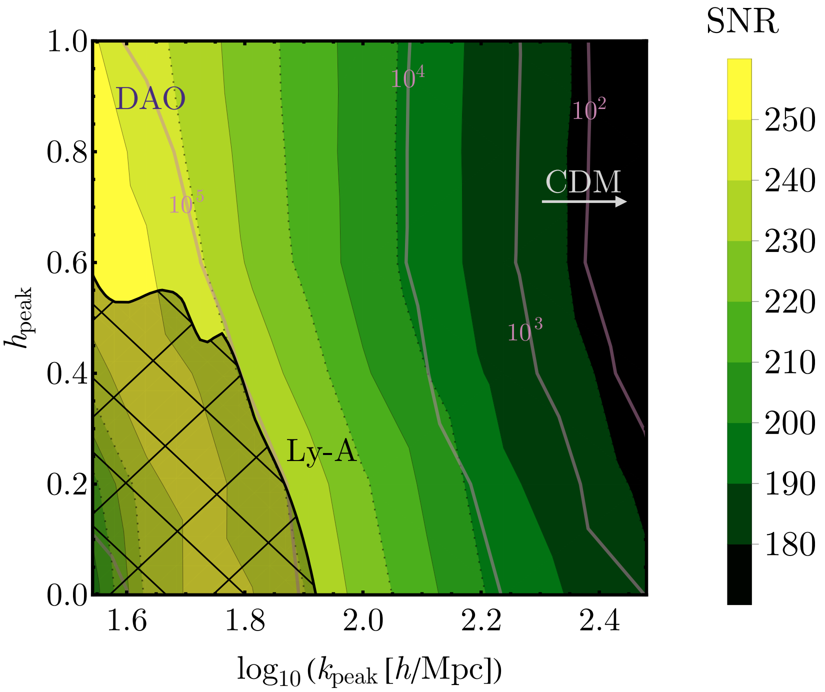

Figure 4.4.12. Forecasted signal-to-noise ratio (SNR) of the 21-cm PS for different ETHOS models, in color map. In all cases we assume 540 days of HERA data, and moderate foregrounds. The thin purple lines follow the contours of constant χ^2 difference between each ETHOS model and CDM (with no feedback), which grows towards the left of the plot.

4.4.3 Detectability

We will use two metrics to study how detectable—and differentiable from each other—our ETHOS models are. The first is the signal-to-noise ratio (SNR), and the second is the χ^2 statistic. In all cases we will assume a diagonal covariance matrix, ignoring correlations between different k_{21} and z bins, for simplicity.

We begin by calculating the SNR for each of our models, computed through

$$\mathrm{SNR}^2 = \sum_{i_k, i_z} \frac{\Delta^2_{21}(k_{21}, z)}{\sigma^2_{\rm full}(k_{21}, z)}, \tag{4.15}$$

where the sum runs over all wavenumber i_k and redshift i_z bins. We show the SNR for all our models, assuming moderate foregrounds, in Fig. 4.4.12. We find SNR $\approx 150-250$, varying smoothly as a function of the ETHOS parameters. Interestingly, more-suppressed ETHOS models haver higher SNR than their CDM-like counterparts. The reason is that a stronger suppression of power delays structure formation, and moves all the 21-cm landmarks to lower z, where the noise is smaller (as $T_{\rm sky}$ sharply rises at lower frequencies—or high redshifts). This trend is reversed for ETHOS models with $k_{\rm peak} \lesssim 10^{1.6}\,h/{\rm Mpc}$, however, as their cosmic-dawn evolution is late enough that it is not completed by $z = 10$, when our simulations end. Nevertheless, the models for which this is true are already in tension with Lyman-α observations (Bohr et al., 2020), as clear in Fig. 4.4.12.

As all our ETHOS models are detectable at high SNR, we now perform a χ^2 test to distinguish between them, similar to the previous section. Given the difference $\Delta^2_{\rm diff}$ between the 21-cm power spectra of two models, we define their χ^2 to be

$$\chi^2 = \sum_{i_k,i_z} \frac{\Delta^2_{\rm diff}(k_{21},z)}{\sigma^2_{\rm full}(k_{21},z)}, \tag{4.16}$$

where the noise in the denominator is evaluated for the first of the two models (which will always be the one plotted). While this χ^2 for the 21-cm PS shares some of the same caveats as that of the GS (as we are not simultaneously varying astrophysical parameters due to the computational cost), it is fundamentally more robust. The reason for that is twofold. First, here we do not have to subtract foregrounds, as we only consider data outside of the wedge, which is expected to be foreground clean. Second, here we are taking realistic forecasted noises for HERA, as opposed to using the “ideal" radiometer equation for the GS, which results in lower overall values of the χ^2 for the PS than for the GS, thought these can be trusted more. Nevertheless, any potential systematics are not included in our `21cmSense` noise, which could change the forecasted results.

Looking at Fig. 4.4.12 once more, we see that essentially all ETHOS models are very different from the vanilla CDM scenario, as the χ^2 difference between them is always larger than 10, and grows dramatically as $k_{\rm peak}$ decreases, especially below $10^{2.4}\,h/{\rm Mpc}$. However, as argued above, some of this difference can be absorbed by a difference in the astrophysics. Moreover, we want to know if ETHOS models can be distinguished from WDM given a fiducial 21-cm observation. We now tackle these two questions.

We begin, as in the previous section, by comparing ETHOS models with DAOs against their closest WDM counterpart. We show the summary of this analysis in Fig. 4.4.13. As before, we find that at fixed $k_{\rm peak}$ models with strong DAOs (large $h_{\rm peak}$) suppress structure less. Now, however, the χ^2 difference between models is slightly smaller, and in fact it is below 10 for $h_{\rm peak} < 0.2$, making those barely distinguishable from WDM. Moreover, all models with $k_{\rm peak} > 10^{2.2}\,h/{\rm Mpc}$ have differences $\chi^2 \lesssim 30$ with respect to their closest WDM counterpart, as such small scales chiefly affect high redshifts where the PS noise is too high to distinguish them. On the opposite side, the difference between models grows for larger values of $h_{\rm peak}$, especially at low $k_{\rm peak}$. For instance the larger-scale DAOs, with $h_{\rm peak} \gtrsim 0.4$ and $k_{\rm peak} \lesssim 100\,h/{\rm Mpc}$, give rise to large $\chi^2 \sim \mathcal{O}(100)$ differences, and thus could be promptly distinguished from WDM. This shows the promise of 21-cm PS measurements to detect and characterize DAOs.

Additionally, we study how well HERA could distinguish ETHOS models from CDM+feedback. A summary of our findings is in Fig. 4.4.14. As before, ETHOS models with more suppression (lower $k_{\rm peak}$) are matched to CDM models with stronger feedback (larger α). However, here the low-$k_{\rm peak}$ range can be better distinguished from CDM+feedback than when using the GS, given the additional information from different wavenumbers. For the same reason, the best-fit values of the feedback-strength α for each ETHOS model are slightly different for the 21-cm PS than for the GS. As

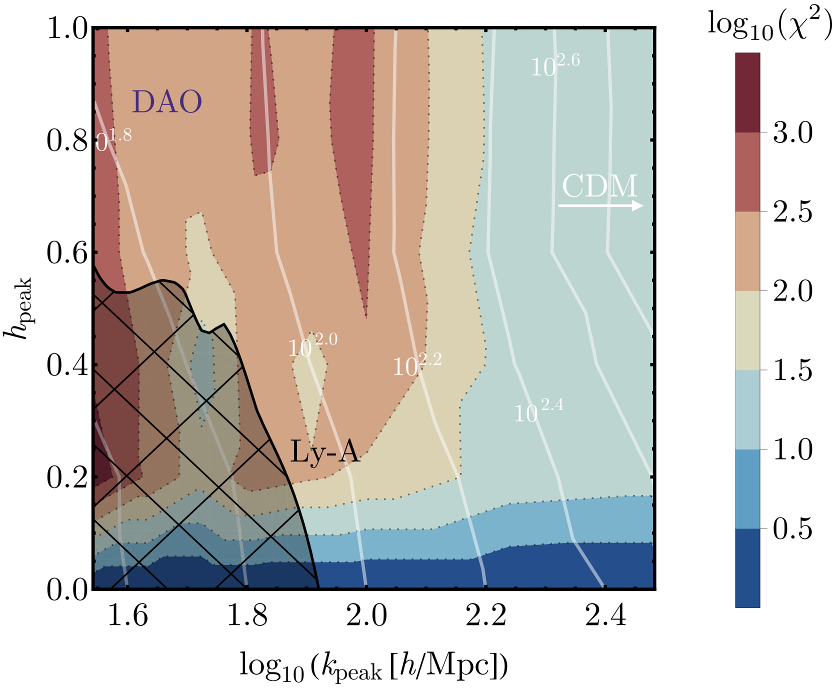

Figure 4.4.13. Different between each ETHOS model and the closest WDM case, as in Fig. 4.3.6, but for the 21-cm fluctuations, assuming 540 days of HERA data and moderate foregrounds. The χ^2 differences reported here (as white contours) are more robust to marginalization than those in Fig. 4.3.6.

was the case in Fig. 4.4.13, the high-$k_{\rm peak}$ part of the parameter space is more difficult to probe with the 21-cm PS, as those models show their most marked suppression at high redshifts, where the noise is large. Nevertheless, we find that ETHOS models with $k_{\rm peak} \lesssim 10^{2.3}\,h/{\rm Mpc}$ give rise to a χ^2 difference larger than 100, showing that HERA has the potential to tell ETHOS apart from CDM+feedback, given our model assumptions.

Throughout this section we have shown results assuming moderate foregrounds, where the vast majority of 21-cm modes observed by HERA would be within the foreground wedge, and thus unusable for our analysis. The extent of the wedge is, as of yet, uncertain at the redshifts we consider, so we have re-done our analyses assuming two other different foreground options, an optimistic one and a pessimistic one. We show the results in Appendix 4.6.1, and simply summarize them here. We find that pessimistic foregrounds reduce the SNR of a prospective 21-cm PS detection by roughly 10% for all ETHOS models, as well as CDM, whereas the optimistic-foreground assumption increases the SNR by roughly a factor of 2. We additionally find that the χ^2 comparisons follow a similar trend as in the moderate-foreground case considered in the main text, though a factor of ~ 5 worse (better) for pessimistic (optimistic) foregrounds. This would change the specific cut of the ETHOS parameter space that is distinguishable

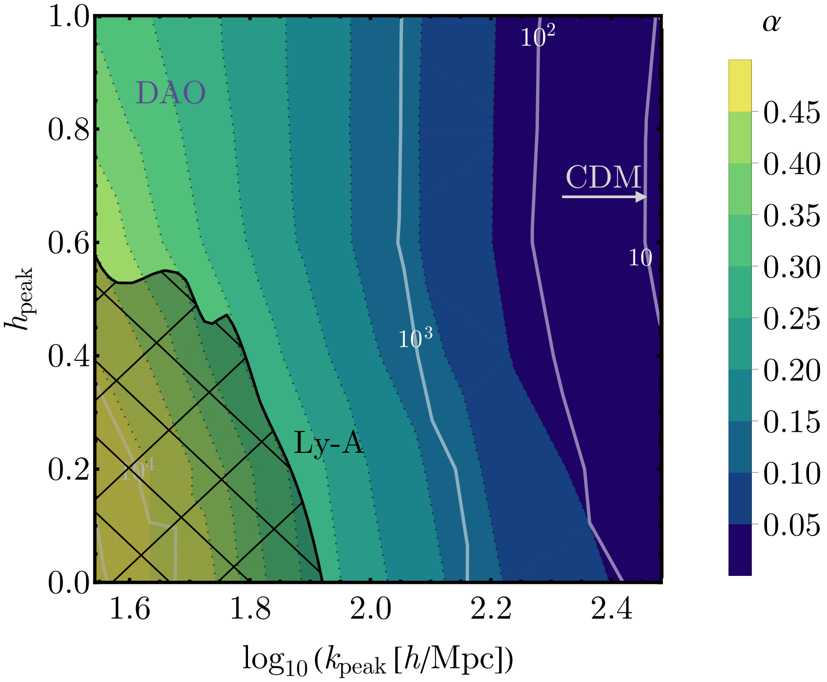

Figure 4.4.14. Same as Fig. 4.4.13 but comparing each ETHOS model to the closest CDM including feedback, whose strength is parametrized through α.

from CDM+feedback or WDM, but would not alter our main conclusions.

4.5 Discussion and Conclusions

In this work we have carried out an exploratory study of how upcoming measurements of the 21-cm line of hydrogen during cosmic dawn can determine the nature of the dark sector, through the small-scale behavior of DM. For that, we have followed the ETHOS paradigm, which translates the microphysical degrees of freedom of the DM and DR interactions into two key variables: the location $k_{\rm peak}$ and amplitude $h_{\rm peak}$ of the first DAO peak. We carried out N-body simulations of each ETHOS model to find their halo mass function down to the atomic-cooling threshold, and used those as input of semi-numeric `21cmvFAST` simulations to find the evolution of the 21-cm signal from the formation of the first stars to reionization. We then studied the prospects to detect, and distinguish, ETHOS models with upcoming measurements of the 21-cm global signal and fluctuations.

Our results can be summarized as follows. All ETHOS models with a suppression scale $k_{\rm peak} \lesssim 10^{2.5}\,h/{\rm Mpc}$ can be distinguished from CDM by both the 21-cm GS and the PS, even when varying the strength of the feedback processes in CDM. More interestingly, in the case that a prospective 21-cm detection shows a lack of power at high k, we have shown that ETHOS models with $h_{\rm peak} \gtrsim 0.4$ can be differentiated from WDM. That is because the cutoff in WDM produces a more sudden turn-on of the 21-cm signal than ETHOS models with strong DAOs, which exhibit a bump in power at smaller scales. Moreover, even models with DAOs can be distinguished from our feedback model, as this is expected to only suppress stellar formation in a smooth manner, rather than the sharper cut of non-CDM models.

Ours is the first study of the evolution of the 21-cm signal across cosmic dawn including DAOs of different heights and locations. As such, we have taken some simplifying assumptions to timely explore the large ETHOS parameter space. First, we have not considered small-mass molecular-cooling haloes, as resolving those requires finer-resolution N-body simulations. Nevertheless, as those haloes are formed out of smaller-scale fluctuations deviations from the standard CDM paradigm will be more apparent, and our analysis is, therefore, conservative. Second, we have only varied one astrophysical parameter (the strength of the stellar feedback in CDM), instead of freely allowing all possible parameters in `21cmvFAST` to vary. Last, in our global-signal forecasts we have ignored foreground marginalization. These simplifying assumptions will be relaxed in subsequent work. Throughout this paper we have assumed some fiducial observation time of 1000 hours for a global-signal experiment, and 4320 hours for a 21-cm fluctuation experiment. These were chosen for convenience only, and our results can be trivially rescaled for different observation times $t_{\rm obs}$. Despite these caveats, this work is a proof-of-concept that data of the 21-cm line of hydrogen at high redshifts ($z \approx 10-25$) can readily distinguish different ETHOS models from the standard CDM, as well as from each other, probing a large swath of parameter space that is currently open.

In summary, we have shown that the cosmic-dawn era holds a trove of information about the small-scale behavior of matter fluctuations. A detection of the 21-cm signal will, therefore, open the window to understanding the nature of DM in a regime currently unprobed, shedding light onto the nature of the dark sector.

Acknowledgements: It is our pleasure to thank Torsten Bringmann and Christoph Pfrommer for comments on a previous version of this draft. JBM was supported by NSF grant AST-1813694 at Harvard and the Clay Fellowship at the Smithsonian Astrophysical Observatory. SB and JZ were funded by a Grant of Excellence from the Icelandic Research Fund (grant number 173929). MV acknowledges support through NASA ATP grants 16-ATP16-0167, 19-ATP19-0019, 19-ATP19-0020, 19-ATP19-0167, and NSF grants AST-1814053, AST-1814259, AST-1909831 and AST-2007355. All simulations were run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University.

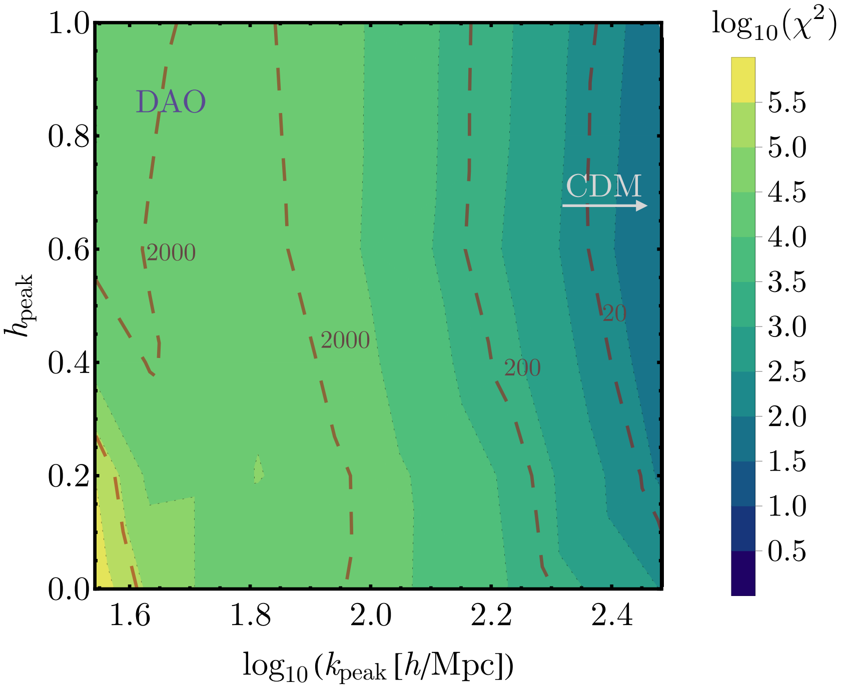

Figure 4.6.15. Difference in χ^2 between each ETHOS model and the closest CDM+feedback, assuming pessimistic (dashed brown contours) and optimistic foregrounds (color map). We do not show the best-fit α here, since they are visually very similar to Fig. 4.4.14.

