Fundamental physics from large-scale structure: Theory modeling and neutrino masses

Caio B. de S. Nascimento

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:
Marilena Loverde, Chair
Matthew McQuinn
Isabel Garcia Garcia

Program Authorized to Offer Degree:
Physics

University of Washington

Abstract

Fundamental physics from large-scale structure: Theory modeling and neutrino masses

Caio B. de S. Nascimento

Chair of the Supervisory Committee:
Marilena Loverde
Department of Physics

In the past few decades the field of cosmology has entered a new era, where specific models for the dynamics of our universe on the largest scales can be tested against a diverse set of precision astronomical surveys, that map out the distribution of matter and energy in our Universe at different stages in its evolution. This revolution in the field was mostly driven by very precise observations of the Cosmic Microwave Background (CMB) radiation, a relic of the early universe from the time when free electrons and protons first combined to form neutral hydrogen. At this point the universe was very homogeneous and isotropic, with deviations from homogeneity of the order of one part in a hundred thousand.

Over the course of the subsequent billions of years, the anisotropies slowly increased due to the instability associated to gravitational attraction, resulting in an intricate pattern for the large-scale distribution of matter known as the cosmic web, where galaxies eventually formed. The next frontier in observational cosmology consists in mapping out the distribution of galaxies in our observable Universe, an effort which is already underway. Such a comprehesive map of the cosmos carries a lot of information about the initial conditions of our Universe, its matter and energy contents, and the nature of gravitational interactions.

The prospect of extracting fundamental physics from large-scale structure probes will only be fully realized if this rich experimental program is matched with both accurate theoretical models and efficient computational tools. This dissertation develops novel methods in both

directions, which help to maximize the scientific return from ongoing and upcoming surveys, and to ensure that the unprecedented amounts of data to be collected in the near future can be efficiently analyzed. Most of these new developments focus on the goal of using cosmology to extract the neutrino mass scale, which in many ways is the best motivated science target of upcoming astronomical surveys. At least two of the three neutrino species are known to be massive due to terrestrial neutrino oscillation experiments, but the mechanism underlying their mass generation is yet unknown, and is likely to be a key piece of the puzzle to reveal new physics beyond the standard model.

We examine the implementation of massive neutrinos in Newtonian simulations of non-linear structure formation, particularly emphasizing the systematic bias introduced by neglecting special relativistic effects. We show that such simulations necessarily overestimate the total distance traversed by neutrino particles, and one must carefully choose the initial conditions to mitigate this effect.

We propose a novel numerical implementation of massive neutrinos in state-of-the-art Boltzmann surveys, the Generalized Boltzmann Hierarchy (GBH), suitable for the linear regime of structure formation. The GBH integrates out the momentum dependence of the neutrino distribution function from the outset, producing a system of ordinary differential equations which are much simpler and faster to solve than the traditional Boltzmann hierarchy. We also introduce a novel fluid approximation for the dynamics of massive neutrino perturbations, which is much more accurate than previously introduced fluid approximation schemes and nicely complements the GBH in solving for the dynamics across all relevant physical scales.

We make significant advancements in the theoretical modeling of neutrino wakes, a new signature of neutrino masses in the large-scale structure, which refers to the preferential accumulation of neutrino particles downstream of moving cold dark matter structures, and results in the emergence of a dipole distortion to the neutrino density field. We derive

the effect from first principles, allowing for a better characterization of its observational signatures, and to forecast the capability of future surveys in detecting this effect.

We finally explore perturbative methods for the nonlinear evolution of cosmological perturbations in the late Universe. We develop a phase-space approach which circumvents the need to assume a perfect fluid from the outset, but show how to recover the results of Standard Perturbation Theory (SPT). We further explain how a key ingredient of the Effective Field Theory approach to Large Scale Structure (EFTofLSS), the sound speed counterterm, naturally emerges within our framework. However, we argue for the necessity of the full EFTofLSS framework to self-consistently model the departures from an ideal fluid. We also provide a novel analytic calculation to estimate the numerical value of the sound speed counterterm, based on separate universe techniques. This calculation is in good agreement with the measurements from full N-body simulations, and provides valuable information about the cosmology dependence of the sound speed and what types of nonlinear structures shape its value.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

These past many years as a graduate student were a period of immense professional and personal growth for me, which was only possible due to the incredible support I received from my mentors. First and foremost, I will forever be grateful to my advisor, Prof. Marilena Loverde, for her outstanding supervision that pushed me to become the independent researcher I am today. Her insights on my research would always nudge me to the right direction, and I will always carry with me her great advices on how to make a good plot, present my research, give talks and write papers, all of which played a decisive role in building my science communication skills. I also thank her for creating such a supportive environment, which was a major influcence to my growing into a much more confident person over the years, both in and out of the workplace.

I am grateful to Prof. Matt McQuinn for the many enlightening discussions and collaboration, Zachary Weiner for debating ideas and offering valuable advice, and everyone in the Dark Universe Science Collaboration (DUSC) (and the physics department more broadly as well) for so many stimulating conversations about science at journal club, and life in general over lunch or at happy hours. Two special mentions are warranted in particular. First to to the remaining members of my committee, Profs. Masha Baryakhtar, Isabel Garcia Garcia and Miguel Morales. Second, to my friends Murali Saravanan, Ella Henry and John Franklin Crenshaw who made my time at the University of Washington so much more enjoyable. I also want to thank Prof. Vivian Miranda for guidence on the process of reaching my long term goal of eventually becoming a faculty member in a Brazillian institution of higher education.

Finally, remaining sane while completing a doctorate degree is no easy task, and it requires a strong support system. I am lucky to have so many amazing people around me to thank

for their continued support (too many to name individually, sorry you all). My family back in Brazil for their love and persistent encouragement. My childhood friends from Belém for free therapy sessions, legal advices, medical consultations, and for taking me to Paysandu games. The friends I made in Brasília during undergrad, which I am still regularly in touch with despite them being spread all around the world. The friends I have made during my brief time at Stony Brook who now have a moral obligation to come and visit me in Waterloo in the next three years. The Brazillian community in Seattle who made life here feel like home to me, including friends I met through UW (and its community of latin students as well), friends I have made trough Forró that feed me all sorts of delicious vegan dishes, that take me to a variety of fun adventures and shared so many moments listening to great music of questionable quality and content. I also could not forget of the friends I occasionally meet to jam samba and pagode, and the ones I play chess with. My life would not have been the same in the past few years without all of you.

DEDICATION

This thesis is dedicated to my mother and older sister, Carina and Mariana Senna, to whom I owe everything. This, and everything else I ever accomplished, was only made possible by your relentless support since as far as I can remember.

Chapter 1

INTRODUCTION

1.1 The standard cosmological model

Over the past few decades the Lambda Cold Dark Matter (ΛCDM) model has prevailed as the leading theory to describe the dynamics of our Universe on the largest physical scales. With just a few free parameters, this standard cosmological model has been shown to agree with a wide variety of astrophysical observations spanning billions of years of cosmic evolution (see Fig. 1.1 for a visual representation of the cosmic timeline), from galactic scales to the entire observable Universe. This striking success is a great triumph of modern physics, since the ΛCDM model follows as a consequence of combining General Relativity (GR) with the Standard Model of particle physics [1].

The standard model of cosmology is based on the cosmological principle, which states that the Universe is homogeneous and isotropic on sufficiently large scales, and is widely supported by astronomical observations. Under this assumption, Einstein's field equations of GR can be exactly solved when the different sources of mass and energy are accounted for [5]. The resulting Friedmann–Lemaître–Robertson–Walker (FLRW) solution describes the spacetime structure of the Universe on very large scales.

All the stable constituents of the standard model of particle physics contribute to the mass and energy budget of the Universe, including baryons, photons and neutrinos. Additionally, the ΛCDM model also features a cold dark matter species which only interacts gravitationally with all the other degrees of freedom, and a dark energy component in the form of a cosmological constant. The microphysical description of both ingredients are yet

[1] With the addition of two important phenomenological ingredients, i.e. a dark matter abudance and a fine-tuned cosmological constant.

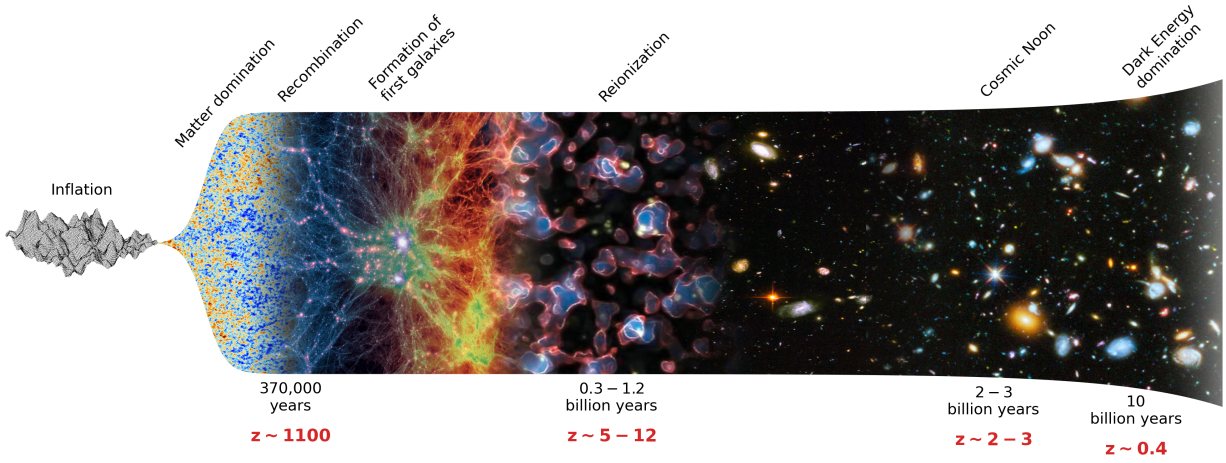

Figure 1.1: Schematic representation of the cosmic timeline, showing important events in the history of the Universe. The most popular model for the very first stage of evolution is cosmic inflation, where the Universe goes through a period of exponential expansion, generating the initial conditions for the subsequent standard big bang evolution. Then a period of radiation domination follows which eventually ends, marking the beginning of matter domination. At recombination free electrons and protons first combine to form neutral hydrogen, and light is allowed to free stream towards us at the present time, forming the so-called Cosmic Microwave Background (CMB) radiation. After millions of years the first galaxies start to form, and they emit ultraviolet radiation which slowly ionize the universe once again during the period of reionization. The peak of star formation activity is denoted cosmic noon, and finally dark energy starts to dominate when the universe is already about ten billion years old. This wonderful diagram is a combination of images from [1–4], and was created by John Franklin Crenshaw.

unknown, and are major scientific drivers of most ongoing efforts in the field.

Many models have been proposed to explain the fundamental nature of dark matter, including weakly interacting massive particles [6], axion-like particles [7] and even primordial black holes [8]. There is a plethora of ongoing investigations, applying both direct and

indirect search strategies, to detect the dark matter whatever it may be. On the dark energy side, quantum mechanics dictates that the vacuum energy should contribute to the cosmological constant. However, quantum field theory calculations overestimate the size of the cosmological constant by at least several tens of orders of magnitude, as the cosmological constant we measure experimentally is fine-tuned to almost (but not exactly) zero. We have yet to understand why the vacuum energy does not contribute to the cosmological constant as dictated by quantum mechanics, and what exactly sets its specific value [9].

The expansion of the Universe is controlled by a single function of cosmic time t, the scale factor $a(t)$. Physical distances are given by $r = a(t)x$, where x represents comoving distances that are agnostic to the expansion. The Hubble rate, $H = \dot{a}/a$, parametrizes the rate at which the Universe is expanding, and it satisfies the Friedmann equation

$$H^2 = \frac{8\pi G}{3}\bar{\rho}_{\rm tot} \equiv \frac{8\pi G}{3}\sum_i \bar{\rho}_i \,, \tag{1.1}$$

where G is the Newtonian constant of gravitation, $\bar{\rho}_{\rm tot}$ is the total background (or average) energy density of the Universe, the sum in the second equality runs over all species contributing to the energy density, and time dependencies are omitted for simplicity of notation.

The scaling of each $\bar{\rho}_i$ with scale factor are as follows. Both baryons and cold dark matter are nonrelativistic species, and hence satisfy $\bar{\rho}_{\rm cb} \propto a^{-3}$ due to three dimensional volume dilation. On top of that, the wavelength of radiation is stretched due to gravitational redshift, and hence $\bar{\rho}_{\rm r} \propto a^{-4}$. This includes both photons and neutrinos, assuming for now that neutrinos are massless particles. The massive neutrino case will be discussed in detail in Sec. 1.3. Within the ΛCDM model, dark energy takes the form of a cosmological constant with $\bar{\rho}_{\rm de} \propto a^0$. If we allow dark energy to evolve with time, its energy density can be parametrized as follows

$$f_{\rm de}(a) = \frac{\bar{\rho}_{\rm de}(a)}{\bar{\rho}_{\rm de,0}} = \exp\left\{3\int_a^1 \frac{da'}{a'}\left[1 + w(a')\right]\right\} \,, \tag{1.2}$$

where we employ the standard convention of setting $a(t_0) = 1$ at the present time t_0 (measuring comoving distances in terms of physical distances evaluated at the present time), and set

$\bar{\rho}_{\rm de,0} = \bar{\rho}_{\rm de}(a=1) \equiv \bar{\rho}_{\rm de}(t=t_0)$ [2]. In Eq. (1.2), $w(a)$ is the so-called equation of state. The case of $w=-1$ corresponds to a cosmological constant. When $w>-1$ the energy density decreases with time, while it increases with time when $w<-1$, known in the literature as the phantom regime.

There is one additional contribution to the right-hand side of Eq. (1.1) coming from the curvature of three-dimensional space, scaling like $\bar{\rho}_k \sim a^{-2}$. It will be convenient to define the present-day fractional contribution of each species to the total energy density, $\Omega_i = \bar{\rho}_{i,0}/\bar{\rho}_{\rm tot,0}$, in terms of which Eq. (1.1) reads

$$H = H_0\sqrt{\Omega_{\rm r}a^{-4} + \Omega_{\rm cb}a^{-3} + \Omega_k a^{-2} + \Omega_{\rm de}f_{\rm de}(a)}\,, \tag{1.3}$$

where $\Omega_{\rm cb} = \Omega_{\rm c} + \Omega_{\rm b}$. Evaluating Eq. (1.3) at the present time yields the constraint $\Omega_{\rm de} = 1 - \Omega_{\rm r} - \Omega_{\rm cb} - \Omega_k$. A comparison of the theoretical model against various cosmological probes can constrain the parameters $\Omega_{\rm r}$, Ω_k, $\Omega_{\rm c}$, $\Omega_{\rm b}$ and H_0 [assuming a known $f_{\rm de}(a)$, as in the ΛCDM model where $f_{\rm de}(a)=1$]. Measurements of the blackbody spectrum of the Cosmic Microwave Background (CMB), to extract its average temperature, directly provide $\Omega_{\rm r}$ when coupled to an understanding of the neutrino decoupling process with standard model interactions [10]. Additionally, observations are consistent with the assumption of a flat universe, so we set $\Omega_k = 0$ [3].

At the end, we need three free parameters to fix the background expansion history in the ΛCDM model: $\Omega_{\rm c}$, $\Omega_{\rm b}$ and H_0. This is not, however, the full story as the universe we live in is not exactly homogeneous nor isotropic. We will now discuss the origin of cosmological perturbations before diving into how astronomical observations can be used to infer the remaining free parameters of the ΛCDM model, including the ones characterizing the primordial fluctuations.

---

[2] A similar notation will also be adopted for other quantities evaluated at the present time.

[3] This parameter is constrained by observations to $|\Omega_k| \lesssim 10^{-3}$ [11], and $\Omega_k \to 0$ is a generic prediction of inflationary models [12]. However, there are hints of nonzero spatial curvature in specific data combinations [13, 14].

1.2 Growth of cosmological perturbations

Thus far our discussion was based on the assumption of homogeneity and isotropy, ignoring the presence of cosmological perturbations without which galaxies and planets would never be able to form. The Universe today has large fluctuations on small scales, with regions of mostly empty space, the so-called cosmic voids, contrasting with the high density environments in galaxy clusters. The situation was different, however, at early times. The CMB radiation offers a glimpse of the Universe from the time of recombination, when it was only around 370.000 years old (see Fig. 1.1). The CMB temperature map, shown in Fig. 1.2, reveals a photon distribution that is extremely homogeneous, with anisotropies of the order of one part in a hundred thousand.

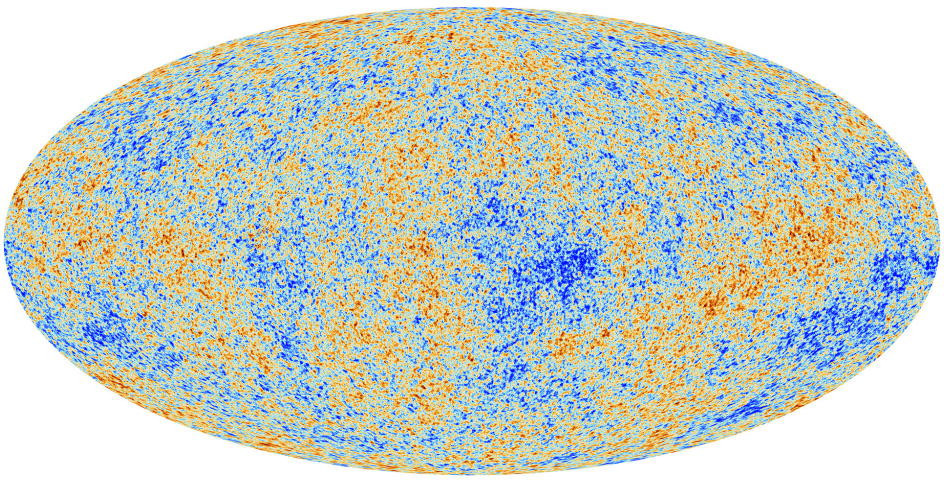

Figure 1.2: Map of temperature anisotropies in the Cosmic Microwave Background radiation. One in a hundred thousand is the typical size of the fluctuations, revealing an early universe that is very homogeneous and isotropic.

Over the course of billions of years of cosmic evolution these small anisotropies grew to become the large fluctuations we now observe, as shown in Fig. 1.3. This is due to the unstable nature of gravitational interactions, which tends to pull things together and build even larger structures over time.

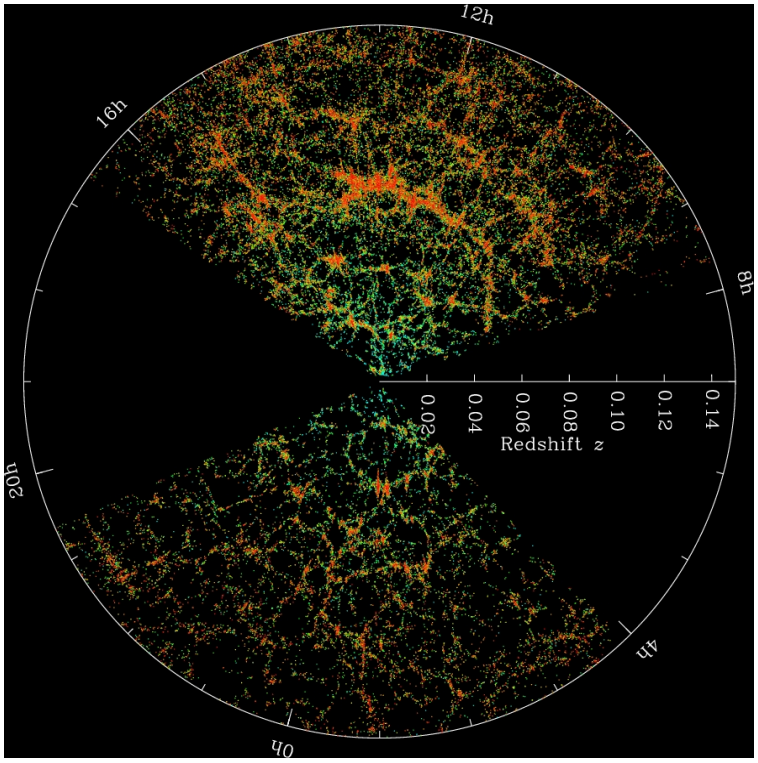

Figure 1.3: Map of the distribution of galaxies in the nearby Universe, produced with the Sloan Digital Sky Survey (SDDS), which traces the underlying distribution of dark matter in the cosmic web. The large fluctuations observed today grew from quantum fluctuations during inflation over billions of years of gravitational evolution.

We can also consider the time-reversal of the previous discussion, and conceive of an even earlier period in the history of the Universe, compared to the time of recombination, when deviations from homogeneity were incredibly tiny. In fact, we now believe that all cosmological perturbations were generated via quantum fluctuations in the very early Universe, during a period of inflation [15] (see Fig. 1.1). The inflationary mechanism not only provides the seeds to all structures in the universe, it also explains the homogeneity of the CMB when different patches of the sky were never in causal contact during the standard big bang evolution [16].

The patterns of anisotropies seen in the CMB and large-scale distribution of galaxies, illustrated in Figs. (1.2) and (1.3) respectively, contain a lot of information about both the inflationary initial conditions of the Universe and its dynamics, which is sensitive to the energy content in the various species and their interactions. To access this information it is convenient to first compress the data into statistics that are easier to handle, i.e. the so-called summary statistics [4]. Then we also need to develop robust theoretical models for these summary statistics, to allow for the interpretation of experimental results in terms of fundamental physics.