4.6 Appendix

4.6.1 Foregrounds in the 21-cm Power Spectrum

In this appendix we describe alternatives for the extent of the foreground wedge, which determines which wavenumbers can be measured by the 21-cm power spectrum, and to which precision. We take a simple model of the foreground wedge, where wavenumbers along the line of sight ($k_{||}$) with

$$k_{||} \leq a(z) + b(z) k_{\perp}, \tag{4.17}$$

are considered to be contaminated by foregrounds, and are thus unusable for our DM studies. The two parameters a and b determine the extent of the wedge (see Pober et al. 2013; Pober et al. 2014 for details and its the implementation in `21cmSense`) as a function of the perpendicular wavenumber $k_{\perp}$, where $b(z)$ determines the extent of the horizon, and $a(z)$ accounts for a supra-horizon buffer where foregrounds may leak out (Orosz et al., 2019). We take three three assumptions for the foreground wedge, following Muñoz (2019a). In the main text we assumed moderate foregrounds, which is

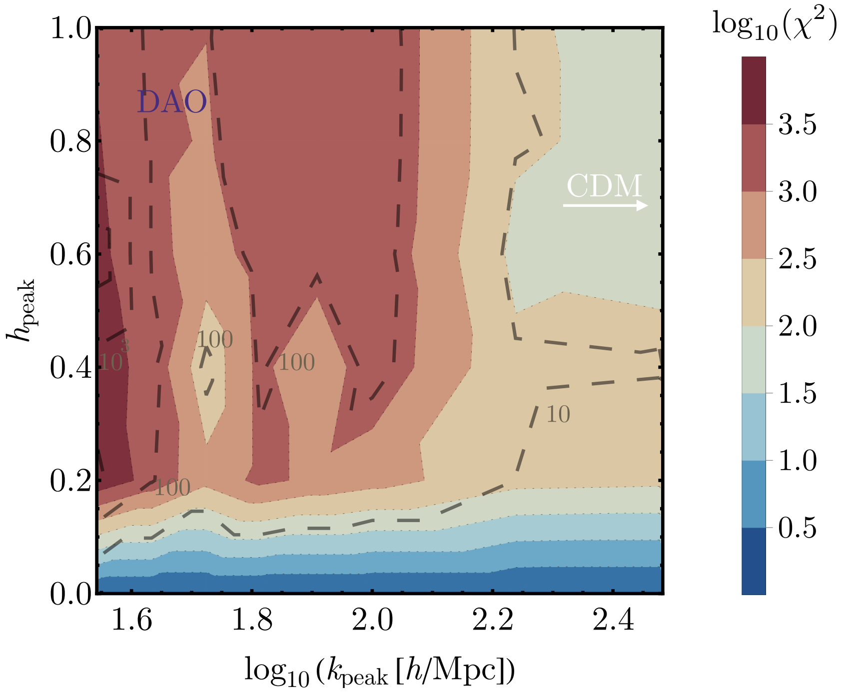

Figure 4.6.16. Same as Fig. 4.6.15 but comparing against the closest WDM case to each ETHOS model.

our best guess for the extent of the wedge. Here, instead, we explore what the results would be if foregrounds were more optimistic, where b is given by the primary beam and we take no buffer ($a = 0$), and a more pessimistic case where $a = 0.1\,h/$ Mpc (instead of half of that in the moderate case).

We show our results for these two foreground assumptions in Figs. 4.6.15 and 4.6.16. The first of these Figures shows the detectability of ETHOS models against CDM and feedback. We find that for the pessimistic-foregrounds case the expected χ^2 is only a factor of ~ 2 worse than for the moderate case. Assuming optimistic foregrounds, however, changes the picture significantly, as the large amount of wavenumbers k_{21} accessible, and the great precision for each of them, allows all ETHOS models we study to be distinguishable from CDM and feedback at $\chi^2 > 100$. The situation is similar in the comparison with WDM, shown in Fig. 4.6.16. Pessimistic foregrounds can still differentiate ETHOS models from WDM at $\chi^2 > 10$ for $h_{\rm peak} \geq 0.4$, as long as $k_{\rm peak} \leq 10^2\,h/$Mpc. Here, again, optimistic foregrounds would open a larger swath of parameter space, as only models with $h_{\rm peak} < 0.1$ can be confounded with WDM in that case. This shows that great progress can be made even when all 21-cm modes within the foreground wedge are discarded, yet the gains from recovering those modes would dramatically enhance our understanding of the dark sector.

Chapter 5

Detectability of dark acoustic oscillations in the Lyman-α forest

This chapter is based on the following article:

ETHOS – an Effective Theory of Structure Formation: detecting dark matter interactions through the Lyman-α forest

Published in Monthly Notices of the Royal Astronomical Society, Volume 487, Issue 1, pp.522-536 (2019), Oxford University Press

Authors:
Sownak Bose[1], Mark Vogelsberger[2], Jesús Zavala[3], Christoph Pfrommer[4], Francis-Yan Cyr-Racine[5,6], Sebastian Bohr[3] and Torsten Bringmann[7]
[1]Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[2]Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Center for Astrophysics and Cosmology, Science Institute, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland
[4]Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany
[5]Department of Physics, Harvard University, Cambridge, MA 02138, USA
[6]Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, USA
[7]Department of Physics, University of Oslo, Box 1048, N-0371 Oslo, Norway

We perform a series of cosmological hydrodynamic simulations to investigate the effects of non-gravitational dark matter (DM) interactions on the intergalactic medium (IGM). In particular, we use the ETHOS framework (Cyr-Racine et al., 2016; Vogelsberger et al., 2016) to compare statistics of the Lyman-α forest in cold dark matter (CDM) with an alternative model in which the DM couples strongly with a relativistic species in the early universe. These models are characterised by a cutoff in the linear power spectrum, followed by a series of 'dark acoustic oscillations' (DAOs) on sub-dwarf scales. While the primordial cutoff delays the formation of the first galaxies, structure builds-up more rapidly in the interacting DM model compared to CDM. We show that although DAOs are quickly washed away in the non-linear clustering of DM at $z \lesssim 10$, their signature can be imprinted prominently in the Lyman-α flux power spectrum at $z > 5$. On scales larger than the cutoff ($k \sim 0.08$ s/km for the specific model considered here), the relative difference to CDM is reminiscent of a warm dark matter (WDM) model with a similar initial cutoff; however, the redshift evolution on smaller scales is distinctly different. The appearance and disappearance of DAOs in the Lyman-α flux spectrum provides a powerful way to distinguish interacting DM models from WDM and, indeed, variations in the thermal history of the IGM that may also induce a small-scale cutoff.

5.1 Introduction

In the standard picture of structure formation, the enigmatic dark matter (DM) particle is assumed to be a kinematically cold, collisionless and non-baryonic entity. This 'cold' dark matter (CDM) model has predictive power, a feature that has been exploited over the past four decades in a rigorous campaign of numerical simulations that has established CDM as part of the standard cosmological paradigm. The great success of this model lies in the finding that the same theory that accounts for the temperature anisotropies in the cosmic microwave background at early times (Spergel et al., 2003; Planck Collaboration et al., 2016) has been similarly successful at reproducing the large-scale clustering of galaxies at present day (Colless et al., 2001; Cole et al., 2005; Eisenstein et al., 2005; Zehavi et al., 2011).

At the regime of dwarf galaxies, however, a number of "small-scale challenges" have been claimed to afflict the CDM paradigm. Chief amongst them are the so-called "Missing Satellites" (e.g. Klypin et al., 1999; Moore et al., 1999), "Too Big to Fail" (Boylan-Kolchin et al., 2011), cusp-core (e.g. Flores & Primack, 1994; de Blok et al., 2001) and plane of satellites problems (e.g. Ibata et al., 2014; Pawlowski et al., 2014). For a thorough overview of this subject, we refer the reader to the recent review by Bullock & Boylan-Kolchin (2017). While these issues have been used to motivate DM candidates beyond CDM, it should be cautioned that the small-scale problems are only firmly established with simulations that include the modelling of the CDM component only. These issues may in fact be resolved within the CDM paradigm once the impact of gas and stellar physics is better understood and fully taken into account.

This has been permitted by the increasing sophistication of hydrodynamical simulations (e.g. Vogelsberger et al., 2014b,c; Dubois et al., 2014; Schaye et al., 2015; Springel et al., 2018), which self-consistently track the co-evolution of dark and baryonic matter. Although hydrodynamical simulations differ in detail, they have shown universally that the interaction between DM and baryons through processes associated with galaxy formation – such as gas cooling, photoionisation and feedback – change both the census of the galactic population (e.g. Kauffmann et al., 1993; Kim et al., 2018; Read et al., 2017), as well as the internal structure of DM haloes (e.g. Pontzen & Governato, 2014; Di Cintio et al., 2014; Sawala et al., 2016a; Fitts et al., 2017) relative to simulations modelling only the DM component.

In fact, perhaps the greatest challenge to the CDM model, at least in terms of its appeal as a complete structure formation theory, is that despite intense efforts at discovering CDM-like particles, the search has been fruitless so far. Most of these efforts have focused on Weakly Interacting Massive Particles (WIMPs), which are one of the best-motivated CDM candidates in great part due to their potential connection with supersymmetry (Jungman et al., 1996). After the successful discovery of the Higgs boson, it was hoped that the Large Hadron Collider (LHC) would find evidence for supersymmetry, giving credence to WIMPs and the CDM model, but thus far the LHC has failed to provide any evidence of this kind. WIMPs have also remained elusive to both direct (Aprile et al., 2018) and indirect detection (e.g. Albert et al., 2017) experiments. Furthermore, promising observational anomalies that might be connected to DM, have either disappeared or explained with non-DM astrophysical sources, e.g., the excess of gamma rays at the Galactic Centre, which has been ascribed to the self-annihilation of WIMPs (e.g. Hooper & Goodenough, 2011; Daylan et al., 2016) may instead be explained by a population of unresolved millisecond pulsars (e.g. Bartels et al., 2016; Lee et al., 2016; Fermi-LAT Collaboration, 2017), or an overdensity of stars in the Galactic bulge (Macias et al., 2018).

As long as the DM remains undetected in the laboratory, it is worth considering well-motivated alternatives to CDM and their implications for structure formation. A popular alternative is *warm* dark matter (WDM, Bond & Szalay, 1983; Colín et al., 2000; Bode et al., 2001), in which the DM particles have a non-negligible velocity dispersion in the early universe. The resulting free-streaming of these particles suppresses density fluctuations below a characteristic scale determined by the rest mass of the particles and their thermal history; this delays the formation of the first structures and reduces the abundance of low-mass galaxies in the process (e.g. Zavala et al., 2009; Macciò & Fontanot, 2010; Lovell et al., 2012; Schneider et al., 2012; Kennedy et al., 2014; Bose et al., 2016a; Bozek et al., 2019). In the linear regime, the free-streaming of WDM particles is manifest as a nearly exponential cutoff in the linear power spectrum relative to CDM. In other well-motivated DM models, there may exist a coupling between the DM and a relativistic species (e.g. neutrinos or 'dark' radiation) in the early universe. In these so-called 'interacting' dark matter (iDM) models, the ensuing radiation pressure inhibits the growth of small-scale fluctuations and also results in a (collisional) cutoff in the linear power spectrum, but with a more complex behaviour than in WDM, exhibiting dark acoustic oscillations (e.g. Carlson et al., 1992; Bœhm et al., 2002; Ackerman et al., 2009; Cyr-Racine & Sigurdson, 2013; Buckley et al., 2014; Bœhm et al., 2014; Bringmann et al., 2016). Another promising alternative is offered by self-interacting dark matter (SIDM) models, in which multiple scattering events between DM particles can significantly change the internal structure of DM haloes compared to CDM in the non-linear regime (e.g. Spergel & Steinhardt, 2000; Yoshida et al., 2000; Davé et al., 2001; Colín et al., 2002; Vogelsberger et al., 2012; Rocha et al., 2013; Zavala et al., 2013; Elbert et al., 2015; Vogelsberger et al., 2014a; Kaplinghat et al., 2016; Robertson et al., 2018; Vogelsberger et al., 2019).

While these different DM species come from a diverse range of particle physics models with vastly different production mechanisms, the resulting effect on structure formation is similar in many cases. This is particularly evident in the case of WDM and iDM, both of which suppress small-scale structure by inducing a cutoff in the linear power spectrum. In this work, we consider examples of these models within the generalised framework of structure formation ETHOS (Cyr-Racine et al., 2016; Vogelsberger et al., 2016), which addresses such degeneracies by providing a mapping between parameters associated with DM physics and parameters relevant for structure formation. The flexibility afforded by this formalism potentially allows the investigation of a general class of model parameters (cutoff scale, self-interaction cross-section, DM-radiation coupling etc.) and their impact on the formation of galaxies without needing to simulate every point in the allowed parameter space. Previous analyses connected to this programme have focused on the predictions of these models for the internal content of DM haloes (Vogelsberger et al., 2016; Brinckmann et al., 2018; Sameie et al., 2018; Sokolenko et al., 2018), the diverse rotation curves of dwarf galaxies (Creasey et al., 2017), the tidal stripping of satellites in the Galactic halo (Dooley et al., 2016) and the possibility of detecting these DM candidates through gravitational lensing (Díaz Rivero et al., 2018). Finally, the most recent investigation has considered the high redshift galaxy population and reionisation history in this general class of models (Das et al., 2018; Lovell et al., 2018b). In this paper, we are particularly interested in the signatures of new DM phenomenology (see Section 5.2 for details) that may be imprinted in statistics of the Lyman-α forest.

The Lyman-α forest has proven to be a remarkably powerful probe of the nature of DM in the mildly non-linear regime. Measurements of the flux spectrum using observed QSO sightlines have been used repeatedly to infer the clustering of matter on these scales (e.g. Croft et al., 1998, 1999; McDonald et al., 2000; Palanque-Delabrouille et al., 2013). The flux spectrum is a particularly powerful probe of processes relating to early galaxy formation and the small-scale behaviour of DM particles by providing an insight on the matter power spectrum at relatively high redshift. In fact, it is now well-established that the *observed* flux spectrum displays a cutoff in power on scales smaller than $k \sim 0.03$ s/km, a feature that has been used to constrain the free-streaming properties of DM particles (e.g. Viel et al., 2005; Seljak et al., 2006; Viel et al., 2013; Baur et al., 2016; Iršič et al., 2017; Kobayashi et al., 2017; Murgia et al., 2018; Nori et al., 2019; Garzilli et al., 2019b).

A cutoff in the flux spectrum towards small scales may, however, originate from purely baryonic processes (e.g. Zaldarriaga et al., 2001; Peeples et al., 2010; Rorai et al., 2013; Nasir et al., 2016). The first effect, known as Jeans smoothing, is a result of increased gas pressure in the IGM due to boosted temperatures induced by the onset of reionisation. The degree of Jeans smoothing depends on both the integrated heat injection and exact timing of reionisation (e.g. Gnedin & Hui, 1998; Kulkarni et al., 2015; Oñorbe et al., 2017). A second effect is brought upon by random thermal motions of the gas, resulting in Doppler broadening of Lyman-α forest lines, further smoothing small-scale power in the flux spectrum. The degenerate behaviour of the thermal history of the IGM and the free-streaming properties of DM make it difficult to pinpoint the physical interpretation of the observed cutoff in the flux spectrum and may indeed relax current constraints on the rest mass of the WDM particle (Garzilli et al., 2017). Nevertheless, it is clear that the Lyman-α forest provides a unique probe into mildly non-linear scales at high redshift, which is the regime where most alternative DM models exhibit the strongest deviations from CDM. The main goal of this paper is to show that, if strong enough, non-gravitational features (dark acoustic oscillations) in the primordial power spectrum of iDM models can remain imprinted in the Lyman-α flux spectrum.

The layout of this paper is as follows. In Section 5.2, we briefly describe the DM particle physics model considered in this work, highlighting its connection to the general ETHOS framework. Section 5.3 describes the numerical setup used for this investigation, detailing our simulations and the analysis pipeline used to extract mock Lyman-α absorption spectra from them. Section 5.4 presents our main findings. Finally, our conclusions are summarised in Section 5.5.

5.2 Dark Matter model

In this work, we study structure formation in DM theories in which early-universe interactions with a relativistic species (see e.g. Carlson et al., 1992; Bœhm et al., 2002; Ackerman et al., 2009; Feng et al., 2009; van den Aarssen et al., 2012; Chu & Dasgupta, 2014; Buen-Abad et al., 2015; Bringmann et al., 2016; Chacko et al., 2016) lead to a modified initial spectrum of density fluctuations as compared to standard CDM. The general phenomenology of such models is described in detail in Cyr-Racine et al. (2016) within the ETHOS framework, while the nonlinear evolution of structure within these models was studied in Buckley et al. (2014) and Vogelsberger et al. (2016) (see also Bœhm et al., 2014; Schewtschenko et al., 2015). In these theories, the DM forms a fluid that is tightly-coupled to a relativistic species (e.g. neutrinos or dark radiation) at early times, much like the standard baryon-photon plasma before the epoch of recombination. Within this ‘dark’ fluid, the large radiation pressure prohibits the growth of DM fluctuations and allows the propagation of acoustic waves to large cosmological distances (Cyr-Racine et al., 2014). Just like the more well-known baryon acoustic oscillations (BAOs), these *dark acoustic oscillations* (DAOs) become imprinted on the spectrum of matter fluctuations at late times, providing us with a potential smoking gun for physical processes taking place in the early Universe. Due to the finite value of the coupling between DM and the relativistic species, the DAOs are usually damped on scales smaller than the radiation mean free path, in a process similar to standard Silk damping (Silk, 1968).

The resulting shape of the linear matter power spectrum is largely determined by how quickly the DM kinetically decouples from the radiation bath. Quantitatively, near the redshift of DM kinetic decoupling, $z_{\rm D}$, we have approximately:

$$(\dot{\kappa}_\chi/\mathscr{H})|_{z\sim z_{\rm D}} \simeq (z/z_{\rm D})^n, \tag{5.1}$$

where $\dot{\kappa}_\chi$ is the 'drag opacity' or interaction rate between DM and the relativistic species, and $\mathscr{H}$ is the conformal Hubble expansion rate. Note that $\dot{\kappa}$ denotes a derivative with respect to conformal time. In general, a larger value for the power law index[1], n, results in a greater number of undamped DAOs on the small-scale linear power spectrum. Once the non-linear evolution of density fluctuations is taken into account, models with low values of $n \lesssim 4$ exhibit structure formation that is reminiscent of standard WDM models (Vogelsberger et al., 2016; Murgia et al., 2017). On the other hand, models characterised by a large value of the power law index $n \gtrsim 6$ (which we hereafter refer to as "strong" DAO models; sDAO) have a structure formation history that is appreciably different from WDM, as first discussed in Buckley et al. (2014). We note that the models used in Vogelsberger et al. (2016), Lovell et al. (2018b), and Das et al. (2018) all have $n = 4$, and thus fall in the former category.