We will now briefly explore the theoretical tools we have in our disposal to model the growth of cosmological perturbation in the Universe, focusing on perturbative techniques that solve for the fluctuations in powers of (small) deviations from homogeneity. Along the way we will also mention some of the summary statistics usually adopted to extract cosmological information out of CMB and galaxy survey datasets.

1.2.1 Linear evolution

In the early Universe deviations from homogeneity were quite small, as illustrated in Fig. 1.2. This means that the growth of cosmological perturbations at early times can be understood within a framework that adds linear perturbations to the FLRW solution of GR [5].

In an isotropic universe the different sources of matter and energy can be modeled as ideal fluids due to the extended symmetry, but this description becomes insufficient as soon as perturbations are turned on. In this case, one generically needs to track the particles trajectories on the perturbed spacetime in phase-space by solving the Boltzmann equation. Since the different species backreact into the spacetime geometry, in reality one needs to solve a couple system of Boltzmann and Einstein equations [19]. A more detailed quantitative analysis of the linear theory of cosmological perturbations will be given in Chapters 3 and

[4] Although it is also possible, and sometimes advantageous, to work directly at the field level [17, 18].

[5] The Universe is still quite homogeneous today when smoothed over sufficiently large scales, such that linear theory can be directly applied even at late times in certain regimes.

4, particularly within the context of implementing massive neutrino species.

We will now focus on a simple description of the main CMB observable, the temperature power spectrum. What is directly measurable from the CMB radiation is the energy flux at a given frequency, coming from a given direction $\hat{n}$ on the sky. One then fits the spectrum to a blackbody template to extract the temperature $T(\hat{n})$ [6]. Averaging over the different directions produces the CMB monopole, $\bar{T} = \langle T(\hat{n}) \rangle \approx 2.725\text{K}$. The anisotropies, on the other hand, are captured by

$$\Theta(\eta) = \frac{T(\eta) - \bar{T}}{\bar{T}} \,. \tag{1.4}$$

The typical value for this quantity is $\Theta \sim 10^{-5}$, and $\langle \Theta(\eta) \rangle = 0$ by definition. The most useful summary statistic of the CMB is the two-point correlation function

$$\langle \Theta(\hat{n}) \Theta(\hat{n}') \rangle = \sum_{\ell} \frac{2\ell + 1}{4\pi} C_\ell P_\ell(\cos\theta) \,, \tag{1.5}$$

which is only a function of $\hat{n} \cdot \hat{n}' = \cos\theta$ due to the underlying statistical isotropy, and can hence be expanded into Legendre polynomials $P_\ell(\cos\theta)$, with coefficients C_ℓ. Since the anisotropies in the early Universe are so small, the assumption of linear evolution holds to an excellent degree. This means that the observed C_ℓ is related to the initial conditions, i.e. the spectrum of perturbations in the curvature ζ at the end of inflation, by a linear transformation. More concretely

$$C_\ell = \frac{4\pi}{(2\ell+1)^2} \int d\log k \; T_\ell^2(k) \Delta_\zeta^2(k) \,, \tag{1.6}$$

where $T_\ell(k)$ are transfer functions encapsulating the linear dynamics from the end of inflation to the observed temperature in the celestial sphere [23], and $\Delta_\zeta^2(k)$ is the primordial (dimensionless) power spectrum of curvature perturbations. The latter ingredient is often parametrized as a generic power-law

$$\Delta_\zeta^2(k) = A_s \left(\frac{k}{k_{\text{piv}}} \right)^{n_s - 1} \,, \tag{1.7}$$

[6] The CMB radiation is the most perfect blackbody spectrum ever observed. The search for deviations from a blackbody template is very well motivated theoretically, and is the science case of many planned experiments [20, 21] and ongoing research [22].

which is agnostic to the inflationary model. In Eq. (1.7), the amplitude A_s and spectral index n_s are free parameters of the ΛCDM model, while the pivot scale $k_{\rm piv} = 0.05{\rm Mpc}^{-1}$ can be chosen arbitrarily.

The two-point correlation function in Eq. (1.5) carries all of the information contained in the CMB map (see Fig. 1.2) under the assumption of a Gaussian distribution. There are two different categories of non-Gaussianities in the CMB. The first would be of primordial origin, in the spectrum of curvature perturbations ζ at the end of inflation. There are upper bounds on the amplitude of primordial non-Gaussianities, but no detections thus far [24, 25]. The second type of non-Gaussianities are sourced by CMB secondaries, which refer to physical effects that are not intrinsic to the state of the Universe at the time of recombination, but are rather imprinted as the CMB photons travel towards us [7].

There is one final free parameter of the ΛCDM model yet to be introduced, that needs to be included when fitting to CMB observations. As illustrated in Fig. 1.1, the period of reionization was triggered by the UV radiation emitted by the first galaxies to be formed in the Universe. As a consequence, some of the CMB photons traveling towards us from the period of recombination end up scattering off free electrons along the line-of-sight, reducing the amplitude of the CMB. This effect is theoretically calculable from first principles, but in reality the physics of galaxy formation and subsequent reionization is too complex and not yet understood. This is mitigated by adopting the the so-called optical depth to reionization, τ, as a free parameter.

Putting it all together, there are six free parameters of the ΛCDM model: $\Omega_{\rm c}$, $\Omega_{\rm b}$, H_0, A_s, n_s and τ. These are to be found by matching the model to cosmological observations. Starting with the CMB, the quantity

$$\mathcal{D}_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell \,, \tag{1.8}$$

is plotted in Fig. 1.4 from the Planck satellite. There is excellent agreement between the

[7] Some examples of CMB secondaries are weak gravitational lensing [26], Sunyaev–Zeldovich effects [27] and the integrated Sachs-Wolfe effect [28].

theory fit (solid blue curve) and experimental data points (in red), fixing the cosmological parameters with percent level uncertainties [11] [8].

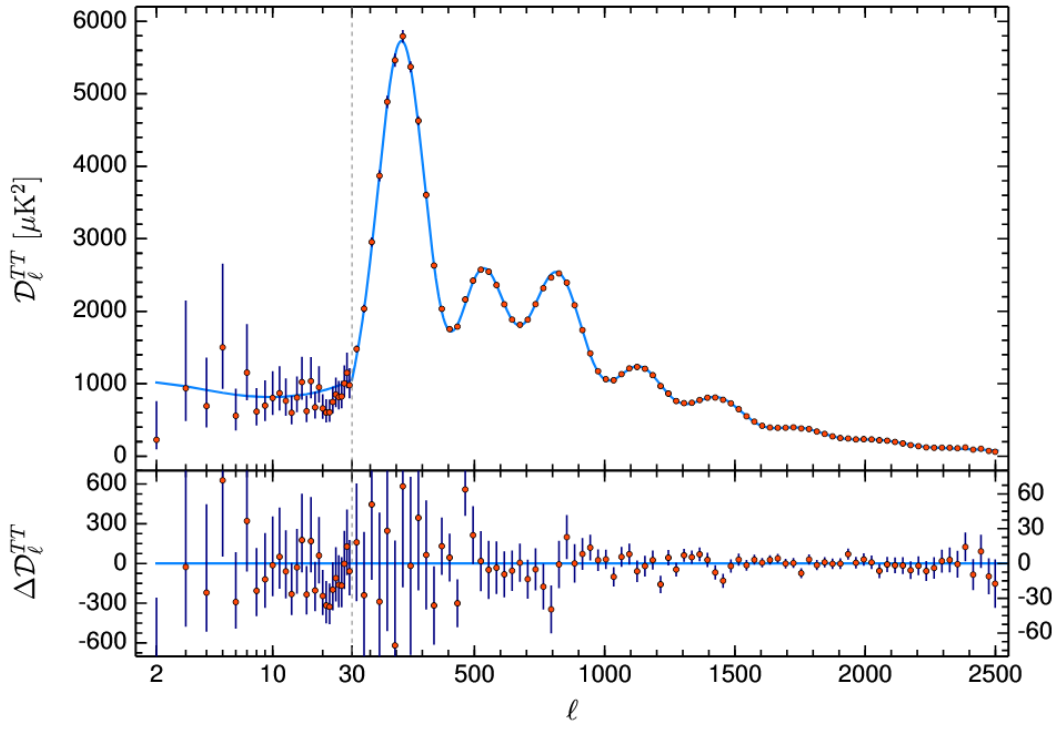

Figure 1.4: Plot of the angular power spectrum of the CMB temperature, with a ΛCDM theoretical fit shown as the solid blue curve, and experimental data points (from the Planck satellite) in red. Residuals between theory and experiment are plotted in the bottom panel. The excellent agreement between theory and experiment highlights the success of the ΛCDM model in describing the large-scale dynamics of the Universe, fixing the six cosmological parameters with percent level uncertainties. Figure extracted from [11].

The different features present in the angular power spectrum in Fig. 1.4 carry a lot of physical content. For example, the wiggles are due to Baryon Acoustic Oscillations (BAO).

[8] The reader may have noticed that some of the data points at low ℓ, such as $\ell = 2$ and $\ell \sim 20$, lie below the theoretical expectation. These are just a few examples of the so-called CMB anomalies [29]. They are likely the result of random statistical fluctuations, but may also be pointing to new physics beyond the ΛCDM model.

In the pre-recombination universe the Baryons are tightly coupled to photons via Thomson scattering, and the dynamics of the resulting photon-baryon fluid is governed by a competition between gravity and pressure support, generating the oscillations. Also, the exponential damping of the CMB at $l \gtrsim 1000$ is due to photon diffusion.

We will now move on to describe how we can model the emergence of nonlinear gravitational effects at late times. This is crucial because significant deviations from homogeneity start to appear, at small scales, around the time of galaxy formation. By the time we reach cosmic noon, it is already important to model the effects of nonlinearities even on cosmological scales. More importantly, the interpretation of maps such as the one illustrated in Fig. 1.3 require nonlinear models to obtain unbiased measurements of cosmological parameters.

1.2.2 Nonlinear evolution at late times

There are various complementary approaches to model the effects of nonlinear evolution in the growth of cosmological perturbations at late times. Perhaps the most reliable of them are numerical simulations, where most employ the N-body technique with Newtonian equations of motion, and Newtonian gravity, in an expanding cosmological background [30] [9].

The disadvantages of numerical simulations are the high computational cost [37], the difficulty in interpreting the results in terms of the underlying physical processes and the lack of flexibility to implement extensions of the standard cosmological model, which require dedicated efforts for each specific case. On the other hand, analytical methods are designed to precisely overcome these difficulties, but are however limited in their range of applicability [38]. For these reasons, numerical and analytical methods offer highly complementary approaches to study the nonlinear growth of cosmological perturbations.

Analytical methods can be further subdivided into two categories. The first attempts to describe the dynamics in the fully nonperturbative regime, producing a model that is

[9] Some simulations of nonlinear structure formation attempt to tackle the problem using a fully relativistic approach [31–33]. However, it is possible to incorporate most relativistic effects in Newtonian N-body simulations by post-processing the simulation outputs [34–36].

not limited by the range of scales where it can be applied. However, to accomplish this one needs to make many simplifying approximations, which results in a lack of control over the accuracy of modeling. A successful example of such a nonperturbative approach is the halo model [39], which is based on excursion set theory [40]. We will focus on a second category of analytical methods, based on perturbation theory. The idea is to systematically expand around small deviations from isotropy and homogeneity, which in one hand enables precise control over theoretical errors, but on the other hand can only be applied on large perturbative scales.

A detailed account on the perturbation theory approach to large scale structure formation will be given in Chapters 6 and 7. For now we will briefly describe the theoretical model, which is based on the Standard Perturbation Theory (SPT). It relies on the assumption that the nonlinear gravitational evolution is driven by a perfect dark matter fluid, with negligible stress. In that case the dynamics is governed by the continuity and Euler equations (coupled to the Poisson equation), which can be solved in a perturbative expansion in powers of the fluid density and velocity fields [41].

A shortcoming of SPT is that it does not account for the emergence of dissipative effects on mildly nonlinear scales. This is cured within the framework of the Effective Field Theory of Large Scale Structure (EFTofLSS), which integrates out the effects of short-wavelength fluctuations backreacting on mildly nonlinear scales, while remaining agnostic to the detailed dynamics on small scales [42, 43]. This is accomplished by writing the most general expansion for the effective stress tensor with all symmetry allowed operators, multiplied by free coefficients to be marginalized over.

The directly measurable quantities from a map of the Universe, such as in Fig. 1.3, are the statistical properties of the galaxy number density, which can be related to the statistical properties of the underlying matter density field, $\rho_{\rm m}(\vec{x})$, via a bias expansion [44]. The average matter density can be defined as $\langle\rho_{\rm m}(\vec{x})\rangle = \bar{\rho}_{\rm m}$, and the two-point correlation function $\xi(r)$ is given by

$$\langle\rho_{\rm m}(\vec{0})\rho_{\rm m}(\vec{r})\rangle = \bar{\rho}_{\rm m}^2\left[1+\xi(r)\right] , \tag{1.9}$$

and its Fourier transform,

$$P(k) = 4\pi \int_0^\infty dr r^2 \xi(r) \frac{\sin(kr)}{kr}\,, \tag{1.10}$$

the matter power spectrum. Just as in the CMB case, the matter power spectrum carries all of the statistical information in the matter density field, for a Gaussian distribution. Any primordial non-Gaussianities imprinted in the CMB would also appear in the maps of large-scale structure, but now there are also non-Gaussianities due to nonlinear gravitational evolution.

The matter power spectrum is plotted in Fig. 1.5, for varying theoretical models, at the present time and assuming the cosmology adopted in the Farpoint simulation [45]. The nonlinear prediction is computed with EuclidEmulator2 (EE2), which interpolates the outputs of high-resolution N-body simulations [46]. The linear result is derived under the same framework used to produce the solid blue curve in Fig. 1.4 for the CMB temperature power spectrum, using the Cosmic Linear Anisotropy Solving System (CLASS) [47]. The SPT calculation is carried out to next to leading order, at one-loop precision. Finally, the EFT adds one extra free parameter to one-loop SPT, the so-called effective sound speed, which is obtained by matching the EFT theoretical template to the nonlinear result from EE2.

Residuals with respect to the nonlinear power spectrum from EE2 are plotted in Fig. 1.6. The EFTofLSS provides an accurate model for the clustering of matter (and biased tracers) at mildly noninear scales, with percent level errors out to $k \sim 0.2\text{Mpc}^{-1}$ [10], and it is by now a key component of the theory modeling applied in the interpretation of real data [52–55].

1.3 Neutrino mass signatures

We have assumed thus far that neutrinos are massless, in accordance with the standard model of particle physics. However, we now know from neutrino flavor oscillation experiments that

[10] An inspection of Figs. 1.5 and 1.6 reveals that the EFT model adopted here does not properly account for the oscillatory features in the matter power spectrum due to BAO (which is also responsible for the wiggles in the two-point function of the CMB as showed in Fig. 1.4). This is because some infrared contributions to the loop integrals need to be resummed to all orders in perturbation theory, after which the EFT model nicely reproduces the wiggly component of the nonlinear power spectrum [48–51].

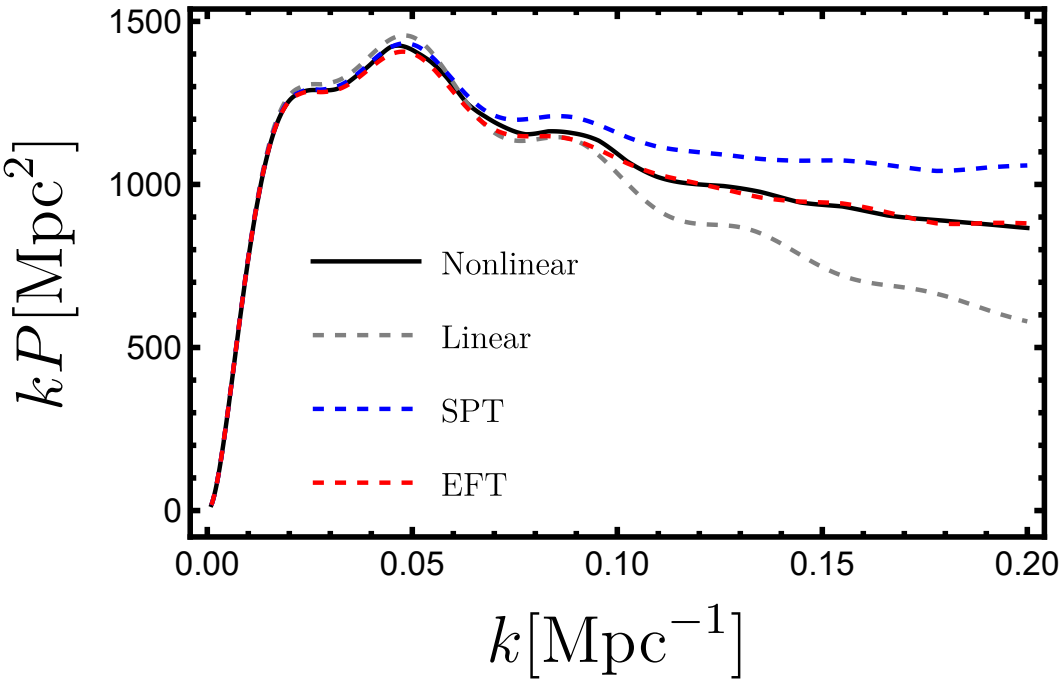

Figure 1.5: Matter power spectrum evaluated with various theoretical models. The solid black curve is the fully nonlinear result from EuclidEmulator2 (EE2), designed to match high-resolution N-body simulations [46]. The dashed gray, blue and pink curves correspond to results from linear perturbation theory, one-loop SPT and one-loop EFT, respectively. The EFT accurately models the matter power spectrum at mildly nonlinear scales, at the cost of adding one free parameter to SPT.

at least two of the three neutrino states are massive, with a lower bound on the sum of the masses of $M_\nu = \sum_\nu m_\nu \gtrsim 0.06\text{eV}$ [56]. The tightest upper bound to a weighted mass average with terrestrial experiments has been obtained by the KATRIN collaboration, $m_{\nu,\beta} < 0.45\text{eV}$ at 90% confidence level [57], with measurements of beta decay.

On the other hand, the existence of nonzero neutrino masses leave distinctive signatures on cosmological observables, which have been exploited to produce tight upper bounds to the sum of neutrino masses, of order $M_\nu \lesssim (0.1 - 0.2)\text{eV}$ [11, 58]. Near term surveys will

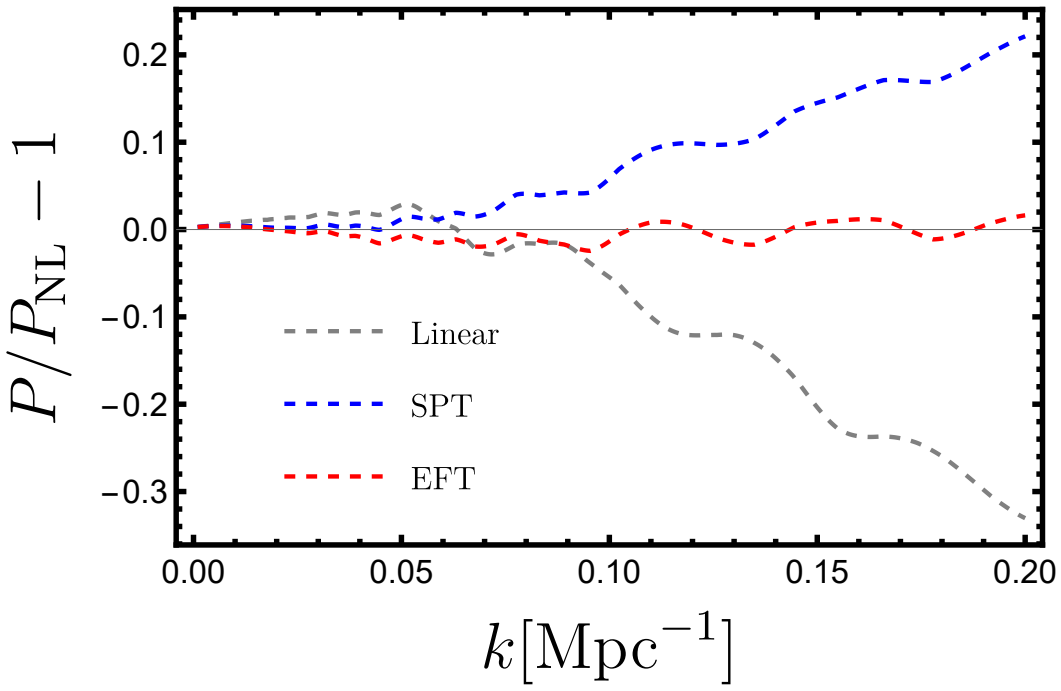

Figure 1.6: Same as in Fig. 1.5, but now showing residuals with respect to the fully nonlinear result from EE2. The EFT accurately models the matter power spectrum at mildly nonlinear scales, with percent level errors out to $k \sim 0.2\mathrm{Mpc}^{-1}$.

have enough sensitivity to measure the neutrino mass scale for the first time ever in the upcoming decade [59–62]. We will now briefly describe the effects of neutrino masses on the expansion history of the universe and the growth of structure, and summarize the current status of neutrino mass constraints with cosmological observations.

In the very early Universe neutrinos were in thermal equilibrium with the primordial plasma, sustained by weak interactions. This changed when the Universe was about a second old, and the weak interaction rate was no longer sufficient to maintain thermal equilibrium against the opposition of Hubble expansion. At this point all neutrinos of a given species were ultra relativistic particles traveling at the speed of light, following a Fermi-Dirac distribution

in phase-space

$$\bar{f}(a_{\rm ini}, p) = g_\nu \left[1 + e^{\frac{p}{T_\nu(a_{\rm ini})}}\right]^{-1} , \tag{1.11}$$

where $g_\nu = 2$ counts the number of degrees of freedom in both left-handed neutrinos and right-handed antineutrinos, $T_\nu(a_{\rm ini})$ is the neutrino temperature at some early time $a_{\rm ini}$ and p is the physical momentum. After decoupling (when they drop out of thermal equilibrium), neutrino particles free-stream following geodesics of the FLRW universe, with the momentum redshifting away like $p \propto a^{-1}$. As a result, the neutrino species keeps the relativistic profile of Eq. (1.11) with a time-dependent temperature of $T_\nu = T_{\nu,0}/a$ [11], where $T_{\nu,0} = 1.95{\rm K} \approx 1.7 \times 10^{-4}{\rm eV}$ is the present-day value of the neutrino temperature. For a massive neutrino species, the transition to the nonrelativistic regime happens when $m_\nu \sim 3T_\nu$, and hence $a_{\rm tr} \sim 3T_{\nu,0}/m_\nu$. This is about $z \sim 200$ for $m_\nu = 0.1{\rm eV}$, where $z \sim 1/a$ is the redshift.

From the discussion around Eq. (1.11), the background neutrino energy density can be written as

$$\bar{\rho}_\nu(a) = \int \frac{d^3\vec{p}}{(2\pi)^3} \bar{f}(a,p) E(p) \, , \tag{1.12}$$

where $E(p) = \sqrt{p^2 + m_\nu^2}$ is the energy of a neutrino particle with momentum p. In the relativistic regime $p \gg m_\nu$, such that $E(p) \approx p$, and Eq. (1.12) scales like $\bar{\rho}_\nu \propto T_\nu^4 \propto a^{-4}$. We then recover the expected scaling of radiation [see lines of text above Eq. (1.2)]. On the other hand, in the nonrelativistic regime $p \ll m_\nu$, such that $E(p) \approx m_\nu$, and Eq. (1.12) scales as $\bar{\rho}_\nu \propto m_\nu T_\nu^3 \propto a^{-3}$, like matter.

When neutrinos are still relativistic in the early Universe, they make a significant fractional contribution to the total energy density of about

$$\frac{3T_\nu^4}{3T_\nu^4 + T_\gamma^4} \sim 40\% \, , \tag{1.13}$$

where the factor of 3 accounts for the three different neutrino states, and $T_\gamma = T_{\gamma,0}/a$ is the photon temperature with $T_{\gamma,0} = 2.725{\rm K}$ the CMB monopole. The existence of three neutrino species can be inferred from probes of the early universe expansion history, such as Big Bang

[11] This remains true even if the neutrinos are massive, and eventually become nonrelativistic.

Nucleosynthesis (BBN) via measurements of the primordial abundance of light chemical elements, and the CMB. These two observables constrain the so-called effective number of relativistic degrees of freedom $N_{\rm eff}$, which roughly counts the number of relativistic species in the early universe (besides the photon), to be $N_{\rm eff} \sim 3 \pm 0.3$ [11] [12].

Any signatures of neutrino masses in cosmology, on the other hand, necessarily involve probing neutrinos after they transition to the nonrelativistic regime, at late times. In this case, they contribute to the background expansion rate of the universe as a matter species as we showed above, such that the late-time expansion history is sensitive to the combination $\Omega_{\rm m} = \Omega_{\rm cb} + \Omega_\nu$, with [64]

$$\Omega_\nu = \frac{\bar{\rho}_{\nu,0}}{\bar{\rho}_{\rm tot,0}} = \frac{M_\nu}{93.14 h^2 {\rm eV}}\,, \tag{1.14}$$

the (massive) neutrino contribution to the total energy density today, where h is defined by $H_0 = 100h$ km/s/Mpc. The fractional contribution of neutrinos to the total matter density is given by $f_\nu = \Omega_\nu/\Omega_{\rm m}$. For $h = 0.7$ and $\Omega_{\rm m} = 0.3$, a neutrino mass sum of $M_\nu = 0.1$eV implies $f_\nu \approx 0.007$, and hence a sub-percent contribution to the total density of matter.

The presence of nonzero neutrino masses also imprints distinctive signatures in the perturbed Universe. This is because massive neutrinos, while being nonrelativistic at late-times, differ from cold dark matter by its significant thermal velocities of about $\sigma_\nu \sim T_\nu/m_\nu$ (for an individual neutrino mass of $m_\nu = 0.05$eV, $\sigma_{\nu,0} \sim 1000$km/s). This significant thermal motion prevents neutrino particles from falling into the gravitational potential wells of cold dark matter, washing out anisotropies below the free-streaming scale $\lambda_{\rm fs} \sim \sigma_\nu/H$ ($\lambda_{\rm fs,0} \sim 15$Mpc for $m_\nu = 0.05$eV and $h = 0.7$).

This manifests itself in the matter power spectrum as a small-scale suppression. In order to separate this effect of neutrino masses on cosmological fluctuations from changes to the background expansion caused by massive neutrinos, one usually computes fractional differences in the matter power spectrum caused by varying the total neutrino mass M_ν,

[12] In reality, $N_{\rm eff}$ parametrizes contributions to the total energy density from light relics during radiation domination. This quantity is sensitive to the decoupling time of the light species, and it does not need to be an integer. Any departures from the standard model prediction of $N_{\rm eff} \approx 3$ would point to new physics beyond the standard model (see Figure 1 in [63]).

while keeping $\Omega_{\rm m}$ fixed. As we explained above, this ensures that the late-time expansion history of the Universe is held fixed as the neutrino mass is varied. The quantity of interest

$$\delta_P(k) = \frac{1}{f_\nu}\frac{P(k|M_\nu) - P(k|M_\nu = 0)}{P(k|M_\nu = 0)}\,, \tag{1.15}$$

is plotted in Fig. 1.7 for a total neutrino mass $M_\nu = 0.1$eV, in the present day Universe and assuming the same cosmological parameters as in Fig. 1.5. The dashed gray curve is the linear theory prediction from CLASS, while the solid black curve is the fully nonlinear result from EE2, showing the characteristic "spoon-like" feature observed from simulations [65]. Many of the technical aspects involved in the implementation of massive neutrinos in simulations of nonlinear structure formation will be addressed in Chapter 2.

The existing cosmological constraints on the total neutrino mass scale are driven by both background and perturbation effects of massive neutrinos alike [66]. The CMB is sensitive to the background effects of neutrino masses via its modulation of the time of matter-radiation equality, and distance to recombination. It is also sensitive to the perturbation effects of massive neutrinos primarily via the weak gravitational lensing of the CMB, whereby the trajectory of CMB photons coming from recombination are distorted due to the intervening large-scale structure. Similar statements also apply for the clustering of galaxies, which are sensitive to cosmological distances that depend on the background effects of massive neutrinos, but also directly probe the nonlinear structures at late-time, which are suppressed in the presence of nonzero neutrino masses as shown in Fig. 1.7.

Combinations of the most recent CMB and galaxy clustering datasets place tight upper bounds on the neutrino mass scale, $M_\nu \lesssim (0.06 - 0.08)$eV, under the minimal ΛCDM + M_ν cosmological model [67–69]. These constraints are starting to dangerously approach the lower bound from oscillation experiments, of $M_\nu \gtrsim 0.06$eV. However, the significance of this emerging "neutrino mass tension" is not yet clear. Both CMB and galaxy clustering datasets are separately consistent with a ΛCDM cosmology, but with different best fit values for $\Omega_{\rm m}$ [70] [13]. The tension between these two probes can be resolved with the introduction

[13]There are other such discrepancies reported in the literature. The one attracting the most atten-

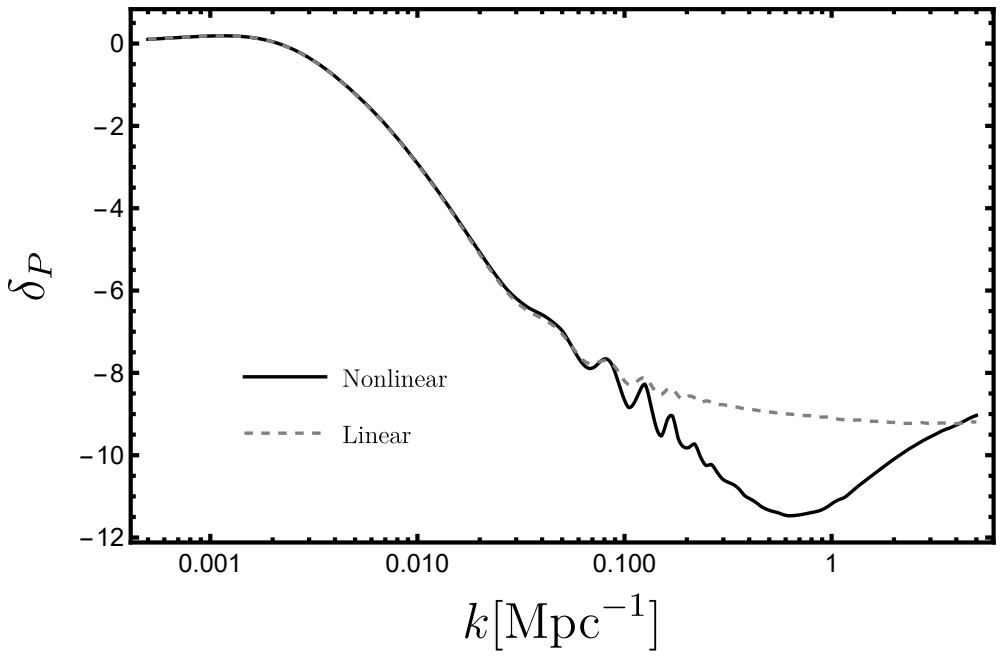

Figure 1.7: Fractional difference in the matter power spectrum (in units of the fractional contribution of neutrinos to the total matter density), between cosmologies with and without massive neutrinos, as given by Eq. 1.15. The dashed gray curve is the linear theory expectation, while solid black is the fully nonlinear result with the characteristic "spoon-shaped" suppression.

of new ingredients, such as neutrino masses, in which case it reflects itself as a preference for negative neutrino masses [74], which pushes the upper bounds on the neutrino mass scale closer towards zero.

Other alternative scenarios can bring these two probes back into agreement, such as a dynamical dark energy [75], a high optical depth to reionization [76, 77] and negative spatial curvature [14]. It is not yet clear whether the discrepancy between recent CMB and galaxy surveys is due to systematics in the data, the manifestation of one (or more) of the new

tion currently is the H_0 tension [71], which refers to the more than 5σ deviation between values of the Hubble constant extracted from early-universe probes (based on the ΛCDM model), and direct late-time measurements via the distance ladder [72, 73].

ingredients discussed above, or something else entirely. Upcoming experiments will play a crucial role in addressing this issue. For now there is a high priority in developing the science case for new signatures of neutrino masses in the large-scale structure, an example of which will be discussed in Chapter 5.

1.4 Structure of the thesis

This thesis consists of my PhD work dedicated to advancing our theoretical and computational models on the formation of structures on cosmological scales, and their underlying signatures of neutrino masses. These efforts are crucial to ensure that the maximal amount of information on fundamental physics will be extracted from ongoing and upcoming astronomical observations, including a robust measurement of the neutrino mass scale which may happen in the upcoming decade.

In Chapter 2, originally published in Phys. Rev. D 104, 043512 [78], we investigate the feasibility of including massive neutrinos in N-body simulations of nonlinear structure formation at late times. We carry out a systematic study of potential biases induced by neglecting special relativistic effects (due to the Newtonian nature of the simulations), and discuss the proper setting of initial conditions to ensure that such biases are mitigated.

In Chapters 3 and 4, originally published in Phys. Rev. D 104, 083535 [79] and Phys. Rev. D 108, 023505 [80] respectively, we propose novel methods to solve for the dynamics of linear cosmological perturbations of massive neutrino species, with the aim of improving the accuracy and reducing the computational cost of standard methods. Chapter 3 introduces a reformulation of the exact system of equations in terms of the so-called Generalized Boltzmann hierarchy, which is much simpler and potentially faster than the traditional Boltzmann hierarchy. This is achieved by integrating out the momentum dependence from the outset, to arrive at a system of equations in configuration space (as opposed to phase-space in standard methods). In Chapter 4, we significantly improve on the fluid approximation for massive neutrinos by accounting for the dispersive nature of the neutrino fluid, i.e. adding a scale dependence to the sound speed of the neutrino fluid based on properties of analytic

solutions that can be obtain in certain regimes. This provides an accurate and inexpensive approach to solve the dynamics of cosmological neutrino fluctuations in the linear regime of structure formation.

In Chapter 5, originally published in JCAP 11 (2023) 036 [81], we advance on the theory modeling and complete characterization of observables associated to a new signature of neutrino masses in the nonlinear regime of large scale structure, the so-called neutrino wakes. This refers to the preferential accumulation of massive neutrino particles downstream of moving cold dark matter structures, which leaves observable signatures on three-point cross correlations of matter and galaxies. We show that a detection of this effect is potentially achievable with future experiments.

Finally, in Chapters 6 and 7, originally published in JCAP 06 (2025) 002 [82] and JCAP 02 (2025) 023 [83] respectively, we explore the foundations of perturbative methods for nonlinear structure formation on large scales. In Chapter 6 we show that the Standard Perturbation Theory (SPT) framework, which originally assumes a perfect fluid description of the cold dark matter, actually follows from perturbatively solving the coupled system of Vlasov-Poisson equations in full phase space. We also use our framework to shed new light on the emergence of deviations from an ideal fluid, parametrized via the sound speed counterterm introduced in the Effective Field Theory for Large Scale Structure (EFTofLSS). In Chapter 7 we provide an analytical estimate for the sound speed counterterm, which is usually extracted by matching the theory model to full simulations. This calculation reveals the physical content of the sound speed, and its dependence on cosmological parameters. These efforts play a key role in ensuring that the theoretical models applied in the interpretation of ongoing and future experiments are robust, and are optimal in the sense that the extraction of fundamental physics is maximized.

Chapter 2

NEUTRINOS IN N-BODY SIMULATIONS

2.1 Abstract

In the next decade, cosmological surveys will have the statistical power to detect the absolute neutrino mass scale. N-body simulations of large-scale structure formation play a central role in interpreting data from such surveys. Yet these simulations are Newtonian in nature. We provide a quantitative study of the limitations to treating neutrinos, implemented as N-body particles, in N-body codes, focusing on the error introduced by neglecting special relativistic effects. Special relativistic effects are potentially important due to the large thermal velocities of neutrino particles in the simulation box. We derive a self-consistent theory of linear perturbations in Newtonian and non-relativistic neutrinos and use this to demonstrate that N-body simulations overestimate the neutrino free-streaming scale, and cause errors in the matter power spectrum that depend on the initial redshift of the simulations. For $z_i \lesssim 100$, and neutrino masses within the currently allowed range, this error is $\lesssim 0.5\%$, though represents an up to $\sim 10\%$ correction to the shape of the neutrino-induced suppression to the cold dark matter power spectrum. We argue that the simulations accurately model non-linear clustering of neutrinos so that the error is confined to linear scales.

2.2 Introduction

Neutrino oscillation experiments have established that at least two of the neutrino mass eigenstates have a non-zero mass, giving a lower bound on the sum of all three masses as $\sum m_\nu \gtrsim 0.06$ eV, 0.1 eV for the normal and inverted hierarchies, respectively (see, e.g. [84, 85]). At present, it is very hard to determine the neutrino mass scale at these lower limits with laboratory experiments. The current constraints from the KATRIN experiment, based

on measurements of β-decay, are $m_{\nu,i} < 1.1$ eV at 90% confidence [86]. Tighter constraints on the upper bound to the sum of neutrino mass eigenvalues come from cosmological datasets, providing $\sum m_\nu \lesssim (0.11 - 0.24)$ eV [11, 87, 88], depending on the specific choice of data. Furthermore, cosmological observations point to a universe that is today dominated by dark energy and dark matter: all the evidence suggests the presence of physics beyond the standard model (see, for instance, [89] for a review).

Current and future large-scale structure surveys [90–96] will shed some light on the nature of this, as of now, mysterious new physics. Concretely, we expect for example, to be able to determine the absolute mass scale of neutrinos (and potentially the hierarchy) [97–99], further constrain the dark energy equation of state [100], and number of effective neutrino species [101]. Large-scale structure surveys will probe small scales where nonlinear gravitational evolution is important. This adds one additional degree of complexity to the mission, as we need to rely on simulations of nonlinear structure formation, typically executed by running N-body codes such as [102–108], to accurately interpret the cosmological observables. The simulations are, however, Newtonian in nature, though large-scale general relativistic (GR) effects can be fully recovered in the case of a dark-matter-only universe [34, 35, 109]. Developing a fully-relativistic N-body code is an active area of research [32, 110, 111].

In this work we will consider limitations to treating neutrinos, implemented as N-body particles, in Newtonian codes. Our analysis relies solely on individual particles evolving with the Newtonian equation of motion. An example of this in the literature would be the codes described in [112–116], all based on GADGET-3, which is itself an improved version of GADGET-2 [117, 118]. The discussion in this paper is restricted to simulations that treat neutrinos as N-body particles, as opposed to a hydrodynamic approach, e.g. [103, 119].

We will be interested in special relativistic (SR) effects that are neglected in a Newtonian treatment of neutrinos. To illustrate the necessity of considering SR effects for neutrinos, consider an individual neutrino of mass $m_\nu = 0.05$eV, the minimum value required for at least one state by neutrino oscillation data, and a value just below that set by current limits on the sum of neutrino mass states [85, 87]. For a simulation starting at redshift

of $z_i \approx 100$, the mass-to-temperature ratio at the initial redshift is $m_\nu/(1+z_i)T_{\nu,0} \approx 3$, where $T_{\nu,0} \approx 1.95K \approx 1.7 \times 10^{-4}$eV is the temperature of relic neutrinos today. In this circumstance, one should worry that evolving neutrinos with the Newtonian limit of the geodesic equation, equivalent to suppressing both GR and order $(v/c)^2$ corrections (what we call, from now on, Newtonian neutrinos), can introduce significant systematic error in the output of N-body simulations.

As we shall see, particles with large thermal velocities evolved with non-relativistic equations of motion will travel faster and further than they would if evolved with the correct, relativistic equations of motion. Consequently, N-body simulations with neutrinos will overestimate the neutrino free-streaming scale, $\lambda_{\rm fs} \sim v_\nu/(aH)$, where v_ν is the neutrino velocity dispersion and H is the Hubble rate. This overestimation disappears at late times, as $z \to 0$, but the neutrino horizon, $\lambda_{\rm h} = \int_{z_i} d(\ln a)\lambda_{\rm fs}$, which depends on the entire history of the free-streaming scale, will be significantly overestimated, even as $z \to 0$. This shift on the free-streaming scale could have a perceptible impact on the usual neutrino-induced suppression to the growth of structure [64].

To study the impact of these errors on the evolution of matter perturbations, we develop the exact linear-theory evolution of inhomogeneities in the distribution of Newtonian neutrinos. We derive the usual fluid-approximation for Newtonian neutrinos (FA) achieved by introducing an effective sound speed, as well as leading-order corrections that include the effects of a non-zero shear stress. For $z \lesssim 100$, these effects impact only the small-scale oscillations in the neutrino transfer functions, and can hence be safely neglected when obtaining matter perturbations. The FA can be used to generate initial conditions for N-body simulations [120], through rescaling of the power spectrum from $z = 0$ to the simulation initial redshift z_i.

The FA is missing both SR and GR corrections. In order to isolate the SR effects, which are only important for fast moving particles such as neutrinos, we also develop a non-relativistic fluid approximation that ignores only SR terms that are missing for neutrinos, but keeps GR terms, and therefore produces the correct evolution of both cold dark matter and

non-relativistic neutrino perturbations on large-scales, by default. By comparing our non-relativistic fluid approximation to the correct fully-relativistic one, we quantify the errors in the usual neutrino-induced suppression to the growth of structure, caused by treating neutrinos as non-relativistic particles.

In Newtonian simulations of a dark-matter-only universe, these GR effects we are including can be obtained with an adjustment of initial and final displacements of particles [34, 109]. In [35], it was realized that this is no coincidence: As long as the output of the simulations is understood in terms of the so-called N-body gauge, GR effects can be fully recovered, i.e. they can be accommodated into a gauge transformation (also see [121]). In this scenario, one can further include the effect of light neutrinos and photons, assuming a linear evolution for these components [122–124]. The N-body gauge approach can also be generalized to cosmologies with massive neutrinos, where the CDM+ν center-of-mass motion is followed by N-body simulations [125]. Relativistic corrections can also be studied in beyond ΛCDM scenarios [126].

This paper is organized as follows. In 2.3, we review the Newtonian limit of the geodesic equation in an isotropic universe, and present expressions for the background energy density and pressure of both Newtonian and relativistic relic neutrinos. In 2.4, we present self-consistent linear perturbation theories of neutrino and cold dark matter (CDM) particles evolved according to the Newtonian and non-relativistic ($(v/c)^2 \ll 1$) equations of motion. In 2.5, we present numerical results using the linear theories from 2.4, to compare the CDM transfer functions in the presence of non-relativistic neutrinos, to the same quantities evolved with the fully-relativistic equations of motion. We also compare the linear theories from 2.4 with the cosmic linear anisotropy solving system (CLASS), in order to illustrate the additional sources of systematic error, in N-body simulations, that are not being accounted for in our analysis (such as radiation). In 2.6 we discuss the implications of our results for N-body simulations with neutrinos implemented as particles, and in 2.7 we summarize our findings. In an appendix 2.8, we present the complete derivation of fluid equations for Newtonian and non-relativistic neutrino components. While our study is motivated by

neutrinos co-evolving with cold dark matter (CDM) and baryons (b), the analysis makes no assumptions about the particle nature of either component, only that they are non-interacting. In what follows we may therefore refer to the cold non-relativistic component as dust and the non-cold dark matter component as NCDM (following the notation of CLASS [127]).