Since our aim is to study Lyman-α constraints on DM theories that have a structure formation history that is distinct from WDM, the present investigation focuses on sDAO models. In particular, we consider an atomic DM (Kaplan et al., 2010, 2011; Cyr-Racine & Sigurdson, 2013) in which DM is composed of two massive fermions that are oppositely charged under a new unbroken $U(1)$ dark gauge force. In this paradigm, the dark sector forms an ionised plasma at early times, until the temperature falls below the binding energy between the two oppositely charged particles, at which point neutral dark atoms form in a process reminiscent to cosmological hydrogen recombination. If this "dark" recombination occurs in or near thermal equilibrium, an extremely rapid kinematic decoupling epoch ensues due to the nearly exponential (Saha-like) nature of bound state formation in this case. This ensures that the power law index appearing in Eq. (5.1) is large ($n = 6$ for our sDAO model), resulting in a linear matter power spectrum composed of a significant number of undamped DAOs on small-scales.

We choose parameters of this model such that the linear matter power spectrum of our sDAO model starts deviating from its CDM counterpart near a comoving wavenumber of $k \sim 10h$ Mpc^{-1}, which is the scale where current observations of the Lyman-α spectrum become a powerful tool to discriminate different DM models. In the nomenclature of the ETHOS framework, the sDAO model is defined by:

$$\{n, a_n, \omega_{\rm DR}, \alpha_2, \alpha_{l\geq 3}\} = \left\{6, 6\times 10^8 {\rm Mpc}^{-1}, 1.25\times 10^{-8}, 9/10, 1\right\}$$

while the ETHOS-4 model[2] is defined by:

$$\{n, a_n, \omega_{\rm DR}, \alpha_{l\geq 2}\} = \left\{4, 414 {\rm Mpc}^{-1}, 1.35\times 10^{-6}, 3/2\right\}$$

---

[1] We note that this power law index is the same as that used to classify DM models within the ETHOS framework (Cyr-Racine et al., 2016).

[2] We note that the a_4 amplitudes given in Vogelsberger et al. (2016) should be divided by h to yield the correct values.

where n is the power law index defined in Eq. (5.1), a_n is the normalisation of the drag opacity at redshift $z_{\rm D} = 10^7$, $\omega_{\rm DR} \equiv \Omega_{\rm DR} h^2$ is the physical energy density in the dark radiation component in units of the critical density and α_l is a set of coefficients that defines the angular dependence of the DM-dark radiation scattering cross section. We refer the reader to Section II E of Cyr-Racine et al. (2016) for further details. The actual particle physics (i.e. Lagrangian) parameters of the SDAO model are listed in Appendix 5.6.2.

Fig. 5.3.1 illustrates the power spectrum appropriate to this model. For comparison, we also display the matter power spectra for CDM, ETHOS-4, as well as WDM thermal relics with mass 1.6 and 3.3 keV. The parameters for ETHOS-4 were especially selected to alleviate the Missing Satellites, Too Big To Fail and core-cusp “problems” in CDM (Section 5.1) through DM physics alone (Vogelsberger et al., 2016). Unlike WDM, in which the cutoff continues indefinitely, the ETHOS models show a resurgence of power on smaller scales due to the aforementioned DAOs. Note that because of these DAOs, ETHOS models, in particular sDAO models, have increased small-scale power compared to WDM models with a cutoff at the same scale. We note that the SDAO model, which is our main focus in this paper, may already be strongly constrained by present observations. Our goal here is to investigate if small-scale DAOs may be at all detectable in the Lyman-α forest, rather than to construct a model that matches the data. For this reason, we opt to simulate an iDM scenario that maximises differences relative to CDM on scales large enough that they may be captured at moderate numerical expense.

5.3 Numerical setup

5.3.1 Simulations and initial conditions

The simulations we present in this work make use of the cosmological simulation code, AREPO (Springel, 2010). AREPO employs a hybrid tree/particle-mesh scheme to solve for gravitational interactions of DM particles, and a moving, unstructured Voronoi mesh to solve equations of hydrodynamics. The moving mesh is adaptive in nature, resolving fluids in regions of high density with many more cells of a smaller size than in low density environments. AREPO has been augmented with a comprehensive model for galaxy formation (Weinberger et al., 2017; Pillepich et al., 2018a) which we use here. In addition, Vogelsberger et al. (2016) presents an updated version of AREPO which, in addition to the galaxy formation models mentioned above, also incorporates elastic, isotropic self-interactions of DM particles, while allowing for arbitrary velocity-dependent interaction cross-sections (using an algorithm adapted from the original described in detail in Vogelsberger et al. 2012). While the self-scatterings of DM particles have a pronounced impact on the internal structure of haloes at late-times,

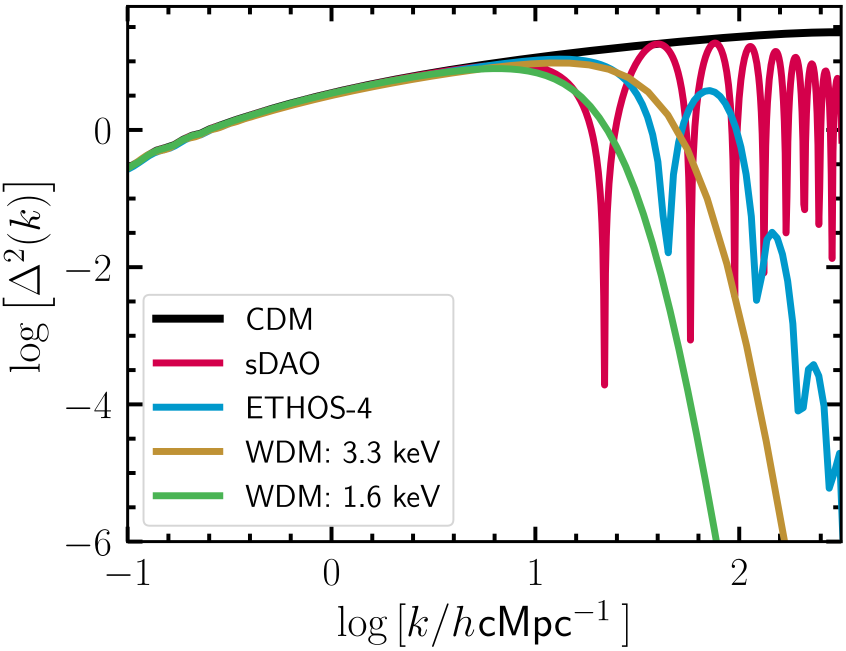

Figure 5.3.1. Dimensionless power spectra $\left[\Delta^2(k) = k^3 P(k)\right]$ *for the CDM (black) and* sDAO *(red) models used in this work. For comparison, we also show the power spectra for the less extreme* ETHOS-4 *model (blue; see Vogelsberger et al. 2016) which exhibits a deviation from CDM at a scale comparable to that of a 3.3 keV thermal relic WDM particle (in yellow). On the other hand, the cutoff scale for the* sDAO *model is closer to that of a 1.6 keV thermal relic (green). Furthermore, the amplitude of the dark acoustic oscillations (DAOs) in the* sDAO *model is considerably larger than in* ETHOS-4.

their influence on the IGM at high redshifts will be sub-dominant to that induced by the cutoff in the power spectrum; we have therefore turned off self-interactions in the simulations.

Our high resolution simulations follow the evolution of 2×512^3 DM and gas particles from $z = 127$ to $z = 0$ in a periodic box of (comoving) size 29.6 cMpc (20 h^{-1}cMpc), resulting in an effective DM particle mass of $6.41 \times 10^6\,\mathrm{M}_\odot$. An individual gas cell has a target mass of $1.01 \times 10^6\,\mathrm{M}_\odot$. This target gas mass also corresponds to the typical mass of a stellar macro-particle representing a stellar population. We enforce that the mass of all cells is within a factor of two of the target mass by explicitly refining and de-refining the mesh cells. The comoving softening length for DM particles is set to 1.19 kpc, while the (adaptive) softening applied to a gas cell is set to a comoving minimum value of 185 pc. To check for convergence, we also run a second set of simulations a factor of two lower in resolution.

We use the fiducial IllustrisTNG galaxy formation model (Weinberger et al., 2017; Pillepich et al., 2018a) with one change. Namely, we have turned off the magnetohydrodynamics solver as it is not relevant for the analysis presented here. As in the fiducial TNG model, each of our simulations is set up with a time-dependent, spatially uniform ionising background as described in the model by Faucher-Giguère et al. (2009). The TNG model is built upon the original Illustris galaxy formation model described in Vogelsberger et al. (2013).

Initial conditions for all simulations were generated using the MUSIC code (Hahn & Abel, 2011), assuming cosmological parameters derived from Planck Collaboration et al. (2016): $\Omega_0 = 0.311$ (total matter density), $\Omega_\mathrm{b} = 0.049$ (baryon density), $\Omega_\Lambda = 0.689$ (dark energy density), $H_0 = 67.5$ kms^{-1}Mpc^{-1} (Hubble parameter) and $\sigma_8 = 0.815$ (linear rms density fluctuation in a sphere of radius 8 h^{-1} Mpc at $z = 0$). The dimensionless linear power spectra used to generate initial conditions are shown in Fig. 5.3.1. While the CDM power spectrum exhibits power on all scales, the two ETHOS models cutoff at $\log[k/h\mathrm{cMpc}^{-1}] \approx 1$. In this paper we will be concerned with the sDAO model, in which the model parameters have been adjusted to amplify the effect of DAOs, as explained in the previous section. Our goal is to investigate the extent to which the characteristics of DAOs in the ETHOS models can be probed using the Lyman-α forest. To put our results in context, we have also run simulations of the ETHOS-4 and 1.6 keV WDM models at our default resolution. The choice of a 1.6 keV thermal relic is motivated by the fact that the free-streaming scale in this model is identical to the cutoff in sDAO; this helps disentangle small-scale differences induced by the acoustic oscillations from those that are caused by a primordial cutoff. The simulations are analysed to perform mock Lyman-α observations using the procedure that we describe in the following subsection. Finally, we note that simulations that resolve the primordial power spectrum cutoff are plagued with artificial fragmentation of filaments that condense into 'spurious' haloes (e.g. Wang & White, 2007; Lovell et al., 2014). These objects are seeded by discreteness of the particle set rather than a true gravitational instability, and must hence be excluded from the analysis. This is a well-known problem in WDM simulations, but is less severe in the ETHOS models which have added small-scale power in the form of DAOs (Buckley et al., 2014, see also Fig. 5.6.11). This is especially true at high redshift, which is the regime of interest in this paper. As such, we do not perform any extra steps to classify these objects in the simulations we have run.

5.3.2 Creating Lyman-α mock absorption spectra

From the outputs of each simulation, we generate synthetic absorption spectra using the methodology outlined in Altay & Theuns (2013). In short, at each output time, we select 1024 randomly-selected skewers[3] oriented parallel to a coordinate axis of the box. Gas cell properties are interpolated onto locations along each skewer using a smoothing kernel; we follow Altay & Theuns (2013) and employ a truncated Gaussian kernel, $G_t(r, \sigma)$, which is defined as:

---

[3] We have checked explicitly that our results are converged for this choice for the number of sightlines (see Fig. 5.6.12).

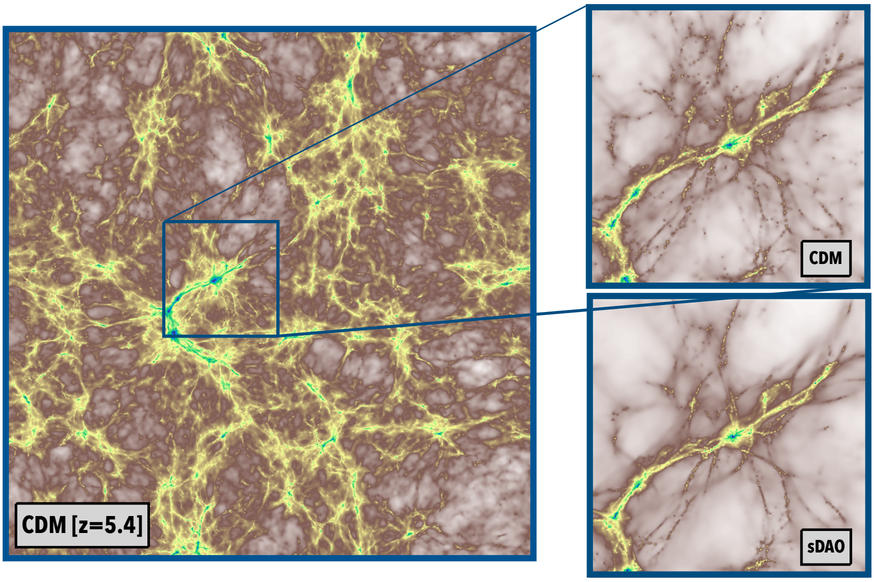

Figure 5.3.2. Images of the gas density at $z = 5.4$ obtained from our hydrodynamical simulations. On the left, we project the simulation box along the z-axis in a projection of comoving dimensions $(20 \times 20 \times 4)\ h^{-1}$cMpc. The smaller panels zoom into a region centred on the most massive halo at this redshift in a window of size $(4 \times 4 \times 2)\ h^{-1}$cMpc in CDM (upper right) and the equivalent region in sDAO *(lower right). While differences are hard to discern on these scales, the small-scale cutoff in the* sDAO *model results in a smoother matter distribution than the CDM volume at the same epoch. These images were processed using the publicly-available* PY-SPHVIEWER *package (Benitez-Llambay, 2015).*

$$G_t(r,\sigma) = \mathcal{N} \begin{cases} \exp\left(-A^2 r^2\right), & \text{for}\, r \leq h_{sml} \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

where:

$$\begin{aligned} \sigma^2 &= \frac{h_{sml}^2}{8\pi^{1/3}}, \\ A^2 &= \frac{4\pi^{1/3}}{h_{sml}^2}, \\ \mathcal{N} &= \frac{8}{\pi h_{sml}^3}\left[\text{erf}(t) - \frac{2t\exp\left(-t^2\right)}{\sqrt{\pi}}\right]^{-1}, \\ t &= 2\pi^{1/6}, \end{aligned} \tag{5.3}$$

and $r = h_{sml}$ is the radius at which the Gaussian kernel is truncated. In smoothed particle hydrodynamics (SPH) simulations, h_{sml} is taken to be the gas particle's smoothing length, calculated using a fixed number of nearest-neighbours. In AREPO, gas is discretised in the form of Voronoi cells rather than SPH particles; we therefore define an 'effective' smoothing length, $h_{sml,i}$, for each gas cell, i, as:

$$h_{sml,i} = f\left(\frac{3m_i}{4\pi\rho_i}\right)^{1/3}, \tag{5.4}$$

where m_i and ρ_i, respectively, are the mass and mass density of the gas cell, and $f = 4$ is some normalisation factor. Our results are insensitive to the precise choice of f.

By dividing each line-of-sight into N bins, we can compute the number density, $n_{\rm H}(j)$, H I-weighted temperature, $T(j)$ and H I-weighted peculiar velocity field, $v(j)$, at each bin j (in velocity space) using only the subset of gas cells that intersect each ray. Following exactly the methodology laid out in Theuns et al. (1998), we can then calculate the optical depth, $\tau(k)$, for the kth velocity bin along the line-of-sight as:

$$\begin{aligned} \tau(k) &= \sum_j \sigma_\alpha \frac{c}{V_{\rm H}(j)} n_{\rm H}(j)\,\Delta \times \frac{1}{\sqrt{\pi}} \exp\left(-\left[\frac{v(k)-v(j)}{V_{\rm H}(j)}\right]^2\right), \\ V_{\rm H}^2(j) &= 2k_{\rm B}T(j)/m_{\rm H}, \end{aligned} \tag{5.5}$$

where c is the speed of light, Δ is the width of each bin in units of physical distance, x, along the line-of-sight, $k_{\rm B}$ is the Boltzmann constant and $\sigma_\alpha = 4.45 \times 10^{-18}$ cm^2 is the cross-section of the hydrogen Lyman-α transition. The corresponding transmitted flux is then given by $F = e^{-\tau}$, where τ is the integrated optical depth along this line-of-sight.

Due to the considerable uncertainty about the level of photoionisation, we follow the standard procedure of rescaling our simulated spectra to the observed optical depth at the corresponding redshifts. At low redshift, $z \sim 2-3$, the rescaling factor is small; at high redshift, $z \sim 4-6$, however, the rescaling becomes increasingly important as fluctuations in the assumed UV background start to become an issue. In particular, at each simulation output, we rescale the optical depths of simulated spectra such that the *mean* transmitted flux matches the *observed* mean flux at that redshift. For the observed mean fluxes, we use the values reported by Walther et al. (2018) for $z < 4$ and by Viel et al. (2013) for $z \geq 4$. The factor by which the CDM and SDAO spectra are rescaled are not too dissimilar at $z > 3$, and are almost identical at lower redshift (see Fig. 5.4.6).

This rescaling procedure is widespread in the simulation community and its validity is worth reflecting on for a moment. As we have mentioned previously, the motivation for rescaling the optical depth is the uncertainty of the photoionisation rate. In practice, one assumes that the H I abundance is directly proportional to the photoionisation rate and hence, the optical depth can be rescaled by the same factor. The assumption implicitly neglects the following effects, which we consider in turn:

1. Deviations from equilibrium: these only play an important role during reionisation and determine how strongly the gas is being heated. At the redshift of the observations of the Lyman-α forest, the ionisation degree should be back in equilibrium. The only remaining effect is the slightly enhanced temperature.

2. Collisional ionisation: this is a negligible effect at IGM temperatures.
3. A constant recombination rate: since the H I fraction is of order $\sim 10^{-4}$, the H II fraction and hence the recombination rate practically do not change for small variations of the H I abundance.
4. Spatial fluctuations of the photoionisation rate: these disappear quickly after the end of reionisation. For very high gas densities, e.g. in damped Lyman-α systems (DLAs), one would have to take into account self-shielding of gas.