2.3 Dynamics of Newtonian neutrinos: Background

Let us start our analysis at the level of background. The relativistic dynamics follows from the geodesic equation in a Friedmann-Robertson-Walker universe with line element

$$ds^2 = -dt^2 + a^2(t)d\vec{x}^2\,, \tag{2.1}$$

$$= a^2(\tau)\left(-d\tau^2 + d\vec{x}^2\right)\,. \tag{2.2}$$

The geodesic equation is given by

$$\frac{d\vec{v}}{dt} + \left(H + \frac{1}{\gamma}\frac{d\gamma}{dt}\right)\vec{v} = 0\,, \tag{2.3}$$

where $\vec{v} = ad\vec{x}/dt$ is the peculiar velocity, $\vec{x}$ is the comoving position vector, $1/\gamma^2 = 1 - v^2$, and $H = a^{-1}da/dt$ is the Hubble rate. The evolution of a Newtonian component arises from the assumption $H \gg \gamma^{-1}d\gamma/dt = \gamma^2\vec{v}\cdot d\vec{v}/dt$, and reads,

$$\frac{d\vec{v}}{dt} + H\vec{v} = 0\,. \tag{2.4}$$

This implies $v \propto a^{-1}$. In contrast, the correct equation of motion for a relativistic particle, Eq. (2.3), will produce $p = \gamma m v \propto a^{-1}$, where p is the momentum, so that the evolution of v is more complicated.

In Newtonian N-body simulations, neutrinos are given a thermal velocity that is obtained by sampling from a Fermi-Dirac distribution,

$$f_0(p) = \frac{2}{(2\pi)^3}\frac{1}{e^{\frac{p}{T_i}} + 1}\,, \tag{2.5}$$

at the initial redshift z_i, where $p = mv$ is the nonrelativistic expression for the physical momentum, and $T_{\nu,i} = (1 + z_i)T_{\nu,0}$ is the neutrino temperature at z_i [1]. Then, under the Newtonian dynamics $p = mv \propto a^{-1}$, as it would if the correct relativistic expression were used, and the distribution retains the form of a Fermi-Dirac distribution, with a temperature that scales as $T_\nu \propto a^{-1}$.

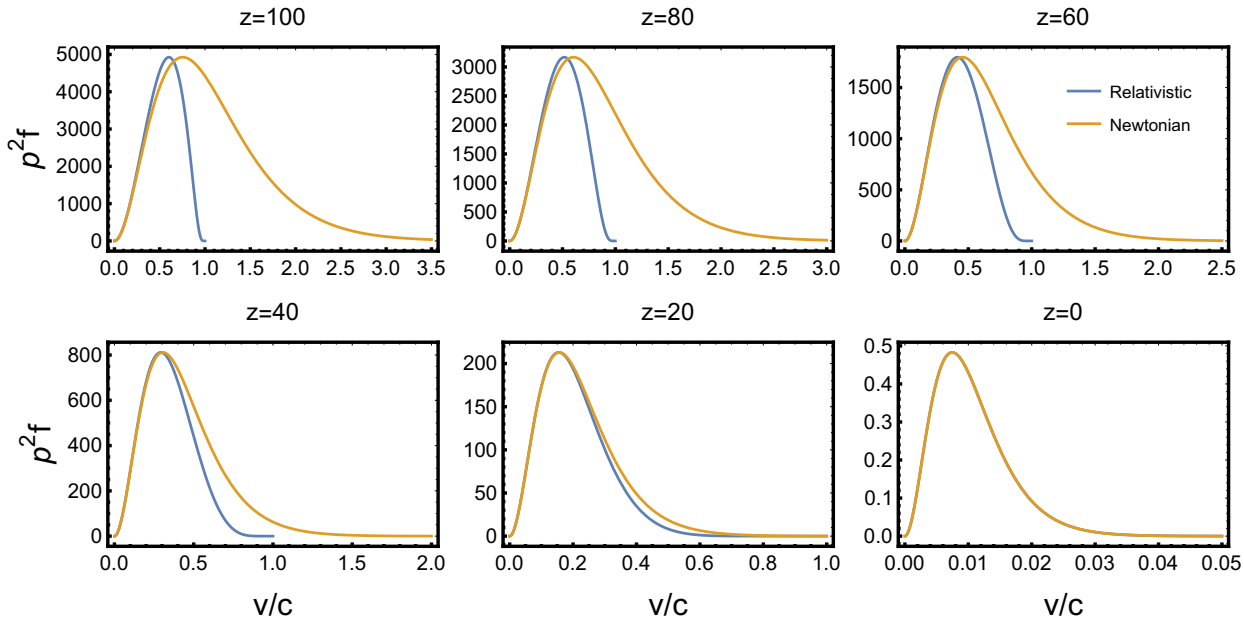

Figure 2.1: Evolution of the velocity distribution function, for $m_\nu = 0.05$eV and varying redshift. A significant portion of particles, in the simulation box, are superluminal for $z \gtrsim 50$. The differences in the velocity distribution function shown above do not depend on z_i, the initial redshift of the simulations.

[1]Note that in converting the exact momentum distribution in Eq. (2.5) into a velocity distribution function at z_i, one could use $p = mv$, as we have assumed, or solve for v in the equation $p = \gamma mv$. We assume the choice $p = mv$, since it is the one made in (at least some) Newtonian codes, e.g. [128]. As shown in Figure 2.1, this leads to a velocity distribution function that deviates from its relativistic counterpart at the initial redshift, while agreeing with the latter today, i.e. as $z \to 0$. If the choice $p = \gamma mv$ was made, then there would be no difference at the initial redshift, by construction, at the expense of getting the $z \to 0$ limit wrong, given the Newtonian evolution.

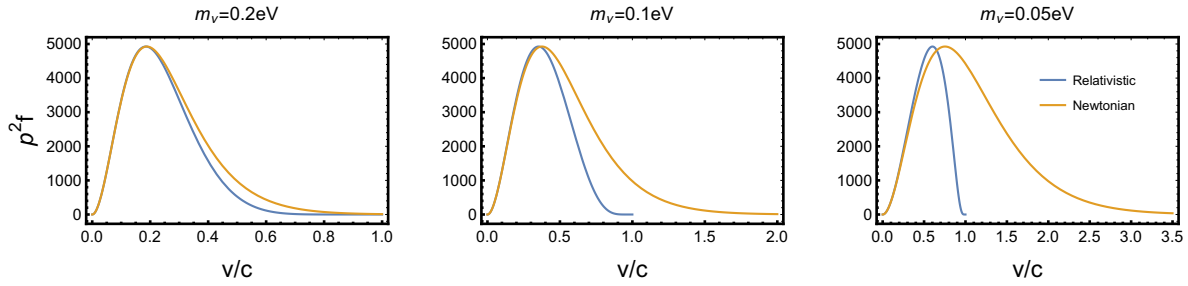

Figure 2.2: The velocity distribution function at $z = 100$ for different values of m_ν. The two distributions approach one another for $m_\nu \gtrsim 0.2$eV.

In Figures 2.1 and 2.2 we plot the velocity distribution function, and compare to its fully relativistic counterpart where $p = \gamma m v$ is used to define the physical momentum. A few comments are in order. First notice that, for a neutrino mass of $m_\nu = 0.05$eV and $z \gtrsim 50$, a significant portion of the particles in the simulation box will be superluminal. This happens at any redshift for which $m_\nu/3T_\nu \lesssim 2$. On the other hand, for $m_\nu \gtrsim 0.2$eV, the difference in the distributions is rather small already at $z = 100$, so we expect the Newtonian limit of the geodesic equation to be a good approximation. This remains true for $m_\nu/3T_\nu \gtrsim 4$. As we shall see, the difference in the velocity distribution functions, shown in Figures 2.1 and 2.2, can potentially introduce errors in both linear and nonlinear clustering of neutrinos.

It is also clear, from Figures 2.1 and 2.2, that the velocity dispersion v_ν of Newtonian neutrinos,

$$v^2_{\nu,\mathrm{New}} = \frac{\int_0^\infty dq q^2 f_0(q) \left(\frac{q}{ma}\right)^2}{\int_0^\infty dq q^2 f_0(q)} \tag{2.6}$$

will be bigger than that of relativistic neutrinos,

$$v^2_{\nu,\mathrm{rel}} = \frac{\int_0^\infty dq q^2 f_0(q) \left(\frac{q}{ma}\right)^2 \frac{1}{1+\left(\frac{q}{ma}\right)^2}}{\int_0^\infty dq q^2 f_0(q)} \tag{2.7}$$

We now consider a fluid description of both Newtonian and relativistic neutrinos. For

three relativistic neutrinos of equal mass, the energy density and pressure are given by,

$$\rho_{\rm rel} = 4\pi N_\nu m a^{-3} \int_0^\infty dq q^2 f_0(q) \sqrt{1 + \left(\frac{q}{ma}\right)^2}\,, \tag{2.8}$$

$$P_{\rm rel} = \frac{4\pi N_\nu}{3} m a^{-3} \int_0^\infty dq q^2 f_0(q) \Big(\frac{q}{ma}\Big)^2 \frac{1}{\sqrt{1 + \left(\frac{q}{ma}\right)^2}}\,, \tag{2.9}$$

where $q = pa$, and $N_\nu = 3$ is the degeneracy factor. For a Newtonian component we have,

$$\rho_{\rm New} = 4\pi N_\nu m a^{-3} \int_0^\infty dq q^2 f_0(q)\,, \tag{2.10}$$

$$P_{\rm New} = \frac{4\pi N_\nu}{3} m a^{-3} \int_0^\infty dq q^2 f_0(q) \Big(\frac{q}{ma}\Big)^2\,. \tag{2.11}$$

In order to define the free-streaming scale, we also introduce the adiabatic sound speed,

$$c_g^2 = \frac{\dot{P}}{\dot{\rho}}\,, \tag{2.12}$$

where $\dot{P}$ is the time-derivative of the neutrino fluid pressure, and $\dot{\rho}$ the time-derivative of the energy density. The relative difference in the Newtonian and relativistic adiabatic sound speeds is presented in Figure 2.3. From this it follows at once that the instantaneous free-streaming scale [2]

$$\lambda_{\rm fs} \equiv \frac{c_g}{aH}\,, \tag{2.13}$$

is overestimated for Newtonian neutrinos and therefore $k_{\rm fs} = 2\pi/\lambda_{\rm fs}$ is underestimated in the simulations. Since $c_{g,\rm New} \to c_{g,\rm rel}$ as $z \to 0$, the free-streaming is correctly reproduced at late times. This difference in free-streaming scales is important, as $\lambda_{\rm fs}$ defines the length scale below which neutrino perturbations get completely washed out. Also relevant is the neutrino (fluid) horizon,

$$\lambda_{\rm h} = \int_{z_i} \frac{dt}{a}\, c_g(a) = \int_{z_i} d(\log a) \lambda_{\rm fs}\,, \tag{2.14}$$

or in k-space $k_{\rm h} = 2\pi/\lambda_{\rm h}$. This integral is dominated by its lower limit, where Newtonian and relativistic free-streaming scales are most different, so the underestimation in the horizon

[2] In order to define the free-streaming scale, the (adiabatic) sound speed should be used instead of the velocity dispersion. Note that $c_{g,\rm New} = \frac{\sqrt{5}}{3} v_{\nu,\rm New}$, while $c_{g,\rm rel} \approx \frac{\sqrt{5}}{3} v_{\nu,\rm rel}$ in the non-relativistic regime (see e.g. the appendix A of [129]).

scale persists in the limit $z \to 0$. This defines the length scale above which free-streaming is no longer relevant, with the neutrino component behaving like cold matter. See [130, 131] for short reviews on the effect of neutrinos in cosmology.

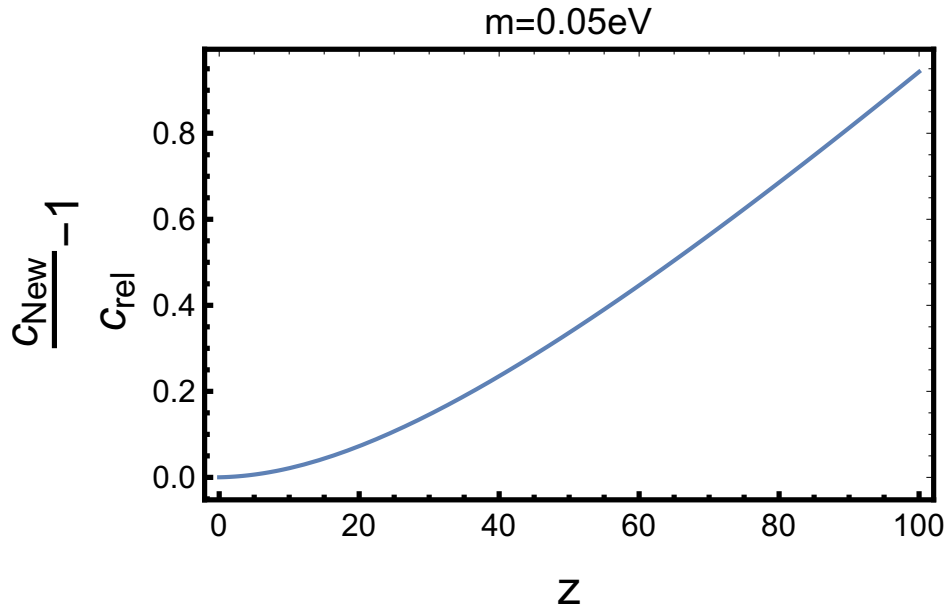

Figure 2.3: Relative difference between Newtonian and relativistic adiabatic sound speeds for a neutrino mass of $m_\nu = 0.05\text{eV}$.

It is clear from the above that the Newtonian neutrino energy density will not have the correct evolution with time. This can potentially impact the evolution of the Hubble rate as a function of time through the Friedmann equation as well. But, in N-body simulations, the background energy densities, and therefore the Hubble rate, are not computed from the dynamics of N-body particles. Moreover, one can straightforwardly provide a tabulated set of values for $H(a)$ computed from the correct relativistic dynamics [120, 132], i.e.

$$\left(\frac{H}{H_0}\right)^2 = \Omega_{\mathrm{r},0}a^{-4} + \Omega_{\mathrm{d},0}a^{-3} + \Omega_{\Lambda,0} + \frac{\rho_{\mathrm{rel}}(a)}{3M_{\mathrm{pl}}^2 H_0^2}\,. \tag{2.15}$$

where $M_{\mathrm{pl}}^2 = 1/8\pi G$, $H_0 = 100h\,\mathrm{km/s/Mpc}$ is the Hubble parameter today, $\Omega_{\mathrm{r},0}$ is the radiation density today, $\Omega_{\mathrm{d},0}$ is the energy density in cold dark matter and baryons today,

$\Omega_{\Lambda,0}$ is the energy density in vacuum, and $\rho_{\rm rel}$ is the relativistic energy density of the relic neutrinos, as computed from Eq. (2.8).

The differences in the Newtonian and correct relativistic neutrino velocity dispersions, established in Figures 2.1-2.3, will impact the evolution of neutrino perturbations in N-body simulations. This could limit the accuracy to which N-body simulations can model neutrinos, at both the linear and non-linear regimes.

2.4 Dynamics of Newtonian neutrinos: Linear perturbations

In this section we will present a self-consistent theory of linear perturbations of neutrinos and dust evolved according to non-relativistic equations of motion. As we shall see, this will amount to taking two independent limits of the fully relativistic equations for linear perturbations. First, as in the last section, we will take the limit $(v/c)^2 \ll 1$ in the equation of motion for neutrinos, as well as for sources to the Poisson equation. Second, to make contact with N-body simulations and the formalism used to generate their initial conditions, we present in Sec. 2.4.2 evolution equations ignoring time derivatives of metric perturbations that appear in the linearized geodesic equation, but are not included in N-body codes. These dropped terms are referred to as *GR-terms*, and their particular form depends on the choice of gauge for the perturbed metric. In the N-body gauge they disappear entirely for a universe comprised of only dust and Λ, but cannot be completely eliminated in the presence of relativistic particles [35]. Since our primary interest is to study the impact of SR effects ($v/c \sim 1$), we will work in the Newtonian gauge, where the interpretation of variables is more straightforward, and present final results that include the GR terms.

2.4.1 Framework

The full details of our calculations of linear perturbations of neutrinos and dust are left to Appendix 2.8, but we shall outline our approach here before proceeding. To study the evolution of linear perturbations in dust and neutrinos we start from the collisionless Boltzmann

Equation[3],

$$\frac{df}{d\tau} = \frac{\partial f}{\partial \tau} + \frac{dx^i}{d\tau}\frac{\partial f}{\partial x^i} + \frac{dq^i}{d\tau}\frac{\partial f}{\partial q^i} = 0\,, \tag{2.16}$$

which governs the dynamics of the distribution function in phase space. In the above, $\vec{q} = a\vec{p}$, with $\vec{p}$ the proper momentum and τ the conformal time. The distinction between neutrinos and dust evolved as Newtonian particles, as is done in N-body codes, or as particles subject to non-relativistic or fully-relativistic dynamics, amounts to implementing different definitions of particle momentum and equations of motion (that is, different expressions for $dq^i/d\tau$ and $dx^i/d\tau$ in Eq. (2.16)). These differences will lead to different evolution equations for $f(x, q, \tau)$ that capture how a distribution of particles would evolve, subject to Newtonian, non-relativistic, or relativistic dynamics.

In the following subsections we will use this approach to derive several different systems of equations for the linear evolution of neutrinos and dust. While it is a bit cumbersome, we will use subscripts to distinguish between the quantities that satisfy the different evolution equations. We will first consider strict Newtonian evolution, as implemented in N-body codes, these quantities being identified by the subscript $_{\text{N-body}}$. Subsequently, we will impose a fluid approximation to the Newtonian $_{\text{N-body}}$ equations, and these quantities are denoted by $_{\text{FA}}$. In Sec. 2.4.4, we present a fluid approximation that ignores SR terms for neutrinos, but keeps GR terms for both neutrinos and dust. We refer to this as the non-relativistic fluid approximation, and identify variables solving those equations with $_{\text{NRF}}$. Finally, we show the usual CLASS relativistic fluid approximation for the coevolution of neutrinos and dust in Sec. 2.4.5. The variables solving these equations will be identified by a subscript $_{\text{RFE}}$. Table 2.1 shows a list of all subscripts with their underlying assumptions.

[3]For dust, of course, one does not need the full Boltzmann equation and can start directly with the equations for a pressureless, shearless fluid.

Subscript	**Assumptions**
N-body	Newtonian evolution
FA	Newtonian evolution + fluid approximation
NRF	Newtonian evolution + GR corrections + fluid approximation
RFE	Relativistic evolution + fluid approximation

Table 2.1: Different subscripts with their associated underlying assumptions. Note that Newtonian evolution + GR corrections is equivalent to dropping the SR terms (taking the $(v/c)^2 \ll 1$ limit) of the fully relativistic expressions.

2.4.2 Linear Perturbations in N-body Fluids

In N-body simulations particles evolve according to the Newtonian equation of motion,

$$\frac{d\vec{v}_{\text{N-body}}}{dt} + H\vec{v}_{\text{N-body}} = -\frac{1}{a}\vec{\nabla}\psi_{\text{N-body}} \tag{2.17}$$

where

$$\vec{v}_{\text{N-body}} = a\frac{d\vec{x}}{dt} \tag{2.18}$$

is the peculiar velocity, and $\psi_{\text{N-body}}$ is the gravitational potential, sourced only by the rest mass of the particles in the simulation,

$$k^2\psi_{\text{N-body}} = -\frac{1}{2M_{\text{pl}}^2}a^2\sum_i \delta\rho_{\text{N-body},i} \tag{2.19}$$

where $\delta\rho_{\text{N-body},i}$ is the perturbation to the mass density of species i, further ignoring the inhomogeneities in the local volume. As shown in Appendix 2.8, one can derive exact fluid

equations for a system of non-interacting particles evolving according to Eq. (2.17) and Eq. (2.19).

For Newtonian neutrinos in N-body simulations we have,

$$\delta'_{\nu,\text{N-body}} = -\theta_{\nu,\text{N-body}}\,, \tag{2.20a}$$
$$\theta'_{\nu,\text{N-body}} = -\mathcal{H}\theta_{\nu,\text{N-body}} + \frac{\delta P_{\nu,\text{N-body}}}{\delta\rho_{\nu,\text{N-body}}}k^2\delta_{\nu,\text{N-body}} - k^2\sigma_{\nu,\text{N-body}} + k^2\psi_{\text{N-body}}\,,$$

along with other evolution equations for higher order multipole moments of the perturbed distribution function. Here $' = d/d\tau$, $\delta_{\nu,\text{N-body}} = \delta\rho_{\nu,\text{N-body}}/\rho_{\nu,\text{New}}$, with $\delta\rho_{\nu,\text{N-body}}$ the perturbation to the neutrino mass density, $\theta_{\nu,\text{N-body}}$ is the velocity divergence, $\delta P_{\nu,\text{N-body}}$ the perturbation to the pressure, $\sigma_{\nu,\text{N-body}}$ the anisotropic stress.

The dust component evolves according to the same equations in the limit of no anistropic stress and no pressure,

$$\delta'_{d,\text{N-body}} = -\theta_{d,\text{N-body}}\,, \tag{2.21a}$$
$$\theta'_{d,\text{N-body}} = -\mathcal{H}\theta_{d,\text{N-body}} + k^2\psi_{\text{N-body}}\,. \tag{2.21b}$$

In Eq. (2.20) and Eq. (2.21), $\psi_{\text{N-body}}$ is the potential computed from the Poisson equation sourced by the N-body fluids,

$$k^2\psi_{\text{N-body}} = -\frac{1}{2M_{\text{pl}}^2}a^2(\rho_d\delta_{d,\text{N-body}} + \rho_{\nu,\text{New}}\delta_{\nu,\text{N-body}})\,. \tag{2.22}$$

Notice that this is sourced by the Newtonian expression for the neutrino energy density, Eq.(2.10), i.e. it just includes the rest mass of neutrino particles.

As described in Appendix 2.8, each of these quantities is computed self-consistently assuming a Newtonian treatment of each of the neutrino properties. This system of equations should therefore describe the exact linear evolution of neutrinos and dust within N-body simulations. To proceed, of course, one needs additional expressions for $\delta P_{\nu,\text{N-body}}/\delta\rho_{\nu,\text{N-body}}$ and $\sigma_{\nu,\text{N-body}}$ to close the system of equations.

2.4.3 Two-fluid approximation (FA) for N-body neutrinos and dust

The standard two-fluid approximation for neutrinos and dust is obtained from the N-body equations, in Sec. 2.4.2, by setting the anisotropic stress to zero, $\sigma_{\nu,\text{N-body}} \approx 0$, and ratio of the pressure and energy perturbations to the Newtonian adiabatic sound speed squared, $\delta P_{\nu,\text{N-body}}/\delta\rho_{\nu,\text{N-body}} \approx c^2_{g,\text{New}}$, obtained from Eqs. (2.10) and (2.11). This yields,

$$\delta'_{\nu,\text{FA}} = -\theta_{\nu,\text{FA}}\,, \tag{2.23a}$$

$$\theta'_{\nu,\text{FA}} = -\mathcal{H}\theta_{\nu,\text{FA}} + c^2_{g,\text{New}}k^2\delta_{\nu,\text{FA}} + k^2\psi_{\text{FA}}\,. \tag{2.23b}$$

This is precisely the fluid approximation, used to generate initial conditions for N-body simulations, by rescaling of the matter power spectrum from $z = 0$ to the initial simulation redshift z_i [120]. Initial particle positions and velocities are obtained with the Zeldovich approximation [133], or the more accurate higher-order Lagrangian perturbation theory schemes [134, 135]. Note that our derived sound speed, as follows from Eq. (2.5), Eq. (2.12), Eq. (2.10), and Eq. (2.11),

$$c^2_{g,\text{New}} = \frac{25}{3}\frac{\zeta(5)}{\zeta(3)}\Big(\frac{T_{0,\nu}}{m_\nu}\Big)^2(1+z)^2 \approx 7.19\Big(\frac{T_{0,\nu}}{m_\nu}\Big)^2(1+z)^2 \tag{2.24}$$

is the expression used in [120, 136]. The dust component in Eq. (2.21) is already treated like a presureless fluid, so the FA equations are identical to the N-body ones

$$\delta'_{d,\text{FA}} = -\theta_{d,\text{FA}}\,, \tag{2.25a}$$

$$\theta'_{d,\text{FA}} = -\mathcal{H}\theta_{d,\text{FA}} + k^2\psi_{\text{FA}}\,. \tag{2.25b}$$

Similarly, the only modification to the Poisson equation in Eq. (2.22) is to change the source terms to the expressions obtained from the FA,

$$k^2\psi_{\text{FA}} = -\frac{1}{2M^2_{\text{pl}}}a^2(\rho_d\delta_{d,\text{FA}} + \rho_{\nu,\text{New}}\delta_{\nu,\text{FA}})\,. \tag{2.26}$$

The system of Eqs. (2.23), (2.25) and (2.26) is the so-called *two fluid approximation*. What we have then is a first principles derivation of the two-fluid approximation, starting from the Boltzmann equation, and assuming the Newtonian evolution, given by Eq. (2.17), of individual dust and neutrino particles.

2.4.4 Non-relativistic fluid (NRF) equations for neutrinos and dust

To study the impact of treating fast-moving neutrinos as non-relativistic particles, we will go beyond the two-fluid approximation in two ways. First, we will include the anisotropic stress, which should physically be present for a neutrino fluid modeled by N-body particles and allows us to obtain a non-relativistic analog of the CLASS fluid approximation [127]. Including anisotropic stress also enables us to check that neutrino shear stress has a negligible impact on matter perturbations at late times, which validates the use of the FA to generate initial conditions for the simulations. Second, since our concern here is on SR effects, we will keep all GR terms that are missing in the two-fluid approximation of Sec. 2.4.3. This second choice allows the expressions in this section to correctly reproduce the relativistic dynamics on large scales, as $z \to 0$ (see 2.5). This is in contrast to the FA equations in Sec. 2.4.3, which disagree with the exact linear theory expressions, on the large scales, even for dust [120].

We will work with the Newtonian gauge metric,

$$\begin{aligned} ds^2 &= -(1+2\psi)dt^2 + a^2(t)(1-2\phi)d\vec{x}^2\,, \\ &= a^2(\tau)\left[-(1+2\psi)d\tau^2 + (1-2\phi)d\vec{x}^2\right]\,. \end{aligned} \tag{2.27}$$

In the non-relativistic (NR) limit ($(v/c)^2 \ll 1$), the particle equation of motion (i.e. the geodesic equation) is,

$$\frac{d\vec{v}}{dt} + \left(H - \dot{\phi}\right)\vec{v} + \frac{1}{a}\nabla\psi = 0\,. \tag{2.28}$$

where the peculiar velocity is now given by,

$$\vec{v} = a(1-\phi-\psi)\frac{d\vec{x}}{dt}\,. \tag{2.29}$$

and includes local inhomogeneities in both time and length intervals. Working with the above equation of motion, along with the Newtonian expressions for the energy density and pressure in Eq. (2.10) and Eq. (2.11), produces the following fluid approximation for non-relativistic

neutrinos,

$$\delta'_{\nu,\mathrm{NRF}} = -\theta_{\nu,\mathrm{NRF}} + 3\phi'_{\mathrm{NRF}}\,, \tag{2.30a}$$

$$\theta'_{\nu,\mathrm{NRF}} = -\mathcal{H}\theta_{\nu,\mathrm{NRF}} + c^2_{g,\mathrm{New}}k^2\delta_{\nu,\mathrm{NRF}} - k^2\sigma_{\nu,\mathrm{NRF}} + k^2\psi_{\mathrm{NRF}}\,, \tag{2.30b}$$

$$\sigma'_{\nu,\mathrm{NRF}} = -\left(2\mathcal{H} + \frac{3}{\tau}\right)\sigma_{\nu,\mathrm{NRF}} + 8\frac{w_{\mathrm{New}}}{1+w_{\mathrm{New}}}c^2_{g,\mathrm{New}}\theta_{\nu,\mathrm{NRF}}\,, \tag{2.30c}$$

where w_{New} and $c_{g,\mathrm{New}}$ are the equation of state and adiabatic sound speed, as computed using the Newtonian expressions for pressure and energy density in Eqs. (2.10) and (2.11).

The dust component evolves according to

$$\delta'_{d,\mathrm{NRF}} = -\theta_{d,\mathrm{NRF}} + 3\phi'_{\mathrm{NRF}}\,, \tag{2.31a}$$

$$\theta'_{d,\mathrm{NRF}} = -\mathcal{H}\theta_{d,\mathrm{NRF}} + k^2\psi_{\mathrm{NRF}}\,. \tag{2.31b}$$

And finally, the gravitational potentials are subject to,

$$k^2\phi_{\mathrm{NRF}} + 3\mathcal{H}(\phi'_{\mathrm{NRF}} + \mathcal{H}\psi_{\mathrm{NRF}}) = -\frac{1}{2M^2_{\mathrm{pl}}}a^2(\rho_d\delta_{d,\mathrm{NRF}} + \rho_{\nu,\mathrm{New}}\delta_{\nu,\mathrm{NRF}})\,, \tag{2.32a}$$

$$k^2(\phi_{\mathrm{NRF}} - \psi_{\mathrm{NRF}}) = \frac{3}{2M^2_{\mathrm{pl}}}a^2\rho_{\nu,\mathrm{New}}\sigma_{\nu,\mathrm{NRF}}\,. \tag{2.32b}$$

For $z \lesssim 100$, it suffices to apply the approximation $\phi_{\mathrm{NRF}} \approx \psi_{\mathrm{NRF}}$, instead of using Eq. (2.32b) [4] Let us call Eqs. (2.30), (2.31) and (2.32) the non-relativistic fluid (NRF) equations. To summarize, the NRF equations describe co-evolution of dust and neutrinos evolving according to the non-relativistic limit ($(v/c)^2 \ll 1$) of both the geodesic and Einstein equations.

[4]Setting $\phi = \psi$ at late times is a simplifying approximation that is implictly made in N-body codes. Indeed, we can argue that the relative difference between ϕ and ψ is negligible: On the small scales, we can combine Eqs. (2.32) and (2.32b) to arrive at $1 - \psi/\phi \propto (\rho_\nu/\rho_{\mathrm{total}})(\sigma_\nu/\delta_{\mathrm{total}})$, i.e. the relative difference is a product of two small quantities, and is vanishingly small. For $M_\nu = 0.15$eV, we found that $|1 - \psi_{\mathrm{CLASS}}/\phi_{\mathrm{CLASS}}| \lesssim \mathcal{O}(10^{-4})$ for $k \gtrsim 0.01$ Mpc^{-1}, for all redshift. On the large scales, the neutrino anisotropic stress is vanishing at low redshift, and indeed we find that the error associated to having $\phi \neq \psi$ is subleading in comparison to the systematic error associated to neglecting radiation perturbations at $z \sim 100$. This argument can be applied in both cases of non-relativistic and relativistic neutrinos.

2.4.5 *Relativistic fluid equations (RFE) for neutrinos and dust*

The equations in Sec. 2.4.4 can be compared with the fully-relativistic CLASS non-cold dark matter (NCDM) fluid approximation [127]. These expressions continue to use the Newtonian gauge metric, Eq. (2.27), but keep all terms $\mathcal{O}((v/c)^2)$. In this case the particle equation of motion is,

$$\frac{1}{\gamma}\frac{d}{dt}\left(\gamma\vec{v}\right) + \left(H - \dot{\phi}\right)\vec{v} + \frac{1}{a}\nabla\psi + \frac{1}{a}\vec{v}\times\left(\nabla\phi\times\vec{v}\right) = 0\,, \tag{2.33}$$

where v continues to be defined through Eq. (2.29). The neutrino fluid equations are,

$$\delta'_{\nu,\mathrm{RFE}} = -(1+w_{\mathrm{rel}})(\theta_{\nu,\mathrm{RFE}} - 3\phi'_{\mathrm{RFE}}) - 3\mathcal{H}(c^2_{g,\mathrm{rel}} - w_{\mathrm{rel}})\delta_{\nu,\mathrm{RFE}}\,, \tag{2.34a}$$

$$\theta'_{\nu,\mathrm{RFE}} = -\Big[\mathcal{H}(1-3w_{\mathrm{rel}}) + \frac{w'_{\mathrm{rel}}}{1+w_{\mathrm{rel}}}\Big]\theta_{\nu,\mathrm{RFE}} + \frac{c^2_{g,\mathrm{rel}}}{1+w_{\mathrm{rel}}}k^2\delta_{\nu,\mathrm{RFE}} - k^2\sigma_{\nu,\mathrm{RFE}} + k^2\psi_{\mathrm{RFE}}\,, \tag{2.34b}$$

$$\sigma'_{\nu,\mathrm{RFE}} = -\Big\{\Big[(2-3w_{\mathrm{rel}}) - \frac{\mathcal{P}_{\mathrm{rel}}}{P_{\mathrm{rel}}}\Big]\mathcal{H} + \frac{w'_{\mathrm{rel}}}{1+w_{\mathrm{rel}}} + \frac{3}{\tau}\Big\}\sigma_{\nu,\mathrm{RFE}} + 8\frac{w_{\mathrm{rel}}}{1+w_{\mathrm{rel}}}c^2_{g,\mathrm{rel}}\theta_{\nu,\mathrm{RFE}}\,, \tag{2.34c}$$

where $w_{\mathrm{rel}} = P_{\mathrm{rel}}/\rho_{\mathrm{rel}}$,

$$\mathcal{P}_{\mathrm{rel}} = \frac{4\pi}{3}ma^{-3}\int_0^\infty q^2 f_0(q)dq\Big(\frac{q}{ma}\Big)^4\frac{1}{\left[1+\left(\frac{q}{ma}\right)^2\right]^{\frac{3}{2}}}\,, \tag{2.35}$$

is a higher velocity weight background pressure, and

$$c^2_{g,\mathrm{rel}} = \frac{5}{3}\frac{w_{\mathrm{rel}}}{1+w_{\mathrm{rel}}}\Big(1 - \frac{1}{5}\frac{\mathcal{P}_{\mathrm{rel}}}{P_{\mathrm{rel}}}\Big)\,, \tag{2.36}$$

is the relativistic expression for the adiabatic sound speed squared, as follows from Eq. (2.12), Eq. (2.8), Eq. (2.9), and Eq. (2.35). We will refer to this as the relativistic fluid equations (RFE).

The dust RFE are identical to those in Eq. (2.31), with the potentials replaced by the ones sourced by the full (kinetic plus rest mass) energy density of neutrinos,

$$\delta'_{d,\mathrm{RFE}} = -\theta_{d,\mathrm{RFE}} + 3\phi'_{\mathrm{RFE}}\,, \tag{2.37a}$$

$$\theta'_{d,\mathrm{RFE}} = -\mathcal{H}\theta_{d,\mathrm{RFE}} + k^2\psi_{\mathrm{RFE}}\,. \tag{2.37b}$$

where $\phi_{\rm RFE}$, $\psi_{\rm RFE}$ are subject to,

$$k^2\phi_{\rm RFE} + 3\mathcal{H}(\phi'_{\rm RFE} + \mathcal{H}\psi_{\rm RFE}) = -\frac{1}{2M_{\rm pl}^2}a^2(\rho_d\delta_{d,\rm RFE} + \rho_{\nu,\rm rel}\delta_{\nu,\rm RFE})\,, \tag{2.38a}$$

$$k^2(\phi_{\rm RFE} - \psi_{\rm RFE}) = \frac{3}{2M_{\rm pl}^2}a^2(\rho_{\nu,\rm rel} + P_{\nu,\rm rel})\sigma_{\nu,\rm RFE}\,. \tag{2.38b}$$

That is, including the relativistic kinetic energy of neutrinos as a source to the gravitational potential. Also, and again for $z \lesssim 100$, it suffices to set $\phi_{\rm RFE} \approx \psi_{\rm RFE}$, instead of using Eq. (2.38b).

2.5 Numerical results

The overall difference between the output of Boltzmann codes, and the two-fluid approximation presented in Sec. 2.4.3, is described in [120]. Here we study the errors induced specifically from neglecting special relativistic corrections, by comparing our non-relativistic fluid equations, of Sec. 2.4.4, to the relativistic fluid equations described in Sec. 2.4.5. As discussed in Sec. 2.3, we expect errors in the NRF associated to the overestimation in the neutrino free-streaming scale $\lambda_{\rm fs}$, and corresponding to an overestimation in the suppression to the growth of structure, caused by particles moving too fast in the simulation box, as seen in Figures 2.1-2.3.

We consider the specific example of a flat $\nu\Lambda$CDM universe, with three neutrino components in the degenerate case of equals masses, with total mass M_ν. Our choice of cosmological parameters is defined by: $h = 0.67$, $\omega_{\rm d} = \Omega_{\rm d,0}h^2 = 0.1424$, $T_{\rm CMB,0} = 2.725K$, $T_{\nu,0} = 1.95K$. The neutrino masses are left as free parameters and $\Omega_{\Lambda,0}$ is determined from the constraint equation $\sum_I \Omega_{I,0} = 1$.

We will be interested in a redshift range of $z \lesssim 100$, relevant for the study of structure formation. In this case, both CDM and baryons can be treated as a single dust component, and radiation may be ignored. This last assumption leads to the introduction of significant systematic errors in the computation of observables for larger values of redshift, and on the largest scales. In order to avoid this, along with other sources of systematic error on

the large scales, initial conditions are often set in N-body simulations from the rescaling of the matter power spectrum at $z = 0$, to the simulation initial redshift z_i. This procedure enforces recovery of the predictions of the linear theory on linear scales, and low redshift, at the expense of predictivity at high redshift [120]. However, it has been argued that this rescaling is not feasible in some cases [137], including the example of cosmologies with massive neutrinos, though one is still able to choose the initial conditions in a suitable way [138]. Taking this into account, and since our goal is to compare non-relativistic with relativistic neutrinos, and not necessarily reproduce the output of a Boltzmann code (so we don't need to worry about the contribution from radiation or lack thereof), we here choose the initial conditions for the transfer functions at $z_i = 100$ or $z_i = 60$. Notwithstanding, we will comment on how a rescaling procedure (or other suitable methods to generate the initial conditions) would impact our results.

The initial conditions for dust and gravitational potential ϕ are generated in the same way for both the RFE and NRF systems, i.e. directly from the output of CLASS, at the initial redshift (the dust transfer functions, at the initial time, are set to be a weighted average of CDM and baryon transfer functions). We do the same for neutrinos in the RFE, that is

$$\delta_{\nu,\mathrm{RFE}}(z_i) = \delta_{\nu,\mathrm{CLASS}}(z_i)\,, \tag{2.39a}$$

$$\theta_{\nu,\mathrm{RFE}}(z_i) = \theta_{\nu,\mathrm{CLASS}}(z_i)\,, \tag{2.39b}$$

$$\sigma_{\nu,\mathrm{RFE}}(z_i) = \sigma_{\nu,\mathrm{CLASS}}(z_i)\,. \tag{2.39c}$$

For neutrinos in the NRF, however, we have to be more careful. This is because the non-relativistic fluid equations are solving for a different set of variables, e.g. the neutrino mass density as opposed to the energy density. For adiabatic initial conditions, the large-scale super-horizon initial conditions satisfy $\delta_{\nu,\mathrm{RFE}} \approx (1 + w_{\mathrm{rel}})\delta_{d,\mathrm{RFE}}$. On the other hand, the correct evolution of the mass density on large scales would be given by $\delta_{\nu,\mathrm{NRF}} \approx \delta_{d,\mathrm{NRF}}$. On sub-horizon scales the neutrino perturbations are no longer adiabatic. For the examples in this paper, however, we find that the solutions for neutrino perturbations on sub-horizon scales

quickly lose sensitivity to the initial conditions once the neutrinos become non-relativistic [129]. The initial conditions for the neutrino fluid then have a negligible impact on the evolution of matter perturbations on small scales.

This allows us to choose the initial conditions for neutrinos in the NRF system as follows,

$$\delta_{\nu,\mathrm{NRF}}(z_i) = \frac{\delta_{\nu,\mathrm{CLASS}}(z_i)}{1 + w_{\mathrm{rel}}(z_i)}\,, \tag{2.40a}$$

$$\theta_{\nu,\mathrm{NRF}}(z_i) = \frac{\theta_{\nu,\mathrm{CLASS}}(z_i)}{1 + w_{\mathrm{rel}}(z_i)}\,, \tag{2.40b}$$

$$\sigma_{\nu,\mathrm{NRF}}(z_i) = \frac{\sigma_{\nu,\mathrm{CLASS}}(z_i)}{1 + w_{\mathrm{rel}}(z_i)}\,, \tag{2.40c}$$

where we apply the same procedure for $\theta_{\nu,\mathrm{NRF}}$ and $\sigma_{\nu,\mathrm{NRF}}$ in order to correct for the fact that their definition differs from the relativistic expression by a factor of $(1 + w_{\mathrm{rel}})$ (see the appendix 2.8).