In summary, the rescaling procedure is valid for the optical depths and temperatures of the IGM (see, e.g., the discussions in Rauch et al., 1997; Weinberg et al., 1997; Theuns et al., 1998; Bolton et al., 2005). At $z > 5.5$, the assumption of a homogeneous photoionisation rate fails severely and it is then necessary to perform radiative transfer calculations. The validity of rescaling the mean flux at these high redshifts then becomes questionable. In what follows, we limit our analysis to $z \leq 5.4$.

5.4 Results

5.4.1 The clustering of matter

As a precursor to the main analysis in this paper, we show in Fig. 5.3.2 the projected gas density map from our simulation volumes at $z = 5.4$. The largest panel shows a $(20 \times 20 \times 4)\, h^{-1}$cMpc projection from the CDM simulation; the smaller panels zoom into a $(4 \times 4 \times 2)\, h^{-1}$cMpc region centred on the most massive halo at this time as it appears in the CDM (upper right) and sDAO (lower right) simulations. While general large-scale filaments and knots look identical in the two density maps, there is noticeable absence of small-scale structure in the sDAO image, in which the gas density distribution is smoother than in CDM. This situation is identical to what is observed in standard WDM simulations, in which the matter distribution is smoothed through free-streaming induced by the cutoff in the linear power spectrum, although the mechanism in operation here is collisional Silk damping, rather than free-streaming. The smoothed gas distribution in WDM models is manifest as a cutoff in the Lyman-α flux spectrum at small-scales; our aim in the subsequent sections is to investigate if the *resurgence* in power at small-scales – predicted by models with strong DAOs, but not by thermal relic WDM – can be probed by the Lyman-α forest.

Before examining the Lyman-α forest, it is instructive to first look at the DM distribution predicted by these models. Fig. 5.3.3 shows ratios (sDAO to CDM) of the non-linear DM power spectrum at $z = 20, 14, 10, 8$ and 6 (coloured lines), measured directly from the DM particles in each simulation at the corresponding redshifts. For comparison, we also show, in black, the ratio of the *linear* power spectra in these models,

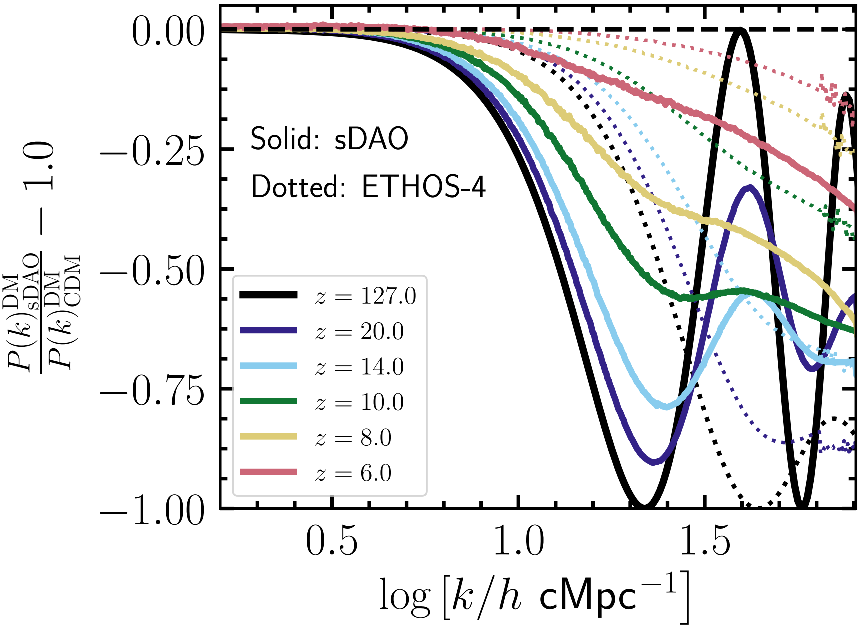

Figure 5.3.3. Redshift evolution of the ratio of the non-linear matter power spectrum at $z = 20, 14, 10, 8, 6$ *for the simulations presented in this paper. The power spectra are measured using only the distribution of DM particles in the simulation snapshots. For comparison, we also show the ratio of the linear power spectra used as the input to making initial conditions at* $z = 127$*; we also show results for the* ETHOS-4 *model with dotted line. While the matter distribution shows significant differences between the* SDAO *and CDM model at high redshift (including the signature of DAOs), these differences are suppressed at lower redshift. In particular, the DAOs are no longer visible in the DM distribution at* $z = 6$*. On the other hand, DAOs are nearly smoothed as early as* $z = 20$ *in the* ETHOS-4 *model due to our finite numerical resolution.*

which were used to generate the initial conditions at $z = 127$. The ETHOS-4 model, comparable to a 3.3 keV thermal relic, is represented by the dotted lines.

From Fig. 5.3.3, one clearly notices that the characteristic DAO peaks are very prominent at early times. At $z = 20$ and $z = 14$, the first DAO peak is still noticeable in the DM distribution (at $\log[k/h\mathrm{cMpc}^{-1}] \approx 1.6$), but only marginally so by $z = 10$. As gravitational collapse continues at $z < 10$, increasing the overall power on all scales, the DAO peak is gradually washed away as a result of mode coupling in the (weakly) non-linear regime of structure formation. By $z = 6$, any signature of DAOs has completely disappeared – qualitatively, the ratio of the non-linear power spectra looks more similar to an ordinary thermal relic WDM particle. This is consistent with the findings of Buckley et al. (2014) and Vogelsberger et al. (2016) who also noted the absence of acoustic peaks in the DM distribution at relatively high redshift ($z \sim 6-8$). This is particularly true for the ETHOS-4 model, where the DAOs are nearly absent as early as

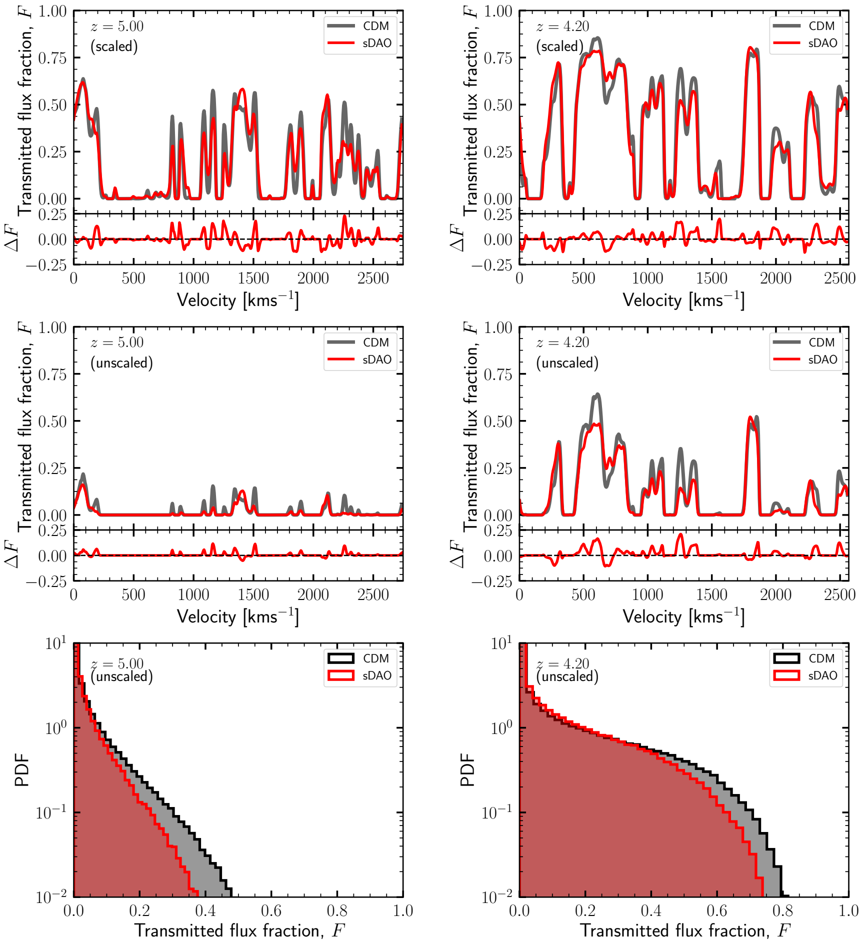

Figure 5.3.4. Top row: Synthetic mock spectra extracted from our simulations at $z = 5$ (left) and $z = 4.2$ (right). Each panel shows the transmitted flux fraction, F, in velocity space for a specific randomly-selected line-of-sight through the simulation volume at the corresponding redshift. The lines-of-sight are chosen such that the same spatial region is probed in the CDM (black) and sDAO *(red) simulation volumes. Clearly, more intervening structure can be inferred from the CDM spectra as evidenced by the deeper transmission lines than in* sDAO*. The lower panels show the relative difference in transmitted flux i.e., $\Delta F = F_{\mathrm{CDM}} - F_{\mathrm{sDAO}}$. Middle row: The same lines-of-sight shown in the top row but before we rescale the mean flux to the observed values. Bottom row: The unscaled flux PDF at these redshifts. These panels show the "true" difference in the transmitted flux between the CDM and* sDAO *models i.e. by removing any artefacts that may be brought in by the different amount of rescaling required for the two DM models.*

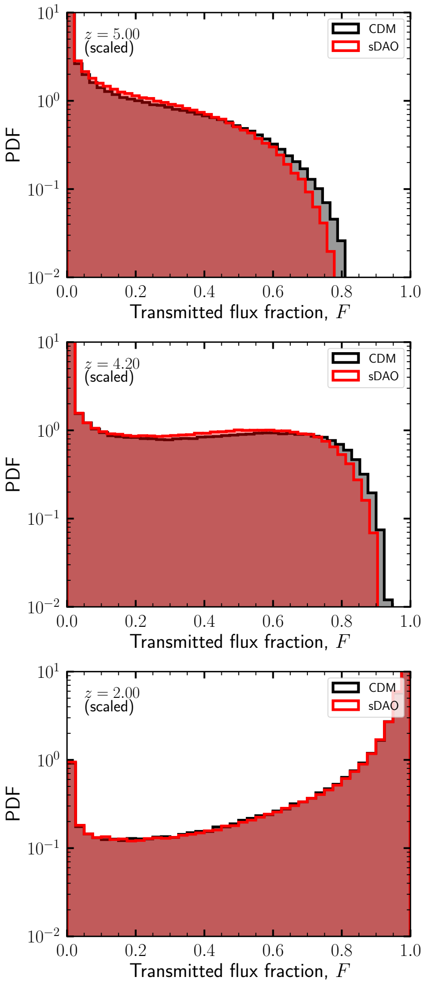

Figure 5.4.5. Comparison of the PDF of the transmitted flux, F, at $z = 5, 4.2, 2$ (top to bottom) after rescaling the mean transmitted flux. While both CDM and sDAO *show identical PDFs at low redshift, the CDM simulations display an extended tail of high flux at higher redshift.*

$z = 20$ (given the finite resolution of our numerical setup).

5.4.2 The Lyman-α forest at a glance

In the top row of Fig. 5.3.4 we present the transmitted flux fraction, F, measured along random lines-of-sight at $z = 5$ (left) and $z = 4.2$ (right) for the CDM and sDAO models. The spectra have been created following the procedure outlined in Section 5.3.2, after rescaling the mean transmission in each simulation box to the observed transmitted flux. In each case, the lines-of-sight have been chosen so as to probe the same spatial region in the two simulation volumes. The lower sub-panels show the difference in the transmitted flux, $\Delta F = F_{\rm CDM} - F_{\rm sDAO}$. This figure highlights the fact that the same line-of-sight probes different intervening structure in the IGM of the two simulations. In particular, it is clear that, just as in WDM, absorption lines are in general deeper for a random line-of-sight in the CDM simulation than for SDAO model, signifying the presence of a more clumpy IGM. This difference is a direct consequence of the cutoff in the initial power spectrum.

The SDAO and CDM models have been rescaled by different amounts when matching the simulated spectra to the observed mean transmitted flux. To disentangle the implications of rescaling from the different line shapes due to the modified cosmology at a given UV background, the middle row of Fig. 5.3.4 shows the *unscaled* line-of-sight spectra at $z = 5$ and $z = 4.2$. Clearly, there are residual differences between the two models even in the unscaled case. Quantitatively, these effects are seen more clearly in the bottom row of Fig. 5.3.4, which shows the PDF of the unscaled transmitted flux for the two models at these redshifts. Considering just CDM to begin with, it is noticeable that as time proceeds, there is a cutoff in the flux PDF at higher values of F, which comes about due to a combination of two effects: i) as the universe expands, the background density drops and a given overdensity needs to be larger (in linear dimensions) in order to produce the same signal strength in the Lyman-α forest, and ii) as structure formation proceeds, the non-linear length scale moves to larger scales, implying that perturbations with larger wavelengths start to collapse, yielding a more structured (clustered) universe. As a result, the extended tail to large F builds as gravitational collapse proceeds through cosmic time. There is a clear extended tail of high flux in CDM which is less prominent in SDAO. This can be ascribed to the delayed collapse of the first haloes as a result of the suppressed small-scale density fluctuations induced by the DM-radiation coupling in the SDAO model.

A more realistic comparison of the two models is shown in Fig. 5.4.5, which displays the *scaled* flux PDFs as a function of redshift. The panels, from top to bottom, show the flux PDFs at $z = 5, 4.2$ and 2. This figure highlights a qualitative difference in the manner in which structure formation proceeds in the SDAO model compared to CDM. At $z = 5$, for example, the flux PDF is truncated at somewhat lower values of F than it is in the CDM simulation. This can be ascribed to the delayed collapse of the first haloes as a result of the suppressed small-scale density fluctuations induced by the DM-radiation coupling in the SDAO model. The same difference, though smaller, is also manifest in the flux PDF at $z = 4.2$. By $z = 2$, however, the flux PDFs are almost indistinguishable between the two models. This is one of the generic features of models that exhibit a primordial cutoff in the linear power spectrum: while the formation of the first galaxies is delayed, structure formation proceeds more rapidly than in CDM afterwards. For the case of WDM, this has been demonstrated in detail in Bose et al. (2016b, 2017). As we show in the following subsection, the same qualitative behaviour is manifest in the 1D flux power spectrum as well.

Fig. 5.4.6 displays another way of visualising the different redshift evolution in the CDM and SDAO models. Here we show the *unscaled* mean transmitted flux as a function of redshift, compared to the observed values that we rescale to. Mean fluxes in

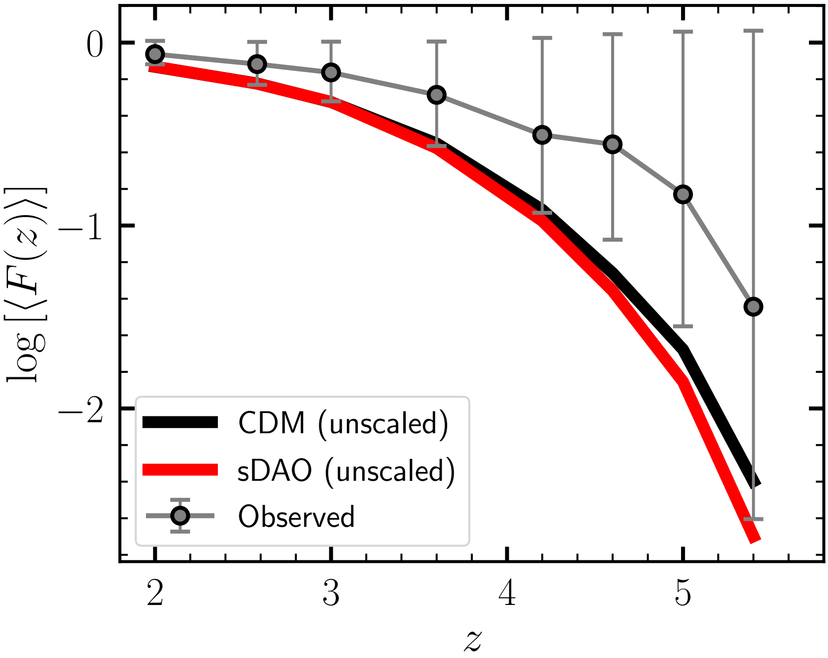

Figure 5.4.6. Redshift evolution of the (unscaled) mean transmitted flux in the two DM models compared to observed value of the mean transmission at each redshift. While SDAO *spectra are rescaled by a larger factor than CDM spectra at $z \geq 4$, the rescaling is almost identical towards lower redshift. Data obtained from Viel et al. (2013) and Walther et al. (2018).*

the SDAO model are systematically lower at $z \geq 4$ but nearly identical at later times. Figs. 5.4.5 and 5.4.6 therefore show explicitly the effects of delayed structure formation in the SDAO model, a direct consequence of the intrinsic cutoff in the linear power spectrum.

5.4.3 The 1D flux spectrum

Next, we investigate if the distinctive feature of the SDAO model, the small-scale acoustic oscillation, is detectable in the Lyman-α forest. To probe this feature, we compute the 1D Lyman-α flux power spectrum. Following Viel et al. (2013), at redshift z, we compute the power spectrum, $P_{1\mathrm{D}}(k)$, of the *fractional transmission*, $\delta_F(z)$, which is defined as:

$$\delta_F(z) = \frac{F(z) - \langle F(z)\rangle}{\langle F(z)\rangle} \tag{5.6}$$

where $\langle F(z) \rangle$ is the mean transmitted flux at redshift z. As described in Section 5.3.2, the mean flux at every snapshot is rescaled to match the observed mean flux at that redshift. The power spectrum is calculated in this way for each of the 1024 lines-of-sight at a given redshift; the resulting value of $P_{\rm 1D}(k)$ at that redshift is then obtained by taking the mean value of the individual power spectra at each k mode.

Fig. 5.4.7 shows the results of this procedure for the CDM and SDAO simulations over a range of redshifts. Each of the models is represented by a shaded region which denotes the uncertainty in observed mean transmission at that redshift reported by Viel et al. (2013) ($z \geq 4$) and Walther et al. (2018) ($z < 4$), which we have propagated through to the normalisation of the simulated power spectra. In each panel we also display the observed flux spectra at each redshift, with data compiled by Kim et al. (2004), Viel et al. (2013), Iršič et al. (2017) and Walther et al. (2018) from the MIKE/HIRES and XQ-100 quasar spectra samples. We do not include large-scale flux power spectrum measurements from BOSS (Palanque-Delabrouille et al., 2013) as the CDM and SDAO models are identical on these scales ($k \lesssim 0.02$ s/km).