We are now ready to state our results. In Fig. 2.4 we compare the dust power spectra in universes where SR terms for neutrinos are included or ignored, computed using the RFE (Sec. 2.4.5) and NRF (Sec. 2.4.4), respectively. These plots deserve careful explanation. Since NRF neutrinos move faster, they wash-out power on larger scales than neutrinos evolved with the RFE, and hence we expect to see an overestimation in the suppression to the growth of structure. In terms of our curves, this implies that they should be above zero up to the horizon of Newtonian neutrinos, with a peak in between the Newtonian and relativistic horizons. In fact, notice that the peaks in the curves shift only slightly towards lower k as $z \to 0$, following the trend expected from the neutrino particle horizon, since its integral is dominated by early times when $c_{g,\mathrm{New}}$ is most different from $c_{g,\mathrm{rel}}$. Furthermore, the position of the peak shifts towards smaller scales as we increase the neutrino mass, as expected. Also notice that, at the initial redshift, the differences are zero by construction, while at later times the free-streaming scales of Newtonian and relativistic neutrinos approach one another. That explains why we see no feature on scales where both NRF and RFE neutrino perturbations get completely washed-out, and the neutrino component has no

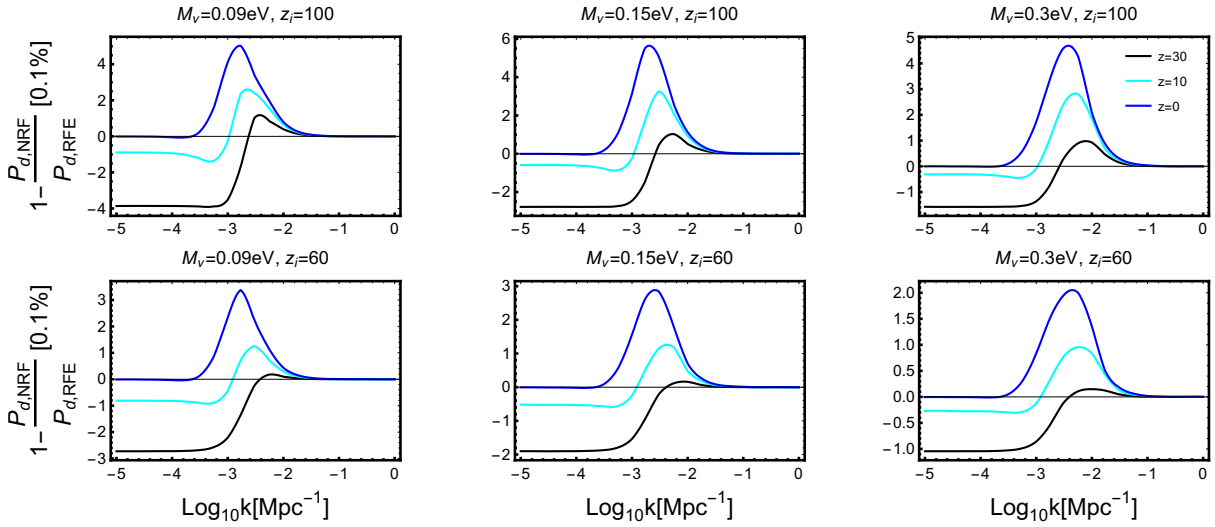

Figure 2.4: Relative difference in the dust (CDM+b) power spectrum, $P = |\delta_d|^2$, between the NRF and RFE, corresponding to Newtonian and relativistic neutrinos, respectively. The six plots are obtained combining two possible values for the initial redshift, $z_i = 60$ or $z_i = 100$, along with three possible choices for the total neutrino mass $M_\nu = 0.09$eV, $M_\nu = 0.15$eV or $M_\nu = 0.30$eV, ranging from (approximately) the smallest to the largest possible value, according to constraints from cosmology and neutrino oscillations. The shift in the neutrino free-streaming scale induces a $\lesssim$0.5% error in the linear matter power spectrum, in all cases, at around the neutrino horizon scale.

effect on dust.

The plots in Fig. 2.4 also show significant disagreement on large scales and for higher values of redshift. This is because the RFE and NRF system are solving for different variables, i.e. neutrino energy density including or ignoring the kinetic energy, respectively. On the large scales, this introduces a $\propto (1 + w_{\mathrm{rel}})$ disagreement on the neutrino density contrast transfer function, as previously discussed, which impacts the evolution of dust through gravitational coupling. This difference disappears as $w_{\mathrm{rel}} \to 0$ at late times.

Another interesting feature observed in Fig. 2.4 is how the error in the matter power

spectrum depends on the neutrino mass scale and the initial redshift of the simulation (for our exercise, this amounts to the redshift at which we start solving the NRF). In agreement with intuition, the error is larger for earlier starting redshifts z_i, since the temperature is larger in comparison to the mass at the initial redshift, and hence the NR approximation becomes worse. On the other hand, for $z_i = 100$, varying the mass seems to leave the size of the peak (approximately) unaltered. We attribute this to two competing effects: for smaller masses the NR approximation is worse, but the overall effects of neutrinos on the power spectrum are also smaller. These two effects don't always cancel each other out, as one can see from the plots with the initial redshift of $z_i = 60$.

The errors in the linear power spectrum illustrated in Fig. 2.4 are $\lesssim 0.5\%$ for the masses and initial redshifts values we consider, and are associated to an overestimation in the neutrino horizon. The effects of neutrino masses in cosmology are traditionally illustrated in the ratio of matter power spectra between cosmologies with and without neutrino masses. For scales above the neutrino horizon, $k < k_{\rm h}$, it assumes the value 1, while on small scales for which $k \gg k_{\rm h}$, below the free-streaming scale, it can be approximated as follows,

$$\frac{P_d(M_\nu)}{P_d(M_\nu = 0)} \approx 1 - 6f_\nu, \quad k \gg k_{\rm h} \tag{2.41}$$

where $\Omega_{\rm m} = \Omega_{\rm d} + \Omega_\nu$ is kept fixed, and $f_\nu = \Omega_\nu/\Omega_{\rm m}$. This is the neutrino-induced suppression to the growth of structure, and it is actually enhanced by non-linear effects [65]. Notice that massless neutrinos behave like photons, with their contribution to the energy budget of the universe becoming negligible at late times. In this case it suffices to simply concentrate all of the matter in the dust component, i.e. to set $\Omega_{\rm m} = \Omega_{\rm d}$ in the model with massless neutrinos.

The ratio of power spectra between cosmologies with and without neutrino masses at $z = 0$, and simulation initial redshift of $z_i = 100$, is plotted in Fig. 2.5 for the neutrino mass scales of $M_\nu = 0.09$eV and $M_\nu = 0.15$eV. The model with massive neutrinos is computed with both the RFE and NRF systems, while the model with massless neutrinos is computed using only the RFE. As can be seen in Fig. 2.5, the neutrino horizon is overestimated when using non-relativistic equations of motion, in agreement with Fig. 2.4. For the smallest

neutrino mass we consider, $M_\nu = 0.09$eV, this causes an up to $\sim 10\%$ correction to the shape of the neutrino-induced suppression to the cold dark matter power spectrum. For heavier masses the correction to the shape is smaller. We see no shift in the location of the transition to the constant suppression of the linear power spectrum at high k. This is because this feature is set by the free-streaming scale, and its overestimation disappears in the limit $z \to 0$.

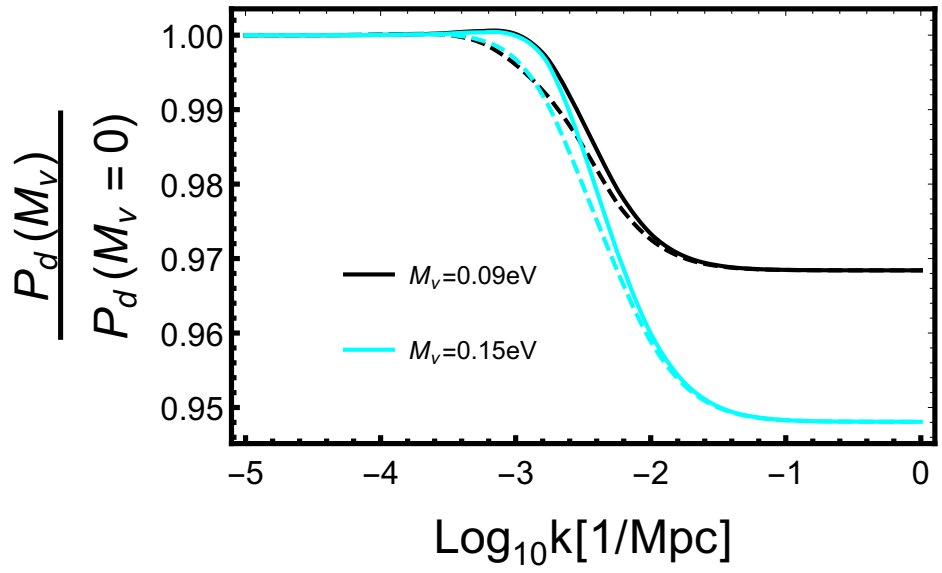

Figure 2.5: Ratio of matter (cdm+b) power spectra between models with massive and massless neutrinos, at $z = 0$, for the neutrino mass scales of $M_\nu = 0.09$eV and $M_\nu = 0.15$eV. The solid lines corresponds to massive neutrinos evolved with the RFE, while dashed lines are computed with the NRF. This also illustrates the overestimation in the neutrino sound horizon, while emphasizing that the neutrino-induced suppression to the growth of structure, in the small scales, is correctly reproduced. The overestimation in the neutrino horizon causes, however, an up to $\sim 10\%$ correction to the shape of the neutrino-induced suppression of the cold dark matter power spectrum, with the largest errors occurring for the lowest values of M_ν.

Finally, a direct comparison of the NRF and RFE with CLASS is illuminating, for it shows the impact of other sources of systematic error in N-body simulations, that are neglected in our analysis. This is presented in Fig. 2.6, for a neutrino mass scale of $M_\nu = 0.15$eV and simulation initial redshift $z_i = 100$.

For $k_1 = 3 \times 10^{-4}\mathrm{Mpc}^{-1} < k_\mathrm{h}$, the difference between the RFE and CLASS is dominated by the contribution from radiation (with a subleading effect coming from the fact that $\phi \neq \psi$), while the additional difference between the RFE and NRF is associated to having different definitions of neutrino energy densities, i.e. including or not the kinetic energy. For $k_2 = 3 \times 10^{-3}\mathrm{Mpc}^{-1} \approx k_\mathrm{h}$, the difference between the RFE and CLASS is again dominated by the contribution from radiation, while the difference between the RFE and NRF is now due to the differences in definitions of fluid properties, as before (this dominates at $z \gg 1$), but also from the overestimation of the neutrino horizon (this dominates at $z \to 0$). This explains the intersection of NRF and RFE curves as the transition between these two regimes. For $k_3 = 3 \times 10^{-2}\mathrm{Mpc}^{-1} > k_\mathrm{h}$, both the NRF and RFE reproduce CLASS exactly, in agreement with our previous results. The differences between the NRF and RFE with CLASS, found to be significant for $k \lesssim k_\mathrm{h}$ (at linear scales), can be alleviated for all redshifts when interpreting the output of the simulations in a suitable gauge [125]. An alternative procedure would be to rescale the power spectrum from $z = 0$ to the simulation initial redshift z_i [120], to generate initial conditions for the simulations. This alleviates the errors at $z \to 0$, but increase the errors at high redshift $z \gg 1$.

2.6 Implications for N-body simulations with massive neutrinos

The systematic errors associated to the overestimation of the free-streaming scale, at higher redshifts, are only present at around the scale of neutrino horizon as $z \to 0$, safely within the regime of applicability of the linear theory, as can be seen in Figs. 2.4 and 2.5. This is in agreement with the results of [124], where the authors add the effects of neutrinos in linear theory through a post-processing after running the simulations. There is, however, a procedure that enforces that the outputs of N-body simulations will match the linear theory

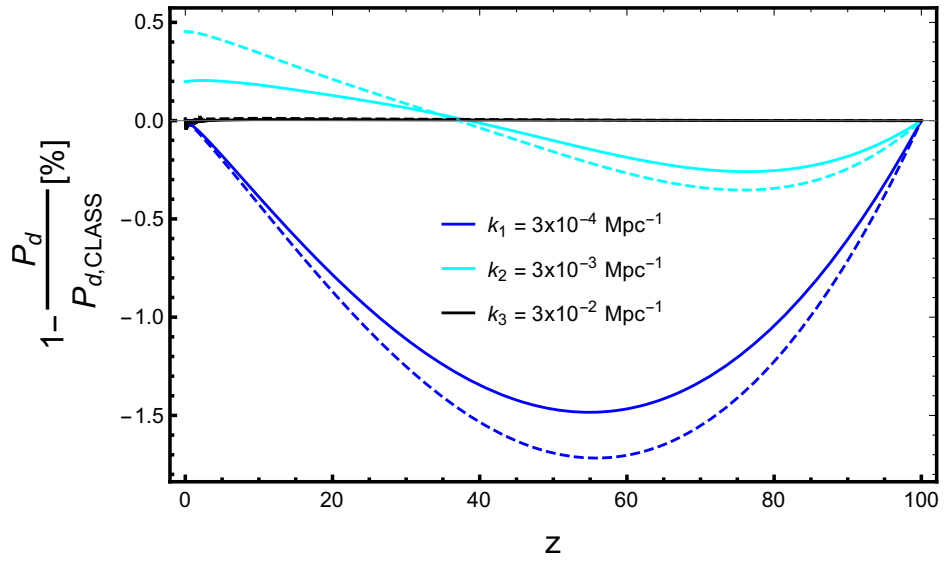

Figure 2.6: Relative difference between dust (cdm+b) power spectra, computed with the RFE (solid lines) or NRF (dashed lines), and CLASS. We consider three fixed values of k: $k_1 = 3 \times 10^{-4}\text{Mpc}^{-1} < k_\text{h}$ (blue), above the neutrino horizon, $k_2 = 3 \times 10^{-3}\text{Mpc}^{-1} \approx k_\text{h}$ (cyan), and $k_3 = 3 \times 10^{-2}\text{Mpc}^{-1} > k_\text{h}$ (black). The neutrino mass scale is fixed at $M_\nu = 0.15\text{eV}$, the simulation initial redshift at $z_i = 100$, and the relative differences are plotted as functions of redshift. For $k > k_\text{h}$, both the RFE and NRF exactly reproduce CLASS, while for $k \lesssim k_\text{h}$, the difference between RFE and CLASS is dominated by the contribution from radiation, which we have neglected.

results on linear scales (with non-linear corrections on top), i.e. the rescaling of the power spectrum, from $z = 0$ to z_i, to generate initial conditions for the simulations [120, 138] (see the discussion in Sec. 2.5). Note that this procedure can only force agreement with linear theory at single redshift, typically chosen to be $z = 0$. An alternative scenario is to interpret the outputs of the simulation in a suitable gauge, which allows N-body simulations to approximate the dynamics of both dust and neutrinos to high accuracy [125].

There is still the question of to what extent the non-linear corrections, generated by N-body simulations, are accurate. This is because the changes to the neutrino velocity

distribution, shown in Figs. 2.1 and 2.2, can also potentially cause errors in nonlinear clustering of neutrinos. However, since the relativistic and Newtonian velocity distributions match in the low-velocity end, we expect these errors to be negligible. While this is hard to compute analytically, one can get a simple estimate of the neutrinos that end up nonlinearly clustered (i.e. bound) in dark matter halos as follows. Consider a halo of mass M. Only neutrinos with $v < v_{\rm esc}$, where $v_{\rm esc}$ is the escape velocity of the halo, will end up bound. It then follows that out of all the neutrinos that encounter the halo, only a fraction

$$f_{\rm bound}(v_{\rm esc}) \propto \int_0^{p_{\rm esc}} dp\ p^2 f_0(p) \tag{2.42}$$

will end up bound to the halo (see, e.g. [139]). For a halo of mass M (comprised of dust), the total neutrino mass in the vicinity of the halo is $M_\nu = (4/3)\pi R_L^3 \rho_{\nu,\rm New}$, where $R_L = (M/(4/3\pi\rho_{\rm d}))^{1/3}$ is the Lagrangian radius of the halo. An estimate of the neutrino mass bound to a halo of mass M is then

$$M_{\nu,\rm bound} = f_{\rm bound}(v_{\rm esc})\frac{4}{3}\pi R_L^3 \rho_{\nu,\rm New} \tag{2.43}$$

$$= f_{\rm bound}(v_{\rm esc})\frac{\rho_{\nu,\rm New}}{\rho_{\rm d}} M \tag{2.44}$$

The escape velocity can be estimated by $v_{\rm esc}^2 \sim GM/R_L \sim M/M_{\rm pl}^2 R_L$. The ratio of relativistic and Newtonian neutrino masses bound to a halo then reads

$$\frac{M_{\rm rel-\nu,bound}}{M_{\rm New-\nu,bound}} = \frac{f_{\rm rel,bound}(v_{\rm esc})}{f_{\rm New,bound}(v_{\rm esc})} \tag{2.45}$$

where $p_{\rm esc} = \gamma_{\rm esc} m_\nu v_{\rm esc}$ or $p_{\rm esc} = m_\nu v_{\rm esc}$ in Eq. (2.42) for relativistic and Newtonian neutrinos, respectively. The relative difference in Eq. (2.45) is plotted as a function of the escape velocity in Fig. 2.7, for redshifts of $z = 0$ and $z = 5$, and a neutrino mass of $m_\nu = 0.05$eV. As expected, the error in non-linear clustering of neutrinos is negligible. Therefore, particle-based N-body simulations accurately model the physics of non-linear structure formation in the presence of massive neutrinos, while the overestimation in the neutrino horizon, along with other soucers of systematic error on linear scales, can be alleviated by either using the rescaling procedure to generate initial conditions for the simulations, or interpreting

the outcome of the simulations in a suitable gauge. For smaller neutrino masses, the sound horizon is pushed to even larger scales where the linear theory produces more accurate results, while for larger masses the non-relativistic approximation becomes better, and the error in the power spectrum smaller. Hence, our results hold for any simulation initial redshift of $z_i \lesssim 100$, and degenerate neutrino masses in the allowed range from neutrino oscillation experiments and constraints from cosmology. For the lightest neutrino mass states with nonzero mass ($m_\nu \approx 0.01$ eV for the minimal mass normal ordering), the errors due to neglecting SR effects on the neutrino evolution are more severe, but these light neutrinos would also comprise a smaller fraction of the neutrino energy density and so the overall errors on the perturbations of dust should be smaller.

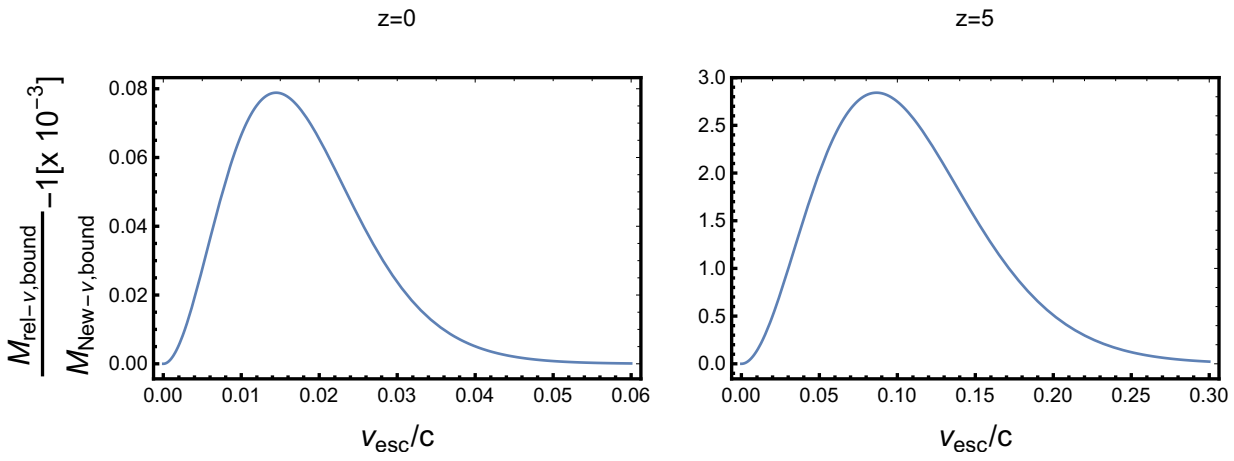

Figure 2.7: Relative difference between relativistic and Newtonian neutrino masses bound to a dark matter halo, as follows from the relativistic or Newtonian velocity distribution functions, for $m_\nu = 0.05$eV. Particle-based Newtonian N-body simulations accurately model non-linear clustering effects of neutrinos.

2.7 Conclusions

We have argued that particles with large thermal velocities, such as neutrinos, evolving with the Newtonian equation of motion travel faster and further than they would if evolved with the correct, relativistic equation of motion. This causes an overestimation in the neutrino free-streaming scale, and could potentially introduce errors in both linear and nonlinear clustering of neutrinos, in particle-based N-body simulations of structure formation. This is a special relativistic effect of neutrinos, that is neglected in N-body simulations, and has not been systematically studied.

In order to determine the impact of this on the evolution of matter perturbations, we developed the exact linear-theory evolution of inhomogeneities in the distribution of both Newtonian and non-relativistic neutrinos, where the non-relativistic equation of motion adds large scale general relativistic corrections to the Newtonian equation of motion. We apply our findings to derive the usual two-fluid approximation, that is used to generate initial conditions for N-body simulations by rescaling of the matter power spectrum from $z = 0$ to the initial simulation redshift z_i [120], along with a fluid approximation for non-relativistic neutrinos that is analogous to its fully-relativistic counterpart used in the code CLASS [127].

It was found that the overestimation in the neutrino free-streaming scale has a sub-percent impact on the linear matter (i.e. CDM +baryon) power spectrum, for neutrino mass scales in the allowed range from neutrino oscillation experiments and constraints from cosmology, and for an initial simulation redshift of $z_i \lesssim 100$. This error is at around the scale of neutrino horizon, safely within the regime of applicability of the linear theory. On the other hand, the free-streaming scale approaches the non-linear regime for higher masses and smaller redshifts, in which case the non-relativistic dynamics becomes a good approximation, and the shift in the free-streaming scale disappears in the limit $z \to 0$. As a consequence, N-body simulations accurately model non-linear clustering of neutrinos. Approaches to limit errors on the linear scales are to rescale the power spectrum from $z = 0$ to the simulation initial redshift z_i [120, 138], to generate initial conditions for the simulations that give the

correct $z = 0$ power spectrum, while having some errors at $z \gg 1$. Alternatively, one may also interpret the output of the simulations in a suitable gauge that allows the simulations to approximate the evolution of both neutrinos and CDM [125].

Particle-based implementations of neutrinos in N-body simulations are then powerful tools to accurately model the non-linear formation of structure in our universe. Including neutrinos as N-body particles has the advantage of accounting for neutrino non-linear clustering effects, with the shortcoming of having to deal with the well-known problem of shot noise, due to large thermal velocities (however, there are ways around it, e.g. [112, 125]). The other possibility, to include neutrinos as a linear component, completely misses effects of non-linear clustering of neutrinos [122]. However, such effects might be negligible for the small neutrino masses constrained by cosmology, though it has been argued that the slow tail in the distribution of neutrinos cluster strongly enough to warrant a non-linear treatment [140].