The redshift evolution of the flux power spectra is reminiscent of the behaviour seen in Figs. 5.3.3 and 5.4.5, in which the stark differences between the CDM and SDAO models diminish with redshift. For example, at $z = 5.4$ the power spectra for the two models match only on the very largest scales ($k \leq 0.2$ s/km); in contrast, their power spectra are identical across all scales by $z = 2$.

While the flux power spectra measurements at $z \leq 4.6$ obtained from the CDM simulation are in good agreement with the data even at the smallest scales (at least within the errors afforded by the uncertainty in the mean transmitted flux), this is not so for the two highest redshift bins ($z = 5, 5.4$). In particular, the simulated flux power spectra show a sharper decrement of power than is observed at scales smaller than $k \sim 0.07$ s/km. Part of this discrepancy may be due to incomplete masking of metal lines, which could add artificial power at small-scales. This effect would be more pronounced at higher redshift where masking all metal contributions is more challenging (e.g. Walther et al., 2018). The bulk of the discrepancy, however, can be pinpointed to numerical resolution, as shown explicitly in the convergence tests performed by Bolton et al. (2017, see their Fig. A4). These authors show that at $z \sim 5$, a simulation where each gas element is $\sim 10^6\,\mathrm{M}_\odot$ (similar to ours) can show a deficit of small-scale power of around 30% at $k \sim 0.1$ s/km compared to a higher resolution simulation with 8 times better mass resolution. This difference diminishes with redshift. Given that we are mostly interested in the *relative* difference between the CDM and SDAO models, however, this difference is not critical; the comparison with observations serves mainly as a consistency check of our procedure for generating mock spectra from our simulations.

Bearing in mind that higher resolution only increases the small-scale power by 10-30%, Fig. 5.4.7 shows that the SDAO model is in clear tension with the data on scales smaller than $k \sim 0.04$ s/km at $z \lesssim 4.6$. This is expected, considering that the linear theory cutoff in the SDAO model is similar to that of a 1.6 keV thermal relic, which may already be ruled out by existing Lyman-α constraints (e.g. Viel et al., 2013; Baur et al., 2016). However, as we have remarked in Section 5.1, constraining models against observed data by means of their relative normalisation is fraught with uncertainties due to the assumed thermal history of the IGM. We are therefore cautious

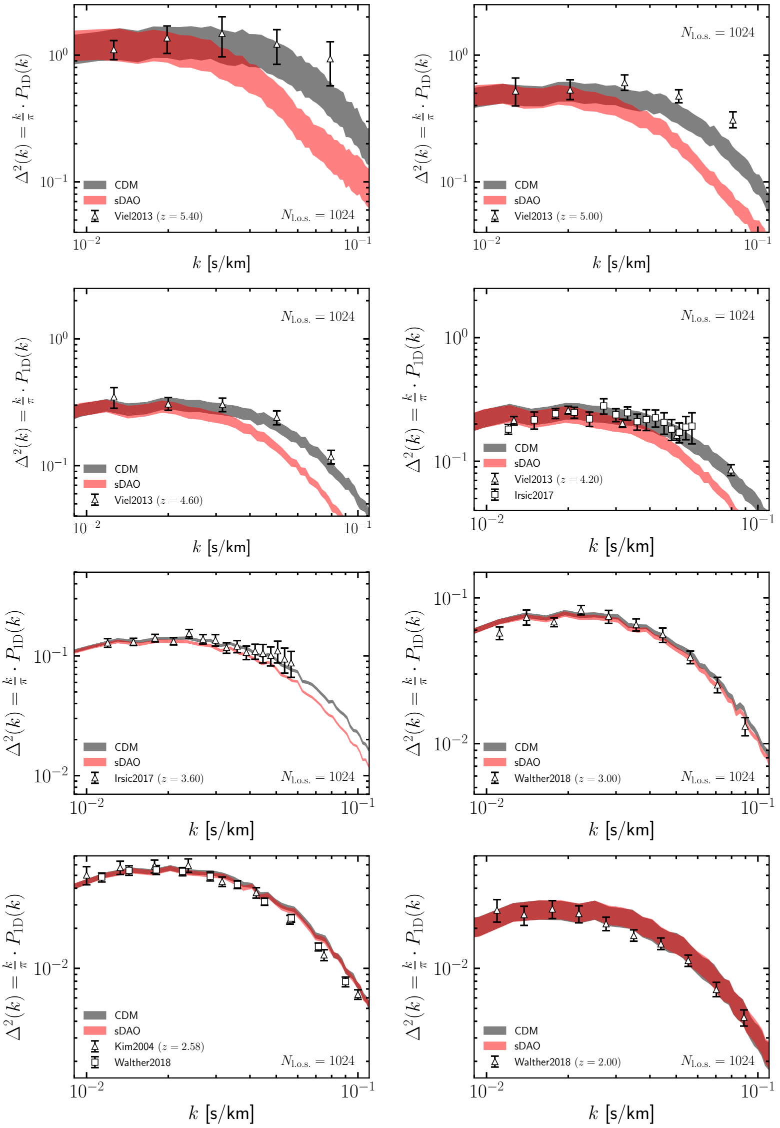

Figure 5.4.7. The 1D Lyman-α flux spectra for CDM and sDAO*, compared to data obtained from MIKE/HIRES and XQ-100 quasar spectra at* $z = 5.4, 5, 4.6, 4.2, 3.6, 3, 2.58$ *and* 2*. To construct the simulated power spectra, we have, where possible, attempted to match the path length of the observed spectra. Each simulated spectrum makes use of 1024 lines-of-sight. The observational measurements are obtained from data compiled by Kim et al. (2004), Viel et al. (2013), Iršič et al. (2017) and Walther et al. (2018). The shaded regions encompass the reported uncertainty in observed mean transmission at that redshift, which translates to an uncertainty in the normalisation of the power spectra after rescaling.*

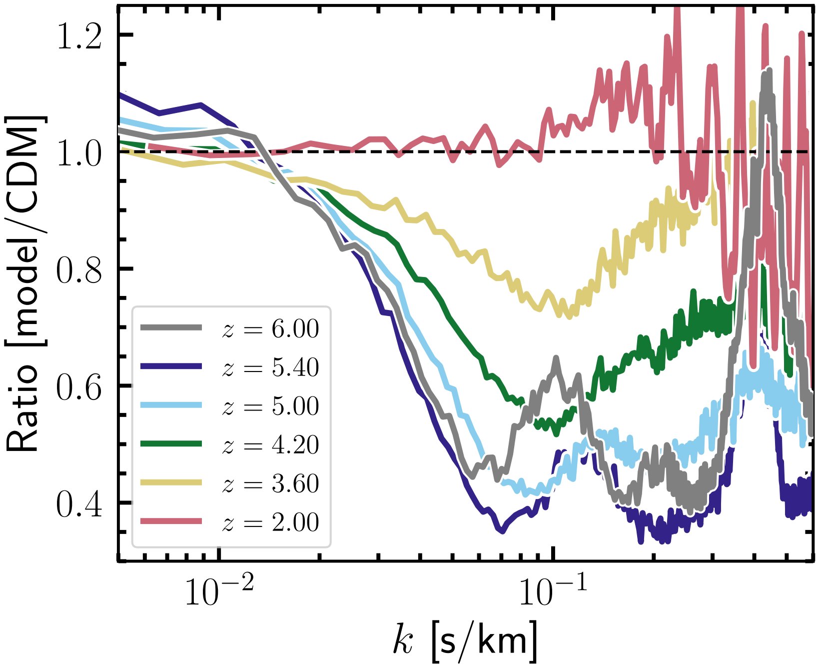

Figure 5.4.8. Ratio of the mean flux power spectra $\left[\Delta^2_{\mathrm{sDAO}}(k)/\Delta^2_{\mathrm{CDM}}(k)\right]$*. For clarity, we do not show the observational data in this figure. The signature of DAOs (at* $k = 0.4$ *s/km) can be noticed in the high redshift spectra, particularly at* $z = 6$ *and* $z = 5.4$*. Any evidence of DAOs is completely washed out by* $z = 2$*. Note that the 'bump' at* $k \sim 0.4$ *s/km is numerical, and is set by the finite resolution of our simulation setup. This secondary feature is not sourced by DAOs.*

of our interpretation of Fig. 5.4.7 with this caveat in mind.

In Fig. 5.4.8 we show the redshift evolution of the ratio of the (mean) flux spectra. This figure reveals the defining characteristics of the sDAO model. At $z \leq 4.2$, the behaviour relative to CDM is similar to what is observed in the case of WDM-like models: agreement with CDM on large-scales[4], followed by largely suppressed power below some characteristic scale. At $z \geq 5$, however, a 'bump' develops at $k \sim 0.13$ s km^{-1}, which becomes increasingly prominent at higher redshift. This is, indeed, the imprint of the DAO in the gas distribution at these early times. This feature is even more prominent at $z = 6$, where even the *second* DAO is visible at $k \sim 0.2$ s/km. In contrast, the 'bump' at $k \sim 0.4$ s km^{-1} that becomes increasingly prominent towards high redshifts is most likely a numerical effect associated with the finite resolution in our simulations, which affects different cosmologies differently. As the overall power increases across all scales between $z = 6$ to $z = 5.4$, mode coupling due to the (mildly) non-linear evolution erases

---

[4] The increased power on large-scales in sDAO compared to CDM is simply an artefact of rescaling the mean flux. As the small-scale power is suppressed heavily in the sDAO simulation, the large-scale power is boosted somewhat in order to achieve the same mean flux in the two models.

the second DAO bump and transfers its power to smaller and larger scales. As a result, the first DAO peak moves towards smaller scales. Note that the flux spectrum at $z = 6$ is shown simply for comparison, and does not necessarily represent the true ratio at this redshift, where the rescaling procedure may no longer be valid due to incomplete reionisation (see the discussion in Section 5.3.2).

Fig. 5.4.8 reveals the value of Lyman-α flux spectrum as a probe of small-scale clustering: while the 3D DM distribution showed no evidence of DAOs at $z \lesssim 10$ (Fig. 5.3.3), the linear scales probed by the flux power spectrum bears memory of the acoustic oscillations in the linear power spectrum of the sDAO model. This may be because the 1D flux spectrum, which can be qualitatively understood as an integrated version of the 3D power spectrum along the line-of-sight, weighted by velocity moments, is more sensitive to small-scale features in the linear power spectrum than the 3D clustering. This is somewhat reminiscent of modified theories of gravity (e.g. the $f(R)$ gravity model), in which the velocity divergence power spectrum (an integral of motion) has been shown to be a much more sensitive probe of deviations from standard gravity than simply the matter density field (e.g. Jennings et al., 2012; Bose et al., 2015). We leave a full understanding of the comparison between 1D and 3D power spectra for future work.

It is illuminating to consider the difference in structure in the sDAO and WDM models at these early times in greater detail. Fig. 5.4.9 compares the (cumulative) halo mass function in CDM, sDAO and 1.6 keV cosmologies at $z = 5.4$. In this calculation, halo mass is defined by M_{200}, which is the mass contained within r_{200}, the radius interior to which the mean density is equal to 200 times the critical density of the universe at that redshift. As expected, all three models agree on the abundance of the most massive haloes in the volume at these times ($M_{200} > 10^{10}\,\mathrm{M}_\odot$). Both the sDAO and WDM models then peel-away from the CDM curve at an identical mass scale; this is a direct consequence of the fact that the linear power spectra of these two models also deviate from CDM at identical scales. There is, however, a clear excess (of around a factor of 3) of haloes with $M_{200} < 3 \times 10^{9}\,\mathrm{M}_\odot$ in sDAO compared to the 1.6 keV simulation. This excess of power is sourced by the DAO, whereas the initial density fluctuations are suppressed indefinitely in the case of WDM. It is also interesting to note that while the effects of artificial halo formation is clear in the WDM case (as evidenced by the unnatural 'upturn' in the mass function at $M_{200} \sim 3 \times 10^{8}\,\mathrm{M}_\odot$; Wang & White 2007), the manifestation of these spurious haloes seems largely reduced in the sDAO model, in which any spurious halo formation is outnumbered by haloes that have collapsed out of true gravitational instability.

The left panel of Fig. 5.4.10 compares the relative difference of the flux spectra to CDM in the two models at $z = 5.4$. Power on scales larger than $k \sim 0.05$ s/km is suppressed by an almost identical amount, but the behaviour of the two models is different on smaller scales. In particular, while power continues to be suppressed in the case of the 1.6 keV thermal relic, the cutoff in the power is halted by the development of the DAO bump around $k \sim 0.13\ \mathrm{s\ km}^{-1}$, which is only present in the sDAO model and not in the WDM model. In practice, this may prove to be difficult to observe since the largest signal is expected to be present at the highest redshift, where the UV background starts to be inhomogeneous due to incomplete reionisation.

We also show predictions for the ETHOS-4 model in which the cutoff is on a smaller

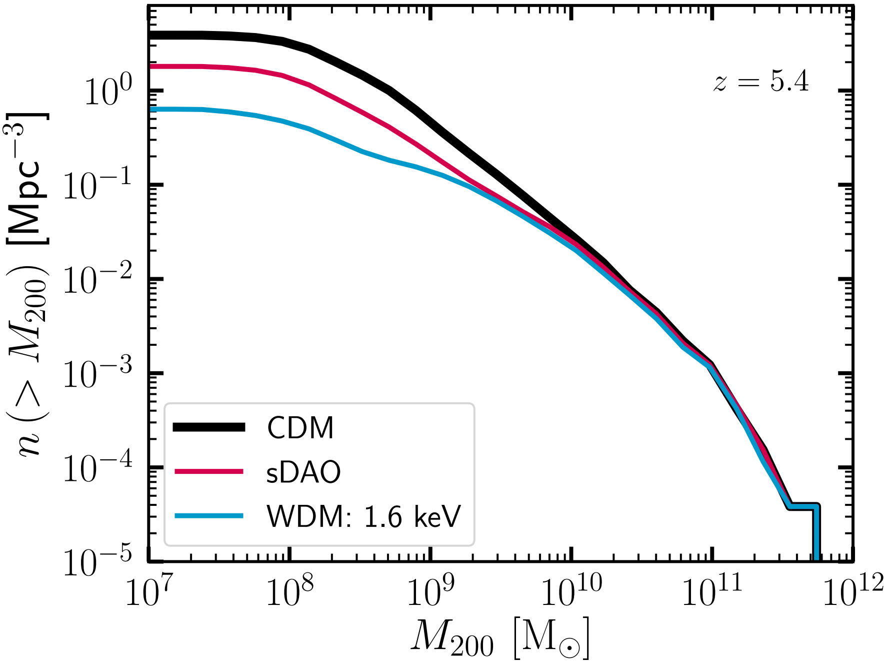

Figure 5.4.9. Comparison of the cumulative halo mass functions at $z = 5.4$ for the CDM, sDAO *and 1.6 keV thermal relic models. While both the* sDAO *and WDM models begin to deviate from CDM at a similar mass scale, there is more small-scale power in* sDAO. *The noticeable upturn at $M_{200} \sim 3 \times 10^{8}\,\mathrm{M}_\odot$ in the WDM mass function is the tell-tale signature of artificial fragmentation (Wang & White, 2007); this is largely absent in the* sDAO *model.*

scale than in the sDAO case, and where the first DAO peak is of lower amplitude than in sDAO and is pushed to smaller scales (see Fig. 5.3.1). The DAO feature in ETHOS-4 is thus unresolved by our simulation (the numerical setup was selected to just resolve the first sDAO peak). Regardless, this comparison highlights the potential of 1D flux spectrum measurements to distinguish not only non-CDM models from CDM, but also different non-CDM models from each other. The major constraining power comes from scales smaller than $k \sim 0.08$ s/km, where there is only limited data available at the moment (but see Boera et al., 2019, for newer data reaching to somewhat smaller scales).

One may be concerned that the DAO features we have identified in the $z = 5.4$ flux power spectrum may be affected by the small-scale noise manifest as the artificial peak at $k \sim 0.4$ s/km. To diagnose this, in the right panel of Fig. 5.4.10 we show the evolution of the flux power spectrum ratio from $z = 5.4$ to $z = 4.2$ for the sDAO and 1.6 keV WDM models. At $z = 5.4$, the DAO is very prominently present in the sDAO case while it is of course absent for the 1.6 keV model; on the other hand, the behaviour of the two models is almost identical by $z = 4.2$. This is consistent with the picture in

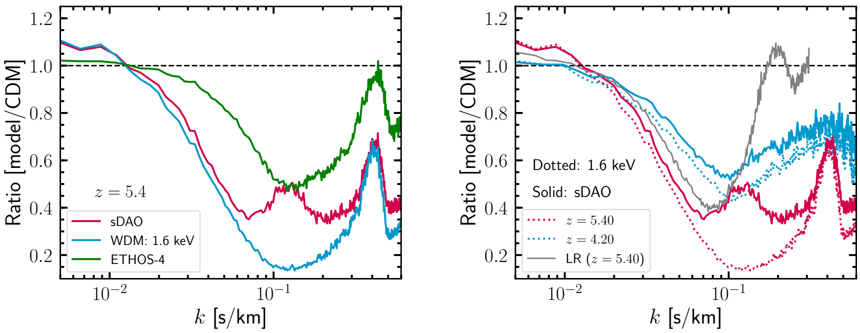

Figure 5.4.10. Left panel: As Fig. 5.4.8, but now comparing the sDAO *model with the 1.6 keV WDM and* ETHOS-4 *models at* $z = 5.4$. *The qualitative behaviour of each model is similar on scales larger than* $k = 0.1$ *s/km, in that power is suppressed relative to CDM. On smaller scales, the* sDAO *model exhibits a prominent bump induced by the first DAO peak. This feature is not observed in* ETHOS-4, *which also exhibits DAOs in the linear power spectrum, but of smaller amplitude than in the* sDAO *case. Each model has been run at the same resolution and each curve therefore exhibits the same numerical 'bump' at* $k = 0.4$ *s/km. Right panel: The evolution of the flux spectrum ratio from* $z = 5.4$ *to* $z = 4.2$ *for the* sDAO *(solid lines) and 1.6 keV WDM models (dotted lines). The* $z = 5.4$ *flux spectrum ratio in the low-resolution (LR)* sDAO *simulation is shown in grey. As time progresses, the overall increase of power causes adjacent modes to couple non-linearly, thus erasing any sharp (DAO) features in the power spectrum, until* $z = 4.2$ *where the behaviour of the* sDAO *and 1.6 keV WDM models is very similar.*

Fig. 5.4.8: the second DAO, which was visible at $z = 6$, is smoothed away by $z = 5.4$ due to non-linear mode coupling; similarly, the first DAO bump, which is visible at $z = 5.4$, is smeared away by $z = 4.2$. This is because the overall power across all scales increases towards lower redshift, giving the illusion of the DAO peaks being smeared with the numerical "noise peak" as time progresses. The effects of noise in the flux power spectrum are manifest more strongly in the 1.6 keV WDM case as there is a lack of "real" power on small-scales, in contrast to the sDAO model where the acoustic oscillation adds physical power on a level larger than the noise at $k > 0.1$ s/km.

How the noise level shifts as a function of resolution (see also Viel et al. 2013) may be evaluated by comparing the $z = 5.4$ flux spectra for the sDAO model at low and high resolution (grey and red curves, respectively). At the lower resolution, the numerical bump is shifted to larger scales by a factor of two (as expected, since the low-resolution simulation retains the same number of particles in a box that is twice as big as the high-resolution simulation). Moreover, the DAO bump, which just starts to develop, blends with the numerical bump and is therefore unresolved in the low-resolution simulation. With increased resolution (i.e., in our default simulations), the DAO is resolved before the noise becomes dominant. Thus, this figure reassures us that our physical interpretation of the first peak in the $z = 5.4$ flux spectrum for the sDAO model is not affected strongly by numerical systematics. As in the case of the cutoff in the small-scale flux spectrum, it may be that quantitative details in Fig. 5.4.10 are affected by assumptions made for the thermal history of the IGM. While varying these assumptions may certainly smear the prominence of the DAO feature, it is not clear that such bumps could be replicated by baryonic mechanisms. In particular, the scale at which these features are manifest, if induced by the nature of the DM, will be set by processes intrinsic to the DM model. We leave the detailed investigation of degeneracies between DAOs and thermal histories to future work.

5.5 Conclusions

We have performed detailed hydrodynamical simulations of non-standard dark matter (DM) species in which the DM is coupled to a relativistic component in the early universe. These interactions alter the primordial linear power spectrum predicted by the concordance cosmological model in a distinctive way: by generating a cutoff at the scale of dwarf galaxies through collisional damping, followed subsequently by a series of ‘dark acoustic oscillations’ (DAOs) towards smaller scales (see Fig. 5.3.1). Early structure formation in these models is therefore modified considerably from standard cold dark matter (CDM), principally in the form of a delay in the formation of the first stars, and a suppression in the abundance of low-mass galaxies (e.g. Lovell et al., 2018b). The structure of DM haloes may be modified as well through strong DM self-interactions at late-times that reshape the phase-space density profiles of galactic haloes (e.g. Vogelsberger et al., 2016). The extent to which these processes impact galaxy formation are, of course, sensitive to parameters specific to the DM theory, such as the duration of DM-radiation coupling, or the self-interaction cross-section.

While it is impossible to explore this parameter space fully, various permutations of these model parameters will predict largely similar galactic populations. The ETHOS framework (Cyr-Racine et al., 2016) provides a formalism for mapping these DM properties to ‘effective’ parameters that shape structure formation, thereby providing a flexible way to explore the implications of a vast range of theories on galaxy formation. In this paper, we focus our attention on an atomic DM model (which we refer to as sDAO) in which DM is composed of two massive fermions that are oppositely charged under a new unbroken $U(1)$ dark gauge force (see Section 5.2). The linear matter power spectrum of this model has a cutoff relative to CDM at $k \sim 10\,h\text{cMpc}^{-1}$, identical to a warm dark matter (WDM) thermal relic with mass 1.6 keV, but differs from WDM on smaller scales where it is composed of a significant number of undamped DAOs. While models as extreme as these may already be strongly constrained, our goal in this paper was to investigate if DAOs may be, in principle, detectable in the Lyman-alpha forest, rather than to present a model that matches the available data. *A priori*, it is not obvious that DAOs would persist in the Lyman-alpha flux spectrum. In particular, we sought to identify observational proxies that are able to distinguish between the different small-scale behaviour of these DAO models from WDM. For this purpose, we have investigated the statistics of the Lyman-α forest extracted from hydrodynamical simulations performed with these models using the AREPO code (Springel, 2010) coupled with a sophisticated galaxy formation model used as part of the IllustrisTNG project (Marinacci et al., 2018; Naiman et al., 2018; Nelson et al., 2018; Pillepich et al., 2018b; Springel et al., 2018).

Our main conclusions from the current study are:

1. On scales smaller than $k \sim 4h$ cMpc^{-1}, the 3D distribution of DM is clustered less strongly in the sDAO model than in CDM, although the differences get smaller with time (Fig. 5.3.3). In particular, while there is a strong DAO signature imprinted in the matter distribution at $z \geq 10$, further epochs of gravitational collapse wash away this feature entirely by $z \sim 6$.

2. A random line-of-sight through the sDAO simulation box reveals far less structure in absorption than the equivalent line-of-sight in the CDM simulation (Fig. 5.3.4). This is a direct consequence of the cutoff in the primordial power spectrum in the sDAO model, delaying the formation of galaxies at these high redshifts ($z \gtrsim 3$).

3. Despite the delayed start to the galaxy formation process in the sDAO model, it catches up with CDM by $z \approx 2$. This faster growth of structure is a fairly generic phenomenon observed in models with a cutoff in the linear power spectrum (including WDM). In our work, this is manifest in the form of the transmitted flux PDFs (Fig. 5.4.5), which are truncated towards high values in the sDAO model at $z \geq 4$, but are identical to CDM by $z = 2$. The probability that a given line-of-sight intersects a region with high transmitted flux increases as the universe transitions from neutral to ionised due to the ionising radiation from high redshift galaxies.

4. While the 1D flux power spectra are identical in CDM and in the sDAO model at $z \leq 3$, there are significant differences at higher redshift. In fact, present data at these redshifts already place the sDAO in significant tension with observations (Fig. 5.4.7), although astrophysical systematics may relax the level of discrepancy.

5. More interestingly, however, we find that the DAO bump characteristic of the sDAO model – which was absent in the 3D matter distribution – is imprinted prominently in the 1D flux power spectrum at $z \geq 5$ on scales smaller than $k \sim 0.1$ s/km (Fig. 5.4.8). At $z \leq 4.2$, the DAO feature is smoothed out, and the behaviour of the model is then reminiscent of standard WDM.

6. The appearance and disappearance of the DAO at different redshifts therefore offers an opportunity to disentangle small-scale features in the flux power spectrum induced by the nature of DM from astrophysical effects (e.g. different reionisation histories). In particular, precise measurements of the flux power spectrum on scales smaller than $k \approx 0.1$ s/km will be fundamental to distinguishing different DM models from each other (Fig. 5.4.10).

While there is a vast parameter space of well-motivated non-standard CDM models, the predictions they make for the formation of structure and the properties of galaxies can be challenging to differentiate. Of fundamental importance is the need to identify sets of statistics that allow the identification of *physical* scales that are characteristic of these theories. DM models in which there is a coupling to a relativistic species in the early Universe are characterised in the linear regime by a cutoff at the scale of dwarf galaxies followed by a series of dark acoustic oscillations towards smaller scales. In this work we have shown that these fundamental scales, while absent in the total matter distribution after the epoch of reionisation, are imprinted in the 1D Lyman-α flux power spectrum in a way that may be constrained with future high-precision observations. In the meantime, it is interesting to consider further statistics that could reveal the scale-dependent behaviour of different DM theories; possible examples include the clustering of DLAs and Lyman-limit systems or cross-correlations of Lyman-α with galaxy properties.

Acknowledgements: We thank the anonymous referee for providing suggestions that have improved this manuscript. We are very grateful to Volker Springel for allowing us access to AREPO, which was used to run all the simulations used in this paper. SB is supported by Harvard University through the ITC Fellowship. MV acknowledges support through an MIT RSC award, a Kavli Research Investment Fund, NASA ATP grant NNX17AG29G, and NSF grants AST-1814053 and AST-1814259. JZ and Sebastian Bohr acknowledge support by a Grant of Excellence from the Icelandic Research Fund (grant number 173929−051). CP acknowledges support by the European Research Council under ERC-CoG grant CRAGSMAN-646955. F-Y C-R acknowledges the support of the National Aeronautical and Space Administration ATP grant NNX16AI12G at Harvard University. This work was made possible in part by usage of computing resources at the University of Southern Denmark through the NeIC Dellingr resource sharing pilot. A fraction of the simulations in this work were carried out on the Garpur supercomputer, a joint project between the University of Iceland and University of Reykjavík with funding from Rannís.

5.6 Appendix

5.6.1 Resolution tests

In Fig. 5.6.11, we compare the $z = 5$ halo mass functions measured in the SDAO high-resolution (HR, 20 h^{-1} Mpc, $N_p = 2 \times 512^3$) and SDAO low-resolution (LR, 40 h^{-1} Mpc, $N_p = 2 \times 512^3$) simulations. The mass functions are converged over the expected range. The LR simulation shows an excess at the very massive end, as these rare haloes are more likely to be found in the larger volume of the LR simulation. It is interesting note, however, that neither simulation shows an upturn towards the low mass end, which is usually the characteristic signature of artificial halo formation in WDM simulations. This problem is exacerbated at low resolution – the scale below which these fragments start to dominate scales with the number of particles roughly as

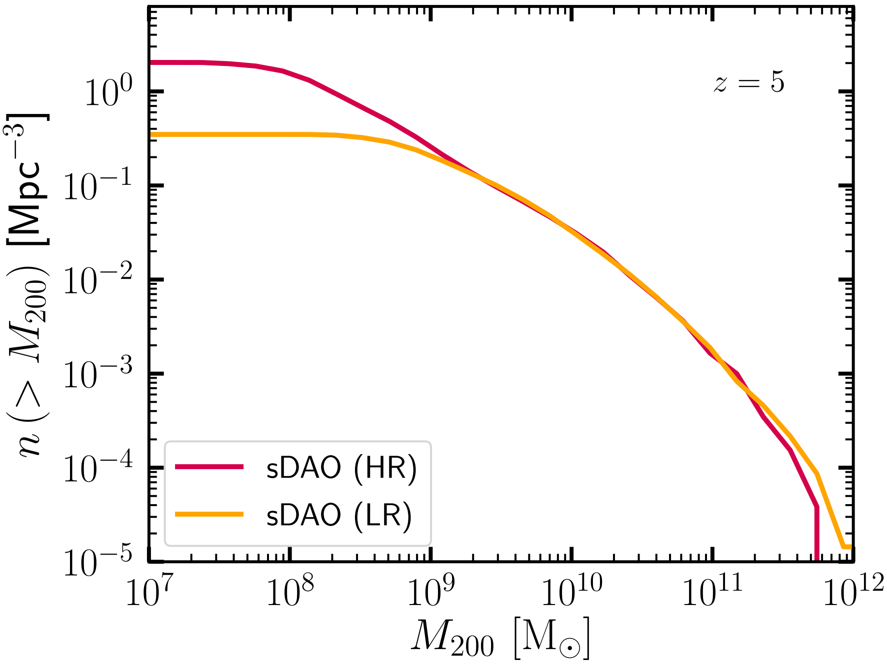

Figure 5.6.11. Comparison of the cumulative halo mass functions at $z = 5$ in the SDAO *LR (40 h^{-1}* Mpc*, $N_p = 2 \times 512^3$) and HR (20 h^{-1}* Mpc*, $N_p = 2 \times 512^3$) simulations. The excess at the high mass end in the LR simulation is due to the larger box size; similarly, the lack of low mass haloes is due to the lower resolution than in the HR version. Note that the characteristic upturn due to spurious fragmentation, which is clearly visible in WDM models (Wang & White, 2007) is no longer clear in the sDAO model at either resolution.*

$N_p^{1/3}$ (Wang & White, 2007). While this scale is clearly present in the 1.6 keV WDM simulation (see Fig. 5.4.9), the effect is largely suppressed in the SDAO model at both LR and HR. The reason for this is that while the cutoff in WDM continues indefinitely, the DAO adds power on scales smaller than the initial cutoff, thereby largely offsetting the instability through discreteness effects that is typical of WDM-like simulations.

Fig. 5.6.12 justifies our use of 1024 lines-of-sight for generating mock absorption spectra from our simulations. In this figure we show the ratio (SDAO/CDM) of the 1D flux power spectrum at $z = 5.4$ using 1024 (red) and 2048 (blue) lines-of-sight. The two curves show excellent convergence across all scales. We have checked explicitly that individual power spectra (rather than simply the ratio) are converged as well. We find that, in general, the truncated Gaussian smoothing kernel (Altay & Theuns, 2013, see also Eq. 5.2) is relatively robust to the number of skewers used to generate the mock spectra.

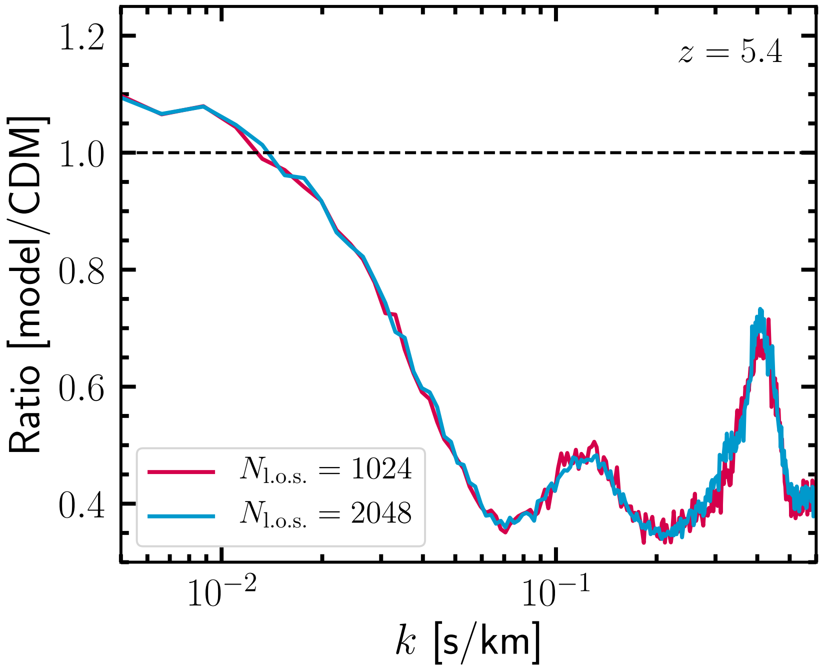

Figure 5.6.12. Demonstrating the convergence of the 1D Lyman-α flux spectrum ratio with respect to the number of lines-of-sight chosen for the analysis.

5.6.2 Dark matter model parameters

We list in Table 5.6.1 the atomic DM parameters used to generate our sDAO model. For more details on the model, see Cyr-Racine & Sigurdson (2013).

Parameter	Description	Value
ξ	$T_{\rm DR}/T_{\rm CMB}$	0.15
α_D	The dark fine structure constant	0.02
B_D	The dark atom binding energy	1.7 keV
$m_{\rm DM}$	The dark atom mass	500 MeV

Table 5.6.1. Atomic DM particle parameters used to generate our sDAO model. $T_{\rm DR}$ is the temperature of the dark radiation bath, while $T_{\rm CMB}$ is the cosmic microwave background temperature.

Chapter 6

Summary and Outlook

Summary. In this Thesis, I studied structure formation models alternative to CDM, particularly those that have a significant galactic-scale damping in the primordial power spectrum with dark acoustic oscillations (DAOs) in the DM-DR "plasma" (analogous to the baryonic acoustic oscillations in the photon-baryon plasma). My analysis starts from the analysis of density fluctuations in the DM-DR fluid using linear perturbation theory and the parametrization of the deviations of the linear power spectrum over CDM with two physically-motivated parameters (Chapter 2). It then continues into the non-linear regime of structure formation focusing on the high-redshift Universe (redshift range $20-5$) using $N-$body simulations with a special tailored application of the zoom-in technique to sample the non-linear power spectrum down to the scales relevant for galaxy formation (Chapter 2). In a general sense, my Thesis develops the ETHOS framework of structure formation to incorporate a self-consistent classification of structure formation models based on how distinct they are in the non-linear regime, but parametrized purely based on the features of their linear power spectrum. This allows for both a simple but general characterization of non-CDM structure formation models (covering both WDM and DAO models), as well as physically-motivated connection to the DM particle physics models behind them. Some of the highlights of my results are summarised as follows:

In **Chapter 2**, I have adapted the $N-$body zoom-in technique, traditionally used to re-simulate a halo of interest with higher resolution, to recover the average DM matter power spectrum across a large dynamical range. This technique reduced significantly the computational cost of the tailored simulations used in this Thesis, which achieve $8\times$ better resolution compared to uniform simulations with similar computational cost (see Table 2.2.1 and Fig. 2.2.3). With this technique, I was able to perform simulations that have a resolution good enough to probe scales up to $\sim 500\,h\,\mathrm{Mpc}^{-1}$(corresponding to a halo mass of $\sim 10^7$ $\mathrm{M}_\odot$), making it possible to perform a suite of 52 simulations to explore a large class of DAO models within the ETHOS framework in the high-redshift non-linear regime ($5 < z < 20$) and down to relevant scales for galaxy formation. This is a key point of this study, by concentrating on scales at and above the galaxy formation edge, we are able to classify models that are truly distinct to CDM at scales relevant to the physics of galaxies and in a potentially observable way. Models that deviate from

CDM at smaller scales might posses a different DM nature, but one that is irrelevant (and hidden from) the physics of galaxies.