2.8 *Appendix: Anisotropies of Newtonian and non-relativistic NCDM components*

We are interested in scalar perturbations of FRW, working in the Newtonian gauge. The metric reads

$$ds^2 = a^2(\tau)[-(1+2\psi)d\tau^2 + (1-2\phi)d\vec{x}^2] \tag{2.46}$$

$$= -(1+2\psi)dt^2 + a^2(t)(1-2\phi)d\vec{x}^2 \tag{2.47}$$

The geodesic equation gives the equation of motion for a point particle as,

$$\frac{1}{\gamma}\frac{d}{dt}(\gamma\vec{v}) + \left(H - \dot{\phi}\right)\vec{v} + \frac{1}{a}\nabla\psi + \frac{1}{a}\vec{v}\times(\nabla\phi\times\vec{v}) = 0 \tag{2.48}$$

where the peculiar velocity is given by

$$\vec{v} = a(1-\phi-\psi)\frac{d\vec{x}}{dt} \tag{2.49}$$

and $1/\gamma^2 = 1 - v^2$. In the Non-relativistic (NR) limit ($(v/c)^2 \ll 1$), this reduces to

$$\frac{d\vec{v}}{dt} + \left(H - \dot{\phi}\right)\vec{v} + \frac{1}{a}\nabla\psi = 0\,. \tag{2.50}$$

Let us contrast Eq. (2.50) with Newton's second law, in the presence of a peculiar gravitational field ψ. It reads, in terms of comoving coordinates $\vec{x}(t)$,

$$\frac{d\vec{v}}{dt} + H\vec{v} = -\frac{1}{a}\vec{\nabla}\psi\,. \tag{2.51}$$

where proper (physical) coordinates are given by $\vec{r}(t) = a(t)\vec{x}(t)$, and here we define the peculiar velocity

$$\vec{v} = a\frac{d\vec{x}}{dt} \tag{2.52}$$

This is the equation of motion used in particle-based N-body simulations. It can be obtained from Eq. (2.50), after dropping the GR term $\dot{\phi}$, along with the factor of $(1-\phi-\psi)$ in Eq. (2.49), that accounts for inhomogeneities in both position and time intervals. It contributes to Eq. (2.50) with further derivatives of the potential and order $(v/c)^2$ terms. From this we conclude that Eq. (2.50) and Eq. (2.51) agree on small scales inside the horizon, in the NR limit we are considering.

We will first consider the linear-theory evolution of a Newtonian NCDM component, according to Eq. (2.51) and Eq. (2.52). The equation of motion can be derived from the action

$$\begin{aligned} S_{\rm New} &= \int dt L_{\rm New} \\ &= m\int dt\left[\frac{1}{2}a^2\left(\frac{d\vec{x}}{dt}\right)^2 - \psi\right] \end{aligned} \tag{2.53}$$

The canonical momentum associated to this is

$$\vec{\Pi}_{\rm New} = \frac{\partial L_{\rm New}}{\partial\dot{\vec{x}}} = ma^2\frac{d\vec{x}}{dt} = ma\vec{v} = \vec{q} \tag{2.54}$$

where $\vec{q} = a\vec{p}$, $\vec{p}$ is the proper momentum, and spatial indices are raised and lowered with the kronecker delta. We may now find the Hamiltonian associated to Eq. (2.53), and from it Hamilton's equation of motion. It yields

$$\frac{d\vec{q}}{d\tau} = -ma\vec{\nabla}\psi\,. \tag{2.55}$$

The dynamics of the distribution function follows from the collisionless Boltzmann equation

$$\frac{df}{d\tau} = \frac{\partial f}{\partial \tau} + \frac{dx^i}{d\tau}\frac{\partial f}{\partial x^i} + \frac{dq^i}{d\tau}\frac{\partial f}{\partial q^i} = 0 \tag{2.56}$$

Splitting the distribution function as $f = f_0(q)(1 + \Psi(\tau, \vec{x}, \vec{q}))$, working to leading order on perturbations (e.g. Ψ, ϕ, ψ), and moving to momentum space ($\partial_i \to ik_i$)

$$\Psi' + i\frac{qk}{ma}(\hat{n}\cdot\hat{k})\Psi = i\frac{mak}{q}(\hat{n}\cdot\hat{k})\psi\frac{d\ln f_0}{d\ln q}\,, \tag{2.57}$$

where $\hat{n} = \vec{q}/q$ and $'$ denotes partial derivative with respect to conformal time. Next we make the usual assumption of axial symmetry of Ψ around $\hat{k}$, at the initial redshift. This allows one to write the usual multipole expansion

$$\Psi(\tau, \vec{k}, q, \hat{n}) = \sum_{l=0}^{\infty}(-i)^l(2l+1)\Psi_l(\tau, k, q)P_l(\hat{n}\cdot\hat{k}) \tag{2.58}$$

where P_ℓ are the Legendre polynomials. Substitution of Eq. (2.58) into Eq. (2.57) yields a hierarchy of evolution equations for the multipoles

$$\Psi_0' + \frac{qk}{ma}\Psi_1 = 0 \tag{2.59a}$$

$$\Psi_1' - \frac{qk}{3ma}(\Psi_0 - 2\Psi_2) = -\frac{mak}{3q}\psi\frac{d\ln f_0}{d\ln q} \tag{2.59b}$$

$$\Psi_l' - \frac{qk}{(2l+1)ma}[l\Psi_{l-1} - (l+1)\Psi_{l+1}] = 0 \tag{2.59c}$$

In principle, this is all one needs to study the evolution of linear perturbations of a Newtonian NCDM component. It is, however, convenient to rephrase this as a set of fluid equations for some suitably defined fluid properties. The starting point for this is the expression for the mass density

$$\begin{aligned}\rho(1+\delta) &= a^{-3}\int d\Omega\int_0^\infty d\Pi_{\rm New}\Pi_{\rm New}^2 f_0(1+\Psi)m \\ &= ma^{-3}\int d\Omega\int_0^\infty dq q^2 f_0(q)(1+\Psi)\end{aligned} \tag{2.60}$$

This definition of Newtonian energy density is motivated by how it is computed in N-body simulations, i.e. where only the rest mass of particles, as opposed to the total energy $E = \gamma m$,

contributes as a source to the gravitational potential. The factor of $(1+3\phi)$, to account for inhomogeneities in the local spatial volume, is also neglected. One can think of choosing a differente gauge, where there are no inhomogeneities in the local spatial volume, e.g. the N-body gauge [35]. This is a GR correction that will be included in the non-relativistic case.

Substitution of Eq. (2.58) into Eq. (2.60) gives, also defining other relevant fluid properties

$$\delta\rho = \rho\delta = 4\pi m a^{-3}\int_0^\infty dq\, q^2 f_0(q)\ \Psi_0 \tag{2.61a}$$

$$\delta P = \frac{4\pi}{3} m a^{-3}\int_0^\infty dq\, q^2 f_0(q)\Big(\frac{q}{ma}\Big)^2\Psi_0 \tag{2.61b}$$

$$\rho\theta = 4\pi k a^{-4}\int_0^\infty dq\, q^2 f_0(q)\ q\ \Psi_1 \tag{2.61c}$$

$$\rho\sigma = \frac{8\pi}{3} m a^{-3}\int_0^\infty dq\, q^2 f_0(q)\Big(\frac{q}{ma}\Big)^2\Psi_2 \tag{2.61d}$$

To derive fluid equations, take derivatives of Eq. (2.61) with respect to conformal time, and use Eq. (2.59) to arrive at

$$\delta' = -\theta \tag{2.62a}$$

$$\theta' = -\mathcal{H}\theta + \frac{\delta P}{\delta\rho}k^2\delta - k^2\sigma + k^2\psi \tag{2.62b}$$

$$\sigma' = -2\mathcal{H}\sigma + \frac{4}{15}\Theta - kF_3\,, \tag{2.62c}$$

where $\mathcal{H} = a'/a$ and the additional variables Θ, F_3 are given by,

$$\rho F_3 = \frac{8\pi}{5} m a^{-3}\int_0^\infty q^2 f_0(q) dq\Big(\frac{q}{ma}\Big)^3\Psi_3 \tag{2.63a}$$

$$\rho\Theta = 4\pi k a^{-4}\int_0^\infty q^2 f_0(q) dq q\Big(\frac{q}{ma}\Big)^2\Psi_1 \tag{2.63b}$$

In order to close the system of Eqs. (2.62), we first write the approximation

$$\Psi_3 \approx \frac{5ma}{qk\tau}\Psi_2 - \Psi_1 \tag{2.64}$$

This is a straightforward generalization, suitable to the Newtonian evolution (i.e. let $\epsilon = \sqrt{q^2 + a^2m^2} \to am$), of the truncation scheme found in [19]. This implies

$$kF_3 \approx \frac{3}{\tau}\sigma - \frac{2}{5}\Theta \tag{2.65}$$

Finally, the approximation for the higher velocity weight fluid properties follows [127]. This amounts to

$$c_{\rm eff}^2 = \frac{\delta P}{\delta \rho} \approx c_g^2 = \frac{P'}{\rho'} = \frac{5}{3} w = \frac{25}{3}\frac{\zeta(5)}{\zeta(3)}\left(\frac{T_0}{m}\right)^2 (1+z)^2 \tag{2.66a}$$

$$\Theta \approx 12 \frac{w}{1+w} c_g^2 \theta \tag{2.66b}$$

For the sound speed and equation of state, we have used the Newtonian expressions for the background pressure and energy density in Eqs. (2.10) and (2.11), along with Eq. (2.5). Substitution of Eq. (2.66) into Eq. (2.62) yields

$$\delta' = -\theta \tag{2.67a}$$

$$\theta' = -\mathcal{H}\theta + c_g^2 k^2 \delta - k^2 \sigma + k^2 \psi \tag{2.67b}$$

$$\sigma' = -\left(2\mathcal{H} + \frac{3}{\tau}\right)\sigma + 8\frac{w}{1+w} c_g^2 \theta \tag{2.67c}$$

As one can see from Eq. (2.67), this is missing GR corrections, as expected of a Newtonian limit. In this work we concentrate on an SR effect, the shift in the neutrino free-streaming scale, and its impact on the matter power spectrum. Hence, we would like to include GR corrections in our fluid equations. We do so in a self-consistent way by considering the dynamics of a non-relativistic NCDM component, i.e. as follows from Eq. (2.50) and Eq. (2.49).

The equation of motion is obtained from the action

$$\begin{aligned} S_{\rm NR} &= \int dt L_{\rm NR} \\ &= m \int dt \left[\frac{1}{2} a^2 \left(1 - 2\phi - \psi\right) \left(\frac{d\vec{x}}{dt}\right)^2 - \psi\right] \end{aligned} \tag{2.68}$$

in the NR limit. This action follows from $-m\int\sqrt{-ds^2}$, after expanding to leading order in metric perturbations and taking the NR limit. The canonical momentum associated to this is

$$\vec{\Pi}_{\rm NR} = \frac{\partial L_{\rm NR}}{\partial \dot{\vec{x}}} = m a^2 \left(1 - 2\phi - \psi\right) \frac{d\vec{x}}{dt} = (1-\phi)\vec{q} \tag{2.69}$$

The analog of Eq. (2.55) is now, again in the NR limit

$$\frac{d\vec{q}}{d\tau} = -ma\vec{\nabla}\psi + \phi'\vec{q} \tag{2.70}$$

The rest of the procedure to derive fluid equations follows closely what is done in the Newtonian case. The only difference is that we now include the factor of $(1+3\phi)$ to account for inhomogeneities in the local spatial volume, in the definition of the mass density. That is,

$$\begin{aligned}\rho(1+\delta) &= (1+3\phi)a^{-3}\int d\Omega\int_0^\infty d\Pi_{\rm NR}\Pi_{\rm NR}^2 f_0(1+\Psi)m \\ &= ma^{-3}\int d\Omega\int_0^\infty dq q^2 f_0(q)(1+\Psi)\end{aligned} \tag{2.71}$$

Note that this is formally identical to Eq. (2.60). The only difference in the fluid equations will then come from the additional GR term in the right hand side of Eq. (2.70), contributing to the evolution equation of the zeroth multipole of the distribution function, and thus we will arrive at

$$\delta' = -\theta + 3\phi' \tag{2.72a}$$

$$\theta' = -\mathcal{H}\theta + c_g^2 k^2\delta - k^2\sigma + k^2\psi \tag{2.72b}$$

$$\sigma' = -\Big(2\mathcal{H} + \frac{3}{\tau}\Big)\sigma + 8\frac{w}{1+w}c_g^2\theta\,. \tag{2.72c}$$

Chapter 3

GENERALIZED BOLTZMANN HIERARCHY FOR MASSIVE NEUTRINOS IN COSMOLOGY

3.1 Abstract

Boltzmann solvers are an important tool for the computation of cosmological observables in the linear regime. In the presence of massive neutrinos, they involve solving the Boltzmann equation followed by an integration in momentum space to arrive at the desired fluid properties, a procedure which is known to be computationally slow. In this work we introduce the so-called generalized Boltzmann hierarchy (GBH) for massive neutrinos in cosmology, an alternative to the usual Boltzmann hierarchy, where the momentum dependence is integrated out leaving us with a two-parameter infinite set of ordinary differential equations. Along with the usual expansion in multipoles, there is now also an expansion in higher velocity weight integrals of the distribution function. Using a toy code, we show that the GBH produces the density contrast neutrino transfer function to a $\lesssim 0.5\%$ accuracy at both large and intermediate scales compared to the neutrino free-streaming scale, thus providing a proof-of-principle for the GBH. We comment on the implementation of the GBH in a state of the art Boltzmann solver.

3.2 Introduction

Neutrino oscillation experiments have established that neutrinos are massive particles (at least two eigenstates), with a lower bound, in the sum of all neutrino masses, of $\sum m_\nu \geq 0.06\text{eV}$, 0.1eV for normal and inverted hierarchies, respectively [84, 85]. The large-scale structure of our Universe gives a sensitive probe of neutrino masses [64, 130]. This allows us to use cosmological data to constrain the sum of neutrino mass eigenstates: $\sum m_\nu \lesssim$

$(0.1-0.3)$eV, e.g. [11, 88], depending on the choice of used datasets. Current and future large-scale structure surveys [90–96] will be used to determine the mass scale of neutrinos [97, 98], but also to constraint beyond-ΛCDM scenarios [100, 101]. It is then of paramount importance that cosmological observables, such as the matter power spectrum, can be computed to a subpercent level accuracy, in both linear and nonlinear scales.

The study of structure formation in the nonlinear regime relies on N-body simulations [118, 141]. On the other hand, the linear theory is much simpler, and there are publicly available codes, such as the code for anisotropies in the microwave background (CAMB) [142] and the cosmic linear anisotropy solving system (CLASS) [127], that can be used to compute the observables. The implementation of neutrinos in the linear theory is somewhat cumbersome, since it involves solving a Boltzmann hierarchy of equations in momentum space. The reason for this can be traced back to the usual statement that the momentum dependence in the distribution function cannot be integrated out [19]. For this reason, fluid approximations have been developed in the past, and incorporated as an optional tool in the Boltzmann solvers [127, 143, 144].

In this work we show that the momentum dependence in the distribution function can, in fact, be exactly integrated out, at the expense of introducing a new countable parameter n, along with the parameter l associated with the multipole expansion, to the infinite system of ordinary differential equations that need to be solved to determine the dynamics, in Fourier space ([144, 145] being examples of this in the literature). This leads us to a novel two-parameter infinite set of equations that determine the evolution of noncold dark matter (or ncdm, borrowing notation from CLASS [127]) perturbations in a flat universe: The generalized Boltzmann hierarchy (GBH). Along with the usual multipole expansion, there is now also an expansion in higher velocity weight integrals of the distribution function [1]. The GBH is simpler than the usual approach, as implemented in Boltzmann solvers, in the sense

[1] An expansion in higher velocity weights is used in CAMB to approximate the evolution of the perturbations in massive neutrinos once already in the non-relativistic regime [144], while in our approach we obtain the exact evolution, starting from initial conditions while still in the relativistic regime.

that it does not require the numerical computation of momentum integrals, after solving the dynamical equations.

The paper is organized as follows: In Sec. 3.3, we introduce the generalized Boltzmann hierarchy for ncdm perturbations, and compare it to the usual approach of evolving the distribution function in phase space. In Sec. 3.4, we implement the equations numerically for a single massive neutrino component, with varying mass m and scale k, obtaining $\lesssim 0.5\%$ agreement with the Boltzmann solver CLASS in high precision settings, for all redshift. We also discuss the dependence of our framework on the neutrino mass m and scale k, and show that when switching to a fluid approximation on the small scales, i.e. once a given mode becomes smaller than the free-streaming scale (hereafter named GBH+FA), we can produce the neutrino transfer function, at $z = 0$, with the same accuracy as CLASS in its default precision settings (or CLASS-DPS), over all scales. We also found that CAMB, when also in its default precision settings (CAMB-DPS), yields more accurate results than both the GBH+FA and CLASS-DPS on the small scales. In Sec. 3.5, we conclude and comment on the implementation of the GBH in Boltzmann solvers, along with its current limitations. In Appendix 3.6, we give a detailed account of the truncation scheme for the GBH. Finally, in Appendix 3.7 we investigate the GBH in the simple case of $l_{\rm max} = 2$ and $n_{\rm max} = 0$, i.e. a viscous fluid approximation (FA). We compare the truncation scheme developed for the GBH with the one employed in CLASS.

3.3 Generalized Boltzmann Hierarchy

We start by introducing the additional expansion in higher velocity weight integrals of the distribution function, followed by a derivation of their associated dynamical equations. We will be considering scalar (linear) perturbations to a flat Friedmann-Robertson-Walker (FRW) universe, in the Newtonian gauge

$$ds^2 = a(\tau)^2[-(1+2\psi)d\tau^2 + (1-2\phi)d\vec{x}^2] \tag{3.1}$$

Conventions and notation follow [19]. Let us first define suitable generalized fluid properties, at the level of background

$$P_n = \rho\omega_n \overset{\text{def}}{=} \frac{4\pi}{3}a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n} \ \forall n \geq 0 \tag{3.2}$$

where $f_0(q) = (2/(2\pi)^3)(1+e^{\frac{q}{T_0}})^{-1}$ is the Fermi-Dirac distribution written in terms of co-moving momentum $\vec{q} = a\vec{p}$, with $\vec{p}$ the proper momentum and $T_0 \approx 1.95$K the temperature of relic neutrinos today. Also $\epsilon = \sqrt{q^2 + a^2m^2} = aE$, with m the neutrino mass and E the proper energy. Then $P_0 = (1/3)\rho$ is a third of the background energy density, $P_1 = P$ is the pressure, $P_2 \equiv \mathcal{P}$, and in general $P_{n+2} \equiv \mathcal{P}^{(n)}, n \geq 0$ are higher velocity weight pressures. Similarly, $\omega_0 = 1/3$, $\omega_1 = \omega$ is the equation of state parameter, and $\omega_{n+2},\ n \geq 0$ are higher velocity weight equation of state parameters.

Taking the derivative of Eq. (3.2) with respect to conformal time gives the following hierarchy of equations

$$\omega_n' = -(2n+3)\mathcal{H}\omega_n + (2n-1)\mathcal{H}\omega_{n+1} - \frac{\rho'}{\rho}\omega_n \ \forall n \geq 0 \tag{3.3}$$

where $\mathcal{H} = a'/a$, and $'$ denotes derivative with respect to conformal time. Notice that $q/\epsilon = v \sim T_0/ma$ is the physical velocity of an individual neutrino particle, such that the additional factors of $(q/\epsilon)^2$ in the integrals in Eq. (3.2) effectively shift the peak of the distribution function to higher particle velocities. In the relativistic regime, all particles travel at the speed of light, and one only needs to consider the $n = 0$ equation. The same holds in the nonrelativistic regime, where $\mathcal{O}(v^2)$ corrections become negligible. During the transition, however, the higher velocity weight fluid properties need to be taken into account, in order to probe the whole spectrum of neutrino particle velocities.

Setting $n = 0$ in Eq. (3.3) yields the familiar equation $\rho' + 3\mathcal{H}(\rho + P) = 0$. At this level, it is easier to simply determine the evolution of the distribution function, and then integrate Eq. (3.2) directly, than to approach the infinite set of Eqs. (3.3). This is because the background distribution function admits a simple, analytic solution (e.g. the Fermi-Dirac distribution). However, this is no longer true when inhomogeneities are introduced.

In this case we have, along with the expansion in higher velocity weight integrals of the distribution function, parametrized by n, the usual expansion in multipoles, parametrized by l. To obtain it, split the distribution function as $f = f_0(q)(1+\Psi)$, and expand Ψ in a Legendre series

$$\Psi(\vec{k},\hat{n},q,\tau) = \sum_{l=0}^{\infty}(-i)^l(2l+1)\Psi_l(k,q,\tau)P_l(\hat{k}\cdot\hat{n}) \tag{3.4}$$

where we are working in Fourier space (set $\vec{\nabla}\to i\vec{k}$), and define $\hat{n}=\vec{q}/q$.