The simulations cover the relevant parameter space of two newly introduced effective parameters $h_{\rm peak}$ and $k_{\rm peak}$, the amplitude and scale of the first DAO peak (see Eq. 2.3). These two physically motivated parameters are sufficient to fully describe the linear matter power spectrum of the ETHOS models described in Cyr-Racine et al. (2016). In the limit of $k_{\rm peak} \to \infty$, the parametrization also covers CDM and for $h_{\rm peak} = 0$, it covers WDM models. The DAO scale $k_{\rm peak}$ is connected to the time of DM decoupling from DR (Fig. 2.3.6), while the amplitude depends on the opacity at the time of decoupling (Fig. 2.3.7). The two new parameters $h_{\rm peak}$ and $k_{\rm peak}$ can be connected to the original ETHOS parameters in Cyr-Racine et al. (2016) and thus, ultimately to particle physics parameters.

To better distinguish different models, I used two statistical measures. First, the ratio of the matter power spectrum of a given ETHOS model with respect to CDM at a specific scale measures the relative difference with CDM. Second, the integrated power down to the same scale can be used to compare different models and identify possible degeneracies between them. Remarkably, despite the large difference seen between DAO models and WDM in the linear power spectrum (given the clear presence of acoustic oscillations in the former), we found that only a small area of DAO models in the parameter space (those with strong DAO features: $h_{\rm peak} \gtrsim 0.7$, $k_{\rm peak} \lesssim 60\,h\,{\rm Mpc}^{-1}$) are actually distinguishable from WDM in the non-linear matter power spectrum on scales above the atomic cooling limit by redshift five (see Fig. 2.3.10). The halo mass function however, is better capable of discriminating between DAO and WDM models (see Fig. 2.4.14), as the slope of the mass function at small masses is sensitive to $h_{\rm peak}$ even for models with weak DAO features (wDAOs). When the DAO peak is at a sufficiently large scale ($k_{\rm peak} < 100\,h\,{\rm Mpc}^{-1}$), the DAO models are distinguishable from WDM regardless of their strength (characterized by $h_{\rm peak}$).

In **Chapter 3**, a closer look is taken at the abundance and structure of DM haloes in the ETHOS simulation suite presented in Chapter 2. One of the highlights is that I show the success of the extended Press-Schechter formalism (EPS) for DAO models in predicting the halo mass function from $z \sim 20$ down to $z = 5$. For sDAO models, the EPS formalism can reproduce the halo mass function as accurately as CDM (see Fig. 3.3.5). For wDAO and WDM models, the accuracy is not as good, but the trend of the halo mass function can still be reproduced (see Fig. 3.3.3). This result offers a powerful way to easily, but accurately, predict the halo mass function of a specific model without the need of a computationally expensive $N-$body simulation. As for the structure of haloes in DAO models, I show that they are well described by the NFW profile. In this regard they are not different to the WDM and CDM cases, however the structure parameter of the profile, namely the concentration, has a scaling with mass that is worthy to highlight. The concentration of low mass DAO haloes is suppressed relative to CDM: in trend, similar to how WDM haloes have lower concentration, but in detail, the DAO models have a distinct behaviour, with the oscillations leaving a trace/impact in the concentration-mass relation. The concentration-mass relation can be reproduced (up to the statistical sampling limitations in the simulations) by the model of Ludlow et al. (2016) (see Fig. 3.4.9). Simulations with better statistics are however needed to reduce the sampling errors and check the $k_{\rm peak}$ scaling of the halo mass function in more detail. This would allow both, to characterize the signatures of DAOs in the concentration-mass relation, and to confirm whether the Ludlow et al. (2016) model is indeed accurate enough to reproduce these features.

In **Chapter 4**, we have shown that the 21-cm signal in the cosmic-dawn era ($z = 10 - 25$) is strongly affected by DAO models (both the global signal and the fluctuations power spectrum). The delayed halo formation of DAO models leads to a (global) 21-cm signal that is both different in shape and delayed relative to CDM (see Fig. 4.3.3). Based on our results, it seems that the shape of the sDAO models cannot be mimicked in CDM by baryonic physics (e.g. stellar feedback) nor by WDM (see Fig. 4.3.6-4.3.7). These results offer a promising way to probe the DM parameter space of structure formation models with upcoming observations such as that from HERA (DeBoer et al., 2017), which could potentially differentiate ETHOS models from CDM. This study only included stellar feedback in the case of CDM, and further studies varying more astrophysical parameters are needed to test for possible degeneracies between dark and baryonic physics.

In another observational consequence of DAOs, we found in **Chapter 5** that the 1D flux power spectrum that is measured by Lyman-α forest observations can retain the DAO features until lower redshifts than the 3D matter power spectrum. Therefore, the Lyman-α forest inferences of the distribution of matter at smaller scales remain a promising way to constrain the allowed parameter space for DM physics, even for DAO models. And in fact, despite the fact that the DAO features tend to be erased by non-linear evolution, and entangled with baryonic physics mechanisms, our results promisingly show that distinct DAO signatures might remain in the 1D flux power spectrum for strong DAO models (sDAOs, see Fig. 5.4.10). As only one sDAO model was studied in Chapter 5, it will be relevant in the future to perform a full Lyman-α pipeline analysis covering a larger area of the ($h_{\rm peak}$,$k_{\rm peak}$) parameter space. Even if weaker models do not show an imprint of their DAO peak anymore, it is important to see if they remain distinguishable from WDM, and what are the current constraints for these models exactly. The latter is something to highlight since current constraints on DAO models are still based on DM-only simulations or guessed based on their similarity to WDM models. My results in Chapter 2 actually allow for the closest equivalence for weak DAO models and WDM models (and thus to their constraints), but it is clear that full mock Lyman-α forest simulations are needed to obtain more accurate constraints.

Outlook. The simulations done in this Thesis only go down to $z \sim 5$, and thus, a natural prospect in the future would be to continue our simulation suite within the (new) ETHOS parameter space down into the low-redshift regime, ultimately all the way to $z = 0$. Although this is more computationally demanding due to both, the larger (cosmic) integration time and much higher degree of clustering at lower redshift, it is certainly achievable. Such a suite would allow for a full analysis of the impact of DAOs across cosmic time in statistics such as the non-linear power spectrum, halo mass function, and inner halo structure. For all of this, my results at high redshift offer a template that would be relatively straightforward to follow. However, at lower redshifts there are two crucial elements that I have (almost) not considered thus far. Firstly, baryonic physics plays a relevant role in setting the minimum scale for galaxy formation and it is ever more important in shaping the structure of dark matter haloes as time goes by. It is thus quite important that our future studies incorporate those physical mechanism that are likely to create an interplay with the primordial DAO cutoff (this could already be quite relevant at the high redshift studied in this Thesis, but even more important at low redshift). For instance, one crucial mechanism at high redshift is the impact of reionization, when neutral hydrogen gets ionized by the UV radiation from early stars and galaxies. This mechanism heats up the intergalactic gas and suppresses galaxy formation (see e.g. Gnedin, 2000; Sawala et al., 2016b). At lower redshifts, the feedback of supernovae pushes matter out of the center of galaxies and therefore changes the inner DM distribution within haloes (see e.g. Pontzen & Governato, 2012). The second crucial element is additional DM physics in the form of DM self-interactions, which is inherent to most DAO models, and plays a relevant role at low redshifts by transforming the inner structure of DM haloes substantially (see e.g. Spergel & Steinhardt, 2000; Vogelsberger et al., 2012).

Beyond these aspects, the parametrization I developed offers a template that can be used in the future to study ETHOS models in many different studies with varying contexts and aiming at different observables. A full Lyman-α analysis of the parameter space and comparison with current observations would offer better constraints on the ETHOS models than the current ones deduced from constraints on WDM. In the near future, new observing facilities will also offer different approaches to constraint the allowed models. The luminosity function at low masses and at high redshifts will be observed by the James Webb telescope (JWST) setting new (indirect) constraints on the halo mass function. The 21-cm signal will be observed with better precision by the upcoming hydrogen epoch-of-reionization array (HERA), offering another promising approach to differentiate ETHOS models.

Bibliography

Abbott T. M. C., et al., 2018, Phys. Rev., D98, 043526

Ackerman L., Buckley M. R., Carroll S. M., Kamionkowski M., 2009, Phys. Rev. D, 79, 023519

Adhikari R., et al., 2017, J. Cosmology Astropart. Phys., 2017, 025

Albert A., et al., 2017, ApJ, 834, 110

Altay G., Theuns T., 2013, MNRAS, 434, 748

Anderson L., et al., 2012, MNRAS, 427, 3435

Ando S., Kamada A., Sekiguchi T., Takahashi T., 2019, Phys. Rev. D, 100, 123519

Aprile E., et al., 2018, Physical Review Letters, 121, 111302

Arcadi G., Dutra M., Ghosh P., Lindner M., Mambrini Y., Pierre M., Profumo S., Queiroz F. S., 2018, European Physical Journal C, 78, 203

Archidiacono M., Hannestad S., Hansen R. S., Tram T., 2015, Phys. Rev. D, 91, 065021

Archidiacono M., Bohr S., Hannestad S., Helboe Jørgensen J., Lesgourgues J., 2017, J. Cosmology Astropart. Phys., 2017, 010

Archidiacono M., Hooper D. C., Murgia R., Bohr S., Lesgourgues J., Viel M., 2019, J. Cosmology Astropart. Phys., 2019, 055

Arkani-Hamed N., Finkbeiner D. P., Slatyer T. R., Weiner N., 2009, Phys. Rev. D, 79, 015014

Avila-Reese V., Colín P., Valenzuela O., D'Onghia E., Firmani C., 2001, ApJ, 559, 516

Baldi M., 2013, MNRAS, 428, 2074

Barkana R., Loeb A., 2001, Phys. Rep., 349, 125

Barnes J., Hut P., 1986, Nature, 324, 446

Bartels R., Krishnamurthy S., Weniger C., 2016, Phys. Rev. Lett., 116, 051102

Baur J., Palanque-Delabrouille N., Yèche C., Magneville C., Viel M., 2016, J. Cosmology Astropart. Phys., 8, 012

Beardsley A. P., et al., 2016, Astrophys. J., 833, 102

Behbahani S. R., Jankowiak M., Rube T., Wacker J. G., 2011, Adv. High Energy Phys., 2011, 709492

Benitez-Llambay A., 2015, py-sphviewer: Py-SPHViewer v1.0.0, doi:10.5281/zenodo.21703, `http://dx.doi.org/10.5281/zenodo.21703`

Bertone G., Hooper D., Silk J., 2005, Phys. Rept., 405, 279

Blas D., Lesgourgues J., Tram T., 2011, J. Cosmology Astropart. Phys., 2011, 034

Blumenthal G. R., Pagels H., Primack J. R., 1982, Nature, 299, 37

Blumenthal G. R., Faber S., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Bode P., Ostriker J. P., Turok N., 2001, ApJ, 556, 93

Bœhm C., Schaeffer R., 2005, A&A, 438, 419

Bœhm C., Riazuelo A., Hansen S. H., Schaeffer R., 2002, Phys. Rev. D, 66, 083505

Bœhm C., Schewtschenko J. A., Wilkinson R. J., Baugh C. M., Pascoli S., 2014, MNRAS, 445, L31

Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101

Bohr S., Zavala J., Cyr-Racine F.-Y., Vogelsberger M., Bringmann T., Pfrommer C., 2020, MNRAS, 498, 3403

Bolton J. S., Haehnelt M. G., Viel M., Springel V., 2005, MNRAS, 357, 1178

Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897

Bond J. R., Szalay A. S., 1983, ApJ, 274, 443

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Bose S., Hellwing W. A., Li B., 2015, J. Cosmology Astropart. Phys., 2, 034

Bose S., Hellwing W. A., Frenk C. S., Jenkins A., Lovell M. R., Helly J. C., Li B., 2016a, MNRAS, 455, 318

Bose S., Frenk C. S., Hou J., Lacey C. G., Lovell M. R., 2016b, MNRAS, 463, 3848

Bose S., et al., 2017, MNRAS, 464, 4520

Bose S., Vogelsberger M., Zavala J., Pfrommer C., Cyr-Racine F.-Y., Bohr S., Bringmann T., 2019, MNRAS, 487, 522

Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67

Boyarsky A., Drewes M., Lasserre T., Mertens S., Ruchayskiy O., 2019, Progress in Particle and Nuclear Physics, 104, 1

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, MNRAS, 422, 1203

Bozek B., et al., 2019, MNRAS, 483, 4086

Brinckmann T., Zavala J., Rapetti D., Hansen S. H., Vogelsberger M., 2018, MNRAS, 474, 746

Bringmann T., 2009, New Journal of Physics, 11, 105027

Bringmann T., Hasenkamp J., Kersten J., 2014, J. Cosmology Astropart. Phys., 1407, 042

Bringmann T., Ihle H. T., Kersten J., Walia P., 2016, Phys. Rev., D94, 103529

Buckley M. R., Peter A. H. G., 2018, Phys. Rep., 761, 1

Buckley M. R., Zavala J., Cyr-Racine F.-Y., Sigurdson K., Vogelsberger M., 2014, Phys. Rev. D, 90, 043524

Buen-Abad M. A., Marques-Tavares G., Schmaltz M., 2015, Phys. Rev. D, 92, 023531

Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559

Burger J. D., Zavala J., 2021, arXiv e-prints, p. arXiv:2103.01231

Campbell D. J. R., et al., 2017, MNRAS, 469, 2335

Carlson E. D., Machacek M. E., Hall L. J., 1992, ApJ, 398, 43

Carrasco J. J. M., Hertzberg M. P., Senatore L., 2012, Journal of High Energy Physics, 2012, 82

Chacko Z., Cui Y., Hong S., Okui T., Tsai Y., 2016, JHEP, 12, 108

Chauhan G., Lagos C. d. P., Obreschkow D., Power C., Oman K., Elahi P. J., 2019, MNRAS, 488, 5898

Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, J. Cosmology Astropart. Phys., 5, 048

Choquette J., Cline J. M., 2015, Phys. Rev. D, 92, 115011

Chu X., Dasgupta B., 2014, Phys. Rev. Lett., 113, 161301

Cline J. M., Liu Z., Xue W., 2012, Phys. Rev., D85, 101302

Cline J. M., Liu Z., Moore G., Xue W., 2014a, Phys. Rev. D, 89, 043514

Cline J. M., Liu Z., Moore G., Xue W., 2014b, Phys. Rev. D, 90, 015023

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109

Cole S., et al., 2005, MNRAS, 362, 505

Colín P., Avila-Reese V., Valenzuela O., 2000, ApJ, 542, 622

Colín P., Avila-Reese V., Valenzuela O., Firmani C., 2002, ApJ, 581, 777

Colless M., et al., 2001, MNRAS, 328, 1039

Creasey P., Sameie O., Sales L. V., Yu H.-B., Vogelsberger M., Zavala J., 2017, MNRAS, 468, 2283

Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44

Croft R. A. C., Weinberg D. H., Pettini M., Hernquist L., Katz N., 1999, ApJ, 520, 1

Cyr-Racine F.-Y., Sigurdson K., 2013, Phys. Rev. D, 87, 103515

Cyr-Racine F.-Y., de Putter R., Raccanelli A., Sigurdson K., 2014, Phys. Rev., D89, 063517

Cyr-Racine F.-Y., Sigurdson K., Zavala J., Bringmann T., Vogelsberger M., Pfrommer C., 2016, Phys. Rev. D, 93, 123527

Dalal N., Kochanek C. S., 2002, Astrophys. J., 572, 25

Das S., Sigurdson K., 2012, Phys. Rev. D, 85, 063510

Das S., Mondal R., Rentala V., Suresh S., 2018, JCAP, 1808, 045

Datta A., Bowman J. D., Carilli C. L., 2010, Astrophys. J., 724, 526

Davé R., Spergel D. N., Steinhardt P. J., Wandelt B. D., 2001, ApJ, 547, 574

Davis M., Lecar M., Pryor C., Witten E., 1981, ApJ, 250, 423

Davis M., Efstathiou G., Frenk C. S., White S. D., 1985, ApJ, 292, 371

Daylan T., Finkbeiner D. P., Hooper D., Linden T., Portillo S. K. N., Rodd N. L., Slatyer T. R., 2016, Physics of the Dark Universe, 12, 1

DeBoer D. R., et al., 2017, Publ. Astron. Soc. Pac., 129, 045001

Dehnen W., Read J. I., 2011, European Physical Journal Plus, 126, 55

Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, MNRAS, 437, 415

DiLullo C., Taylor G. B., Dowell J., 2020, J. Astron. Inst., 09, 2050008

Diamanti R., Giusarma E., Mena O., Archidiacono M., Melchiorri A., 2013, Phys. Rev. D, 87, 063509

Díaz Rivero A., Dvorkin C., Cyr-Racine F.-Y., Zavala J., Vogelsberger M., 2018, Phys. Rev. D, 98, 103517

Dijkstra M., Haiman Z., Rees M. J., Weinberg D. H., 2004, Astrophys. J., 601, 666

Dodelson S., 2003, Modern Cosmology. Academic Press, Amsterdam

Dooley G. A., Peter A. H. G., Vogelsberger M., Zavala J., Frebel A., 2016, MNRAS, 461, 710

Drlica-Wagner A., et al., 2019, arXiv e-prints, p. arXiv:1902.01055

Dubois Y., et al., 2014, MNRAS, 444, 1453

Dutton A. A., Obreja A., Macciò A. V., 2019, MNRAS, 482, 5606

Eastwood M. W., et al., 2019, Astron. J., 158, 84

Efstathiou G., 1992, MNRAS, 256, 43P

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114

Elbert O. D., Bullock J. S., Garrison-Kimmel S., Rocha M., Oñorbe J., Peter A. H. G., 2015, MNRAS, 453, 29

Escudero M., Lopez-Honorez L., Mena O., Palomares-Ruiz S., Villanueva-Domingo P., 2018, JCAP, 1806, 007

Fan J., Katz A., Randall L., Reece M., 2013a, Phys. Dark Univ., 2, 139

Fan J. J., Katz A., Randall L., Reece M., 2013b, Phys. Rev. Lett., 110, 211302

Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, ApJ, 703, 1416

Feng J. L., Kaplinghat M., Tu H., Yu H.-B., 2009, J. Cosmology Astropart. Phys., 0907, 004

Feng J. L., Kamionkowski M., Lee S. K., 2010, Phys. Rev. D, 82, 015012

Fermi-LAT Collaboration 2017, preprint, (`arXiv:1705.00009`)

Field G. B., 1959, ApJ, 129, 536

Fitts A., et al., 2017, MNRAS, 471, 3547

Flores R. A., Primack J. R., 1994, ApJ, 427, L1

Foot R., 2004, Int. J. Mod. Phys. D, 13, 2161

Furlanetto S., Oh S. P., Briggs F., 2006, Phys. Rept., 433, 181

Garrison-Kimmel S., Boylan-Kolchin M., Bullock J. S., Kirby E. N., 2014, MNRAS, 444, 222

Garrison-Kimmel S., et al., 2019, MNRAS, 487, 1380

Garzilli A., Boyarsky A., Ruchayskiy O., 2017, Physics Letters B, 773, 258

Garzilli A., Ruchayskiy O., Magalich A., Boyarsky A., 2019a, arXiv e-prints, p. arXiv:1912.09397

Garzilli A., Magalich A., Theuns T., Frenk C. S., Weniger C., Ruchayskiy O., Boyarsky A., 2019b, MNRAS, 489, 3456

Geller M. J., Huchra J. P., 1989, Science, 246, 897

Gillet N. J., Mesinger A., Park J., 2020, Mon. Not. Roy. Astron. Soc., 491, 1980

Gilman D., Birrer S., Treu T., Nierenberg A., Benson A., 2019, Mon. Not. Roy. Astron. Soc., 487, 5721

Gilman D., Birrer S., Nierenberg A., Treu T., Du X., Benson A., 2020, Mon. Not. Roy. Astron. Soc., 491, 6077

Gnedin N. Y., 2000, ApJ, 542, 535

Gnedin N. Y., Hui L., 1998, MNRAS, 296, 44

Gott J. Richard I., Jurić M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, ApJ, 624, 463

Greig B., Mesinger A., 2015, Mon. Not. Roy. Astron. Soc., 449, 4246

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Hahn O., Abel T., 2011, MNRAS, 415, 2101

Hildebrandt H., et al., 2017, MNRAS, 465, 1454
Hirata C. M., 2006, Mon. Not. Roy. Astron. Soc., 367, 259
Hoekstra H., Franx M., Kuijken K., Squires G., 1998, ApJ, 504, 636
Hooper D., Goodenough L., 2011, Physics Letters B, 697, 412
Hooper D., Weiner N., Xue W., 2012, Phys. Rev. D, 86, 056009
Hsueh J.-W., Enzi W., Vegetti S., Auger M., Fassnacht C. D., Despali G., Koopmans L. V., McKean J. P., 2020, Mon. Not. Roy. Astron. Soc., 492, 3047
Hu W., Sugiyama N., 1996, ApJ, 471, 542
Hui L., Ostriker J. P., Tremaine S., Witten E., 2017, Phys. Rev. D, 95, 043541
Ibata N. G., Ibata R. A., Famaey B., Lewis G. F., 2014, Nature, 511, 563
Iršič V., et al., 2017, MNRAS, 466, 4332
Iršič V., et al., 2017, Phys. Rev. D, 96, 023522
Ishiyama T., et al., 2020, arXiv e-prints, p. arXiv:2007.14720
Jenkins A., 2010, MNRAS, 403, 1859
Jennings E., Baugh C. M., Li B., Zhao G.-B., Koyama K., 2012, MNRAS, 425, 2128
Jungman G., Kamionkowski M., Griest K., 1996, Phys. Rept., 267, 195
Kamada A., Takahashi T., 2018, J. Cosmology Astropart. Phys., 2018, 047
Kaplan D. E., Krnjaic G. Z., Rehermann K. R., Wells C. M., 2010, JCAP, 1005, 021
Kaplan D. E., Krnjaic G. Z., Rehermann K. R., Wells C. M., 2011, J. Cosmology Astropart. Phys., 1110, 011
Kaplinghat M., Tulin S., Yu H.-B., 2016, Physical Review Letters, 116, 041302
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kennedy R., Frenk C., Cole S., Benson A., 2014, MNRAS, 442, 2487
Kim T.-S., Viel M., Haehnelt M. G., Carswell R. F., Cristiani S., 2004, MNRAS, 347, 355
Kim S. Y., Peter A. H. G., Hargis J. R., 2018, Phys. Rev. Lett., 121, 211302
Klypin A. A., Shandarin S. F., 1983, MNRAS, 204, 891
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Klypin A., Karachentsev I., Makarov D., Nasonova O., 2015, MNRAS, 454, 1798
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Kobayashi T., Murgia R., De Simone A., Iršič V., Viel M., 2017, Phys. Rev. D, 96, 123514
Koopmans L. V. E., et al., 2015, PoS, AASKA14, 001
Koposov S., et al., 2008, ApJ, 686, 279
Kulkarni G., Hennawi J. F., Oñorbe J., Rorai A., Springel V., 2015, ApJ, 812, 30
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lee S. K., Lisanti M., Safdi B. R., Slatyer T. R., Xue W., 2016, Physical Review Letters, 116, 051103
Leo M., Baugh C. M., Li B., Pascoli S., 2018a, J. Cosmology Astropart. Phys., 2018, 010
Leo M., Baugh C. M., Li B., Pascoli S., 2018b, J. Cosmology Astropart. Phys., 2018, 001
Lesgourgues J., Marques-Tavares G., Schmaltz M., 2016, J. Cosmology Astropart. Phys., 1602, 037
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Li Y., Hu W., Takada M., 2014, Phys. Rev. D, 89, 083519

Liu A., Tegmark M., Bowman J., Hewitt J., Zaldarriaga M., 2009, Mon. Not. Roy. Astron. Soc., 398, 401

Lopez-Honorez L., Mena O., Moliné A., Palomares-Ruiz S., Vincent A. C., 2016, JCAP, 1608, 004

Lopez-Honorez L., Mena O., Villanueva-Domingo P., 2019, Phys. Rev., D99, 023522

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

Lovell M. R., et al., 2018a, MNRAS, 477, 2886

Lovell M. R., et al., 2018b, Mon. Not. Roy. Astron. Soc., 477, 2886

Lovell M., Zavala J., Vogelsberger M., 2019, Mon. Not. Roy. Astron. Soc., 485, 5474

Ludlow A. D., Navarro J. F., Angulo R. E., Boylan-Kolchin M., Springel V., Frenk C., White S. D. M., 2014, MNRAS, 441, 378

Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214

Ma C.-P., Bertschinger E., 1995, ApJ, 455, 7

Macciò A. V., Fontanot F., 2010, MNRAS, 404, L16

Macias O., Gordon C., Crocker R. M., Coleman B., Paterson D., Horiuchi S., Pohl M., 2018, Nature Astronomy, 2, 387

Marinacci F., et al., 2018, MNRAS, 480, 5113

Mason C., Trenti M., Treu T., 2015, Astrophys. J., 813, 21

McCullough M., Randall L., 2013, J. Cosmology Astropart. Phys., 1310, 058

McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1

Mena O., Palomares-Ruiz S., Villanueva-Domingo P., Witte S. J., 2019, Phys. Rev., D100, 043540

Mesinger A., Furlanetto S., Cen R., 2011, Mon. Not. Roy. Astron. Soc., 411, 955

Mocz P., Vogelsberger M., Robles V. H., Zavala J., Boylan-Kolchin M., Fialkov A., Hernquist L., 2017, MNRAS, 471, 4559

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, Astrophys. J., 752, 137

Muñoz J. B., Cyr-Racine F.-Y., 2021, Phys. Rev. D, 103, 023512

Muñoz J. B., Loeb A., 2018, arXiv e-prints, p. arXiv:1802.10094

Muñoz J. B., Bohr S., Cyr-Racine F.-Y., Zavala J., Vogelsberger M., 2021, Phys. Rev. D, 103, 043512

Muñoz J. B., 2019a, Phys. Rev. Lett., 123, 131301

Muñoz J. B., 2019b, Phys. Rev., D100, 063538

Muñoz J. B., Dvorkin C., Loeb A., 2018, Phys. Rev. Lett., 121, 121301

Muñoz J. B., Dvorkin C., Cyr-Racine F.-Y., 2020, Phys. Rev. D, 101, 063526

Murgia R., Merle A., Viel M., Totzauer M., Schneider A., 2017, J. Cosmology Astropart. Phys., 11, 046

Murgia R., Iršič V., Viel M., 2018, Phys. Rev. D, 98, 083540

Nadler E. O., Gluscevic V., Boddy K. K., Wechsler R. H., 2019, Astrophys. J., 878, L32

Nadler E. O., et al., 2021, Phys. Rev. Lett., 126, 091101

Naiman J. P., et al., 2018, MNRAS, 477, 1206

Nasir F., Bolton J. S., Becker G. D., 2016, MNRAS, 463, 2335
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Navarro J. F., et al., 2010, MNRAS, 402, 21
Nelson D., et al., 2018, MNRAS, 477, 450
Neto A. F., et al., 2007, MNRAS, 381, 1450
Nori M., Murgia R., Iršič V., Baldi M., Viel M., 2019, MNRAS, 482, 3227
Oñorbe J., Garrison-Kimmel S., Maller A. H., Bullock J. S., Rocha M., Hahn O., 2014, MNRAS, 437, 1894
Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017, ApJ, 847, 63
Oh S. P., Haiman Z., 2002, Astrophys. J., 569, 558
Oman K. A., et al., 2015, MNRAS, 452, 3650
Oman K. A., Marasco A., Navarro J. F., Frenk C. S., Schaye J., Benítez-Llambay A., 2019, MNRAS, 482, 821
Orosz N., Dillon J. S., Ewall-Wice A., Parsons A. R., Thyagarajan N., 2019, Mon. Not. Roy. Astron. Soc., 487, 537
Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, MNRAS, 443, 678
Pagels H., Primack J. R., 1982, Phys. Rev. Lett., 48, 223
Palanque-Delabrouille N., et al., 2013, A&A, 559, A85
Papastergis E., Martin A. M., Giovanelli R., Haynes M. P., 2011, ApJ, 739, 38
Papastergis E., Giovanelli R., Haynes M. P., Shankar F., 2015, A&A, 574, A113
Park J., Mesinger A., Greig B., Gillet N., 2019, Mon. Not. Roy. Astron. Soc., 484, 933
Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, Astrophys. J., 756, 165
Pawlowski M. S., Kroupa P., Jerjen H., 2013, MNRAS, 435, 1928
Pawlowski M. S., et al., 2014, MNRAS, 442, 2362
Peñarrubia J., Pontzen A., Walker M. G., Koposov S. E., 2012, ApJ, 759, L42
Peebles P. J. E., 1984, ApJ, 277, 470
Peeples M. S., Weinberg D. H., Davé R., Fardal M. A., Katz N., 2010, MNRAS, 404, 1281
Philip L., et al., 2019, Journal of Astronomical Instrumentation, 8, 1950004
Pillepich A., et al., 2018a, MNRAS, 473, 4077
Pillepich A., et al., 2018b, MNRAS, 475, 648
Planck Collaboration et al., 2016, A&A, 594, A13
Planck Collaboration et al., 2020, A&A, 641, A6
Pober J. C., et al., 2013, The Astronomical Journal, 145, 65
Pober J. C., et al., 2014, Astrophys. J., 782, 66
Pontzen A., Governato F., 2012, MNRAS, 421, 3464
Pontzen A., Governato F., 2014, Nature, 506, 171
Pontzen A., Roškar R., Stinson G. S., Woods R., Reed D. M., Coles J., Quinn T. R., 2013, pynbody: Astrophysics Simulation Analysis for Python
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Preskill J., Wise M. B., Wilczek F., 1983, Physics Letters B, 120, 127
Press W. H., Schechter P., 1974, ApJ, 187, 425
Price D. C., et al., 2018, MNRAS, 478, 4193

Pritchard J. R., Furlanetto S. R., 2007, MNRAS, 376, 1680

Pritchard J. R., Loeb A., 2008, Phys. Rev. D, 78, 103511

Pritchard J. R., Loeb A., 2012, Rept. Prog. Phys., 75, 086901

Qin Y., Mesinger A., Park J., Greig B., Muñoz J. B., 2020, Mon. Not. Roy. Astron. Soc., 495, 123

Qin Y., Mesinger A., Greig B., Park J., 2021, MNRAS, 501, 4748

Randall L., Scholtz J., 2015, J. Cosmology Astropart. Phys., 1509, 057

Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173

Rauch M., et al., 1997, ApJ, 489, 7

Read J. I., Iorio G., Agertz O., Fraternali F., 2017, MNRAS, 467, 2019

Robertson A., et al., 2018, MNRAS, 476, L20

Robles V. H., Matos T., 2012, MNRAS, 422, 282

Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, MNRAS, 430, 81

Rorai A., Hennawi J. F., White M., 2013, ApJ, 775, 81

Roszkowski L., Sessolo E. M., Trojanowski S., 2018, Reports on Progress in Physics, 81, 066201

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471

Sameie O., Creasey P., Yu H.-B., Sales L. V., Vogelsberger M., Zavala J., 2018, MNRAS, 479, 359

Sameie O., Benson A. J., Sales L. V., Yu H.-b., Moustakas L. A., Creasey P., 2019, ApJ, 874, 101

Sánchez-Conde M. A., Prada F., 2014, MNRAS, 442, 2271

Sawala T., et al., 2016a, MNRAS, 456, 85

Sawala T., et al., 2016b, MNRAS, 457, 1931

Schaye J., et al., 2015, MNRAS, 446, 521

Schewtschenko J. A., Wilkinson R. J., Baugh C. M., Bœhm C., Pascoli S., 2015, MNRAS, 449, 3587

Schneider A., 2018, Phys. Rev., D98, 063021

Schneider A., Smith R. E., Macciò A. V., Moore B., 2012, MNRAS, 424, 684

Schneider A., Smith R. E., Reed D., 2013, MNRAS, 433, 1573

Schneider A., Trujillo-Gomez S., Papastergis E., Reed D. S., Lake G., 2017, MNRAS, 470, 1542

Seljak U., Makarov A., McDonald P., Trac H., 2006, Physical Review Letters, 97, 191303

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Silk J., 1968, Astrophys. J., 151, 459

Singh S., Subrahmanyan R., Shankar N. U., Rao M. S., Girish B. S., Raghunathan A., Somashekar R., Srivani K. S., 2018, Experimental Astronomy, 45, 269

Sitwell M., Mesinger A., Ma Y.-Z., Sigurdson K., 2014, MNRAS, 438, 2664

Sobacchi E., Mesinger A., 2013, Mon. Not. Roy. Astron. Soc., 432, 3340

Sokolenko A., Bondarenko K., Brinckmann T., Zavala J., Vogelsberger M., Bringmann T., Boyarsky A., 2018, J. Cosmology Astropart. Phys., 2018, 038

Spergel D. N., Steinhardt P. J., 2000, Phys. Rev. Lett., 84, 3760
Spergel D. N., et al., 2003, ApJS, 148, 175
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Springel V., et al., 2005, Nature, 435, 629
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Springel V., et al., 2008, MNRAS, 391, 1685
Springel V., et al., 2018, MNRAS, 475, 676
Tacchella S., Bose S., Conroy C., Eisenstein D. J., Johnson B. D., 2018, Astrophys. J., 868, 92
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Trenti M., Stiavelli M., Bouwens R., Oesch P., Shull J., Illingworth G., Bradley L., Carollo C., 2010, Astrophys. J. Lett., 714, L202
Trujillo-Gomez S., Schneider A., Papastergis E., Reed D. S., Lake G., 2018, MNRAS, 475, 4825
Tulin S., Yu H.-B., 2018, Phys. Rep., 730, 1
Tyson J. A., Kochanski G. P., Dell'Antonio I. P., 1998, ApJ, 498, L107
Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 043502
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M., Zavala J., Simpson C., Jenkins A., 2014a, MNRAS, 444, 3684
Vogelsberger M., et al., 2014b, MNRAS, 444, 1518
Vogelsberger M., et al., 2014c, Nature, 509, 177
Vogelsberger M., Zavala J., Cyr-Racine F.-Y., Pfrommer C., Bringmann T., Sigurdson K., 2016, MNRAS, 460, 1399
Vogelsberger M., Zavala J., Schutz K., Slatyer T. R., 2019, MNRAS, 484, 5437
Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Rev. Phys., 2, 42
Voytek T. C., Natarajan A., Jáuregui García J. M., Peterson J. B., López-Cruz O., 2014, Astrophys. J., 782, L9
Walker M. G., Peñarrubia J., 2011, ApJ, 742, 20
Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Evans N. W., Gilmore G., 2009, ApJ, 704, 1274
Walther M., Hennawi J. F., Hiss H., Oñorbe J., Lee K.-G., Rorai A., O'Meara J., 2018, ApJ, 852, 22
Wang J., White S. D. M., 2007, MNRAS, 380, 93
Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, Nature, 585, 39
Weinberg D. H., Hernsquit L., Katz N., Croft R., Miralda-Escudé J., 1997, in Petitjean P., Charlot S., eds, Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System. p. 133 (`arXiv:astro-ph/9709303`)

Weinberger R., et al., 2017, MNRAS, 465, 3291
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Wouthuysen S. A., 1952, AJ, 57, 31
Yoshida N., Springel V., White S. D. M., Tormen G., 2000, ApJ, 544, L87
Yoshiura S., Takahashi K., Takahashi T., 2020, Phys. Rev. D, 101, 083520
Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2018, Monthly Notices of the Royal Astronomical Society, 483, 2983–3006
Zaldarriaga M., Hui L., Tegmark M., 2001, ApJ, 557, 519
Zavala J., Frenk C. S., 2019, Galaxies, 7, 81
Zavala J., Jing Y. P., Faltenbacher A., Yepes G., Hoffman Y., Gottlöber S., Catinella B., 2009, ApJ, 700, 1779
Zavala J., Vogelsberger M., Walker M. G., 2013, MNRAS, 431, L20
Zavala J., Lovell M. R., Vogelsberger M., Burger J. D., 2019, Phys. Rev. D, 100, 063007
Zehavi I., et al., 2011, ApJ, 736, 59
Zjupa J., Springel V., 2017, MNRAS, 466, 1625
Zwicky F., 1933, Helvetica Physica Acta, 6, 110
de Blok W. J. G., McGaugh S. S., 1997, MNRAS, 290, 533
de Blok W. J. G., McGaugh S. S., Bosma A., Rubin V. C., 2001, ApJ, 552, L23
van Haarlem M. P., et al., 2013, Astron. Astrophys., 556, A2
van den Aarssen L. G., Bringmann T., Pfrommer C., 2012, Phys. Rev. Lett., 109, 231301