In terms of the multipole expansion in Eq. (3.4), the fluid properties directly sourcing the gravitational field, i.e. in the energy momentum tensor, are [19]

$$\delta\rho = \rho\delta = 4\pi a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Psi_0 \tag{3.5a}$$

$$\delta P = \frac{4\pi}{3}a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^2\Psi_0 \tag{3.5b}$$

$$(\rho+P)\theta = 4\pi k a^{-4}\int_0^\infty dqq^2 f_0(q)q\Psi_1 \tag{3.5c}$$

$$(\rho+P)\sigma = \frac{8\pi}{3}a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^2\Psi_2 \tag{3.5d}$$

We now wish to generalize this to higher multipoles, and also include higher velocity weight integrals of the distribution function, in analogy to Eq. (3.2), to make sure that the whole spectrum of particle velocities is being probed during the transition from the relativistic to nonrelativistic regimes. The following is then a natural choice of dynamical variables:

$$\delta P_n = \rho\delta_n \stackrel{\text{def}}{=} \frac{4\pi}{3}a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n}\Psi_0 \tag{3.6a}$$

$$(\rho+P)\theta_n \stackrel{\text{def}}{=} 4\pi k a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n+1}\Psi_1 \tag{3.6b}$$

$$(\rho+P)f_{l,n} \stackrel{\text{def}}{=} 4\pi\frac{l!}{(2l-1)!!}a^{-4}\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n+l}\Psi_l \quad \forall\ l\geq 1 \tag{3.6c}$$

and $n\geq 0$ everywhere. Notice that $\delta_0 = \frac{1}{3}\delta$ is a third of the density contrast, $\rho\delta_1 = \delta P$ is the perturbation to the pressure, $\rho\delta_2 \equiv \delta\mathcal{P}$, and $\rho\delta_{n+2}\equiv\delta\mathcal{P}^{(n)}, n\geq 0$ are the perturbations to the higher velocity weight pressures. Also $\theta_0=\theta$ is the divergence of the velocity, $\theta_1\equiv\Theta$

and $\theta_{n+1} \equiv \Theta^{(n)}, n \geq 0$ are its higher velocity weight counterparts. We also define $f_{2,n} \equiv \sigma_n$, the anisotropic shear stress, with a similar notation for its higher velocity weight integrals (i.e. $\Sigma^{(n)}$), and set $\theta_n \equiv k f_{1,n}$, when it is convenient to do so. In order to derive a set of equations for the variables in Eq. (3.6), we need the time evolution of the multipoles. It follows from the substitution of Eq. (3.4) into the Boltzmann equation, and reads [19]

$$\Psi_0' = -\frac{qk}{\epsilon}\Psi_1 - \phi'\frac{d\ln f_0}{d\ln q} \tag{3.7a}$$

$$\Psi_1' = \frac{qk}{3\epsilon}(\Psi_0 - 2\Psi_2) - \frac{\epsilon k}{3q}\psi\frac{d\ln f_0}{d\ln q} \tag{3.7b}$$

$$\Psi_l' = \frac{qk}{(2l+1)\epsilon}[l\Psi_{l-1} - (l+1)\Psi_{l+1}] \;\forall\; l \geq 2 \tag{3.7c}$$

Now take the derivative of each expression in Eq. (3.6) with respect to conformal time, and use Eqs. (3.2), (3.3), (3.6) and (3.7) to arrive at

$$\delta_n' = -(2n-3w)\mathcal{H}\delta_n + (2n-1)\mathcal{H}\delta_{n+1} - \frac{1}{3}(1+w)\theta_n + [(2n+3)\omega_n - (2n-1)\omega_{n+1}]\phi' \tag{3.8a}$$

$$\theta_n' = -\Big[(2n+1-3w)\mathcal{H} + \frac{w'}{1+w}\Big]\theta_n + 2n\mathcal{H}\theta_{n+1} + \frac{1}{1+w}k^2\delta_{n+1} - k^2\sigma_n \tag{3.8b}$$
$$+ \frac{1}{1+w}[(2n+3)\omega_n - (2n-1)\omega_{n+1}]k^2\psi$$

$$f_{l,n}' = -\Big[(2n+l-3w)\mathcal{H} + \frac{w'}{1+w}\Big]f_{l,n} + (2n+l-1)\mathcal{H}f_{l,n+1} + \frac{l^2}{4l^2-1}kf_{l-1,n+1} \tag{3.8c}$$
$$- kf_{l+1,n} \;\forall l \geq 2$$

with $n \geq 0$. This is the generalized Boltzmann hierarchy (GBH) for a ncdm component. Setting $n = 0$ in Eq. (3.8), one recovers the usual ncdm fluid equations (and up to $l = 2$, including only dynamical equations for the fluid properties that directly source the gravitational field)

$$\delta' = -(1+w)(\theta - 3\phi') - 3\mathcal{H}\Big(\frac{\delta P}{\delta\rho} - w\Big)\delta \tag{3.9a}$$

$$\theta' = -\Big[(1-3w)\mathcal{H} + \frac{w'}{1+w}\Big]\theta + \frac{\delta P/\delta\rho}{1+w}k^2\delta - k^2\sigma + k^2\psi \tag{3.9b}$$

$$\sigma' = -\Big[(2-3w)\mathcal{H} + \frac{w'}{1+w}\Big]\sigma + \mathcal{H}\Sigma + \frac{4}{15}\Theta - kf_3 \tag{3.9c}$$

As is well known, Eq. (3.9) involves variables, i.e. $\delta P/\delta\rho$, Σ, Θ and kf_3, that need to be somehow approximated, in terms of the dynamical variables in the system, in order to close the equations. On the other hand, the two-parameter hierarchy of Eqs. (3.8) is closed as it is, and we have achieved our goal: To get rid of the momentum integrals altogether. Of course, any practical implementation of the GBH requires a good truncation scheme: A choice of a given number of multipoles, $l_{\rm max}+1$, and higher velocity weight variables, $n_{\rm max}+1$, to dynamically evolve, together with a recipe for approximating higher order quantities. The resulting system of equations is of dimension $(l_{\rm max}+1)\times(n_{\rm max}+1)$. We discuss this at length in Appendix 3.6. Here we will just spell out the recipe. The truncation in multipoles is done with the approximation

$$f_{l_{\rm max}+1,n}\approx(l_{\rm max}+1)\Big(\frac{1}{k\tau}f_{l_{\rm max},n}-\frac{l_{\rm max}}{4l_{\rm max}^2-1}f_{l_{\rm max}-1,n+1}\Big) \tag{3.10}$$

while the truncation in higher velocity weight integrals is handled with

$$\frac{\delta_{n_{\rm max}+1}}{\delta_{n_{\rm max}}}\approx\frac{2n_{\rm max}+5}{2n_{\rm max}+3}\frac{\omega_{n_{\rm max}+1}}{\omega_{n_{\rm max}}}\frac{1-\frac{2n_{\rm max}+1}{2n_{\rm max}+5}\frac{\omega_{n_{\rm max}+2}}{\omega_{n_{\rm max}+1}}}{1-\frac{2n_{\rm max}-1}{2n_{\rm max}+3}\frac{\omega_{n_{\rm max}+1}}{\omega_{n_{\rm max}}}} \tag{3.11a}$$

$$\frac{f_{l,n_{\rm max}+1}}{f_{l,n_{\rm max}}}\approx\frac{2(n_{\rm max}+l)+3}{2(n_{\rm max}+l)+1}\frac{\omega_{n_{\rm max}+l}}{\omega_{n_{\rm max}+l-1}}\frac{1-\frac{2(n_{\rm max}+l)-1}{2(n_{\rm max}+l)+3}\frac{\omega_{n_{\rm max}+l+1}}{\omega_{n_{\rm max}+l}}}{1-\frac{2(n_{\rm max}+l)-3}{2(n_{\rm max}+l)+1}\frac{\omega_{n_{\rm max}+l}}{\omega_{n_{\rm max}+l-1}}}\quad\forall\ l\geq 1 \tag{3.11b}$$

Furthermore, as explained in Appendix 3.6, both the l and n expansions are controlled by the parameter $x=k\mathrm{T}$, with T the neutrino horizon (average comoving distance traveled by neutrino particles through cosmic history, see Eq. (3.19) and comments below Eq. (3.25), along with the plot in Fig. 3.1). Specifically, if one wishes to follow the neutrino transfer function up to a time x, we found that

$$l_{\rm max}\approx\frac{x}{2} \tag{3.12a}$$

$$n_{\rm max}\approx\frac{x^{1.6}}{5} \tag{3.12b}$$

are approximately sufficient for convergence, up to $x=30$. These are plotted in Fig. 3.2. This concludes our discussion on the truncation scheme.

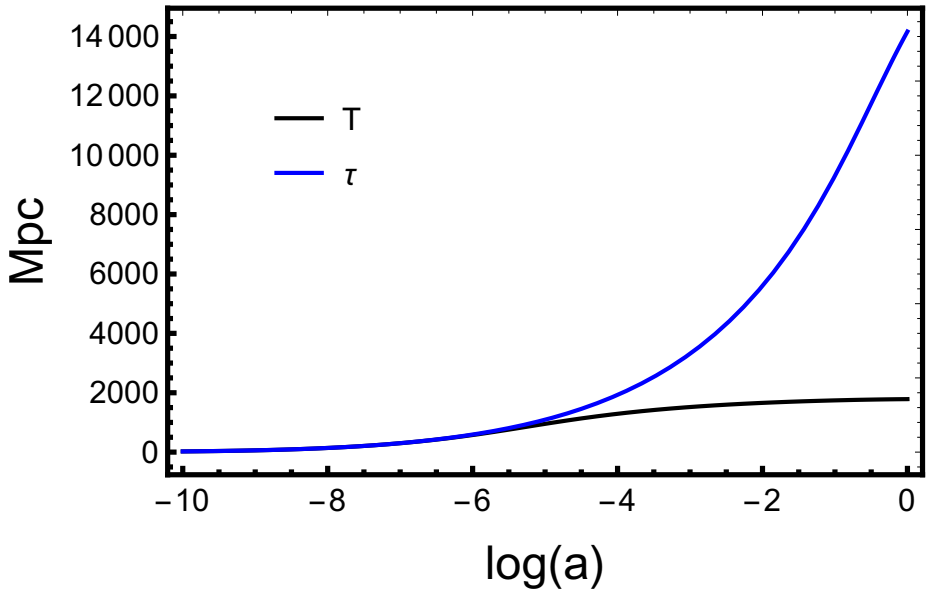

Figure 3.1: Evolution of the neutrino horizon, i.e. T for $q/T_0 = 3$, and $m = 0.1$eV. It grows like the conformal time τ up to the time of transition, when it effectively freezes as it approaches the nonrelativistic regime. The significant difference between T and τ in the nonrelativistic regime explains why one needs a much higher $l_{\max}$ for radiation than for massive neutrinos [19].

Now we move on to setting the initial conditions for the GBH, starting at early times when all the individual neutrinos still move at the speed of light. Notice that in the relativistic regime $\frac{q}{\epsilon} \to 1$ and all higher velocity weight integrals approach one another: We recover the usual hierarchy of equations for radiation. One can then set the (say adiabatic) initial conditions for the $n = 0$ fluid properties as usual [19]

$$\delta = -2\psi \tag{3.13a}$$

$$\theta = \frac{1}{2}(k^2\tau)\psi \tag{3.13b}$$

$$\sigma = \frac{1}{15}(k\tau)^2\psi \tag{3.13c}$$

$$f_{l,0} = 0 \ \forall l > 2. \tag{3.13d}$$

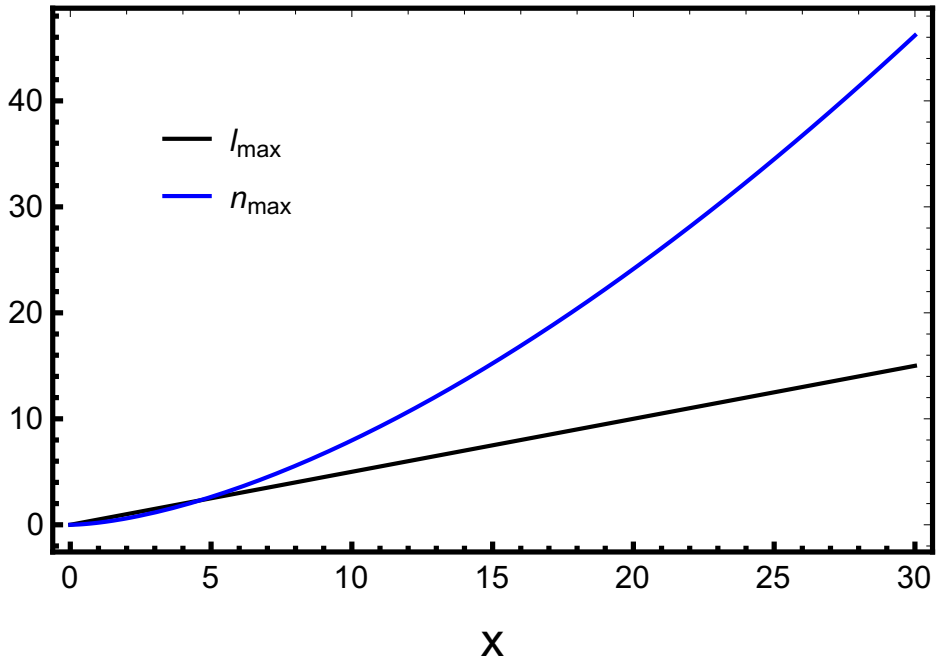

Figure 3.2: Choices of $l_{\max}$ and $n_{\max}$ as a function of x, as given by Eq. (3.12), to approximately ensure convergence, up to $x = 30$. Note that $n_{\max}$ grows faster than linearly with x.

Along with $f_{l,n} = f_{l,0}\ \forall n > 0$ to set the initial condition for the remaining higher velocity weight fluid properties.

In the nonrelativistic regime, these variables get suppressed by powers of $(q/\epsilon)^2 = v^2 \sim (T_0/ma)^2$, and similarly $n = 0$ should suffice for most applications. During the transition, however, higher order contributions become important, and must be included for an accurate computation of the ncdm transfer functions.

In the standard approach for including massive neutrinos in the computation of cosmological observables in linear theory, Eq. (3.7) are solved for a given number N_q of momentum bins, and up to some $l_{\max}$, i.e. a system of dimensionality $N_q \times (l_{\max} + 1)$. The solution is then used to perform the q integrals in Eq. (3.5) for the neutrino fluid properties, which in turn are coupled to the perturbations to the other components in the universe via the Einstein equations. This is known to be a computationally slow procedure [19, 144]. Orig-

inally, these q integrals were evaluated with a fixed grid of many equally spaced samples. A significant acceleration in the integration procedure was achieved once a kernel-weighted sampling scheme was introduced, allowing for a much smaller number of momentum bins N_q (now employed in both CLASS and CAMB [127, 146]).

The GBH is simpler than the standard approach in that it removes the intermediate step of performing the q integrals, i.e. the neutrino fluid properties are directly coupled to the perturbations to the other species in the universe. It is then plausible that the GBH may be faster than the standard approach (a performance comparison between the GBH and the standard approach is left to future work, see Sec. 3.5). However, it could be the case that the dimensionality of the GBH, i.e. $(l_{\rm max}+1)\times(n_{\rm max}+1)$ is significantly bigger than $N_q\times(l_{\rm max}+1)$ (and it is actually what happens on the small scales, given the rapid increase of $n_{\rm max}$ with x, as seen in Fig. 3.2), for the same achieved accuracy. Notwithstanding, in some cases a very large number of momentum bins is actually necessary, e.g. to accurately obtain the effective sound speed [147]. The GBH is not plagued with the same issue since all momentum dependence is integrated out of the dynamical equations.

We now have everything we need, i.e. a closed system of dynamical equations plus suitable initial conditions, to consider the numerical implementation of the GBH. This will allow us to compare it with the Boltzmann solvers CLASS and CAMB.

3.4 Numerical implementation

As an example of the numerical implementation of the GBH, we will first consider an individual neutrino component, with varying mass: $m=0.02$eV, $m=0.1$eV and $m=0.5$eV. Also, we implement the GBH at intermediate scales compared to the neutrino horizon today, i.e. $k=x/\mathrm{T}$ with $x=15$ and $x=30$. On larger scales ($x\lesssim 1$), neutrino velocities are unimportant and a simple viscous fluid approximation suffices i.e. $l_{\rm max}=2, n_{\rm max}=0$ should be enough, while on smaller scales (say $x>30$) accurately obtaining the neutrino transfer functions is not so important because of free-streaming: Neutrino perturbations get washed out and have a negligible impact on matter perturbations. Nonlinear effects also start to kick

in. Furthermore, because of the rapid increase of $n_{\max}$ with x found in Fig. 3.2, integration time also rapidly increases with x.

The neutrino density contrast transfer function, as a function of the scale factor, obtained from the GBH is compared to the output from CLASS. Boltzmann solvers do not produce accurate neutrino transfer functions at their default precision settings, as the codes are tailored to accurately produce the matter power spectrum, and relic neutrinos only have a subleading impact on this observable. To obtain accurate results, we follow the improved settings found in Appendix B of [147]: Turn off the CLASS ncdm fluid approximation, use a quadrature strategy to perform the q-integrals, with $N_q = 30$ momentum bins, and set $l_{\max} = 30$, i.e. a system of 930 equations. We then expect to get subpercent level accuracy for all redshifts and scales of interest. Note that CAMB does not provide the option of outputting neutrino transfer functions as a function of redshift, for a given fixed scale. This is why a direct comparison of the GBH with CAMB is not included at this stage. However, in high precision settings the neutrino transfer functions from CLASS and CAMB are known to agree to a percent level [147].

We develop a toy code, where the sources $\phi(a)$ and $\psi(a)$, as given by CLASS, are used to evolve the GBH. In that way, we do not need to solve the Einstein equations, i.e. we do not need to consider the dynamics of perturbations to the other components in the universe, and can evolve the neutrino perturbations alone. The full problem of implementing the GBH in a Boltzmann code, followed by a comparison of performance with the standard method, is beyond the scope of this paper, and is left to future work. This is discussed in Sec. 3.5. Our goal here is to demonstrate that the GBH indeed can be used to produce accurate neutrino transfer functions, i.e. to provide a proof-of-principle for the GBH.

The initial conditions are set, according to Eq. (3.13), at some arbitrary early time, when all modes of interest are in superhorizon scales, and neutrinos are relativistic. Our choices of $l_{\max}$ and $n_{\max}$ are guided by Eq. (3.12). For $x = 15$, we set $l_{\max} = 8$ and $n_{\max} = 16$, while for $x = 30$, we choose $l_{\max} = 15$ and $n_{\max} = 49$.

The relative difference in the transfer functions generated from CLASS and the GBH are

shown in Fig. 3.3: There is $\lesssim 0.5\%$ agreement for all redshift, so the GBH is accurately producing neutrino transfer functions.

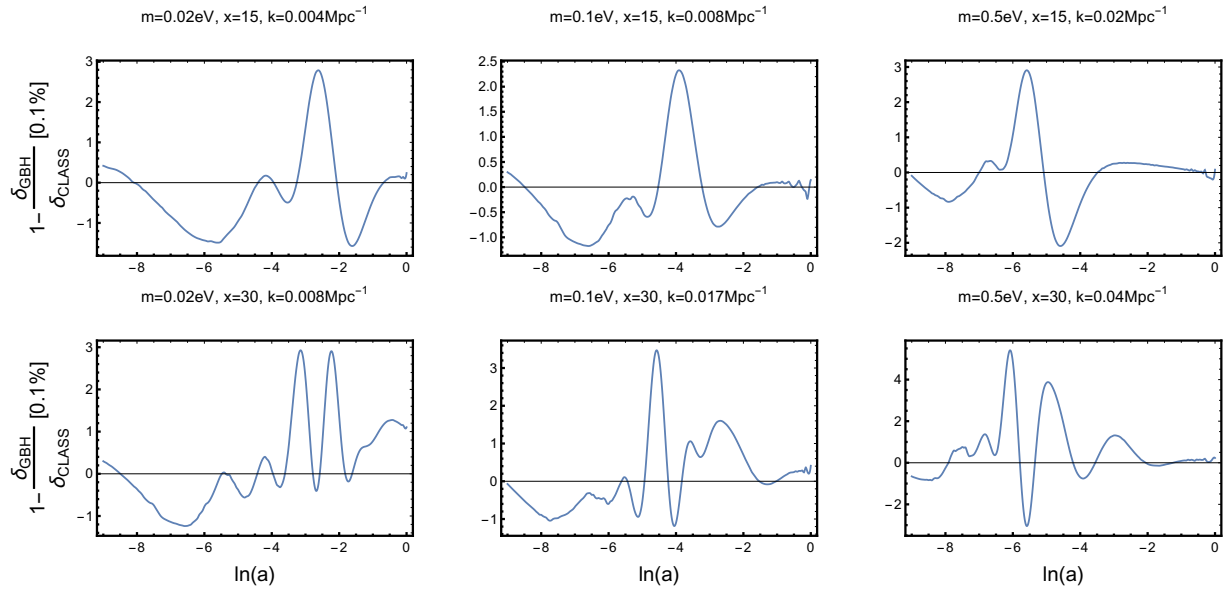

Figure 3.3: Relative difference in the density contrast neutrino transfer function from the GBH and CLASS (in high precision settings) for neutrino masses of $m = 0.02$eV, $m = 0.1$eV, and $m = 0.5$eV, and at intermediate scales defined by $k = x/\mathrm{T}$, with $x = 15$ (top) and $x = 30$ (bottom). The agreement is in the $\lesssim 0.5\%$ level for all redshift. We conclude that the GBH is accurately producing neutrino transfer functions.

The case $x > 30$ requires a much larger system of equations to obtain subpercent level agreement with CLASS, with the n expansion converging rather slowly on the small scales. There are also some technical issues with the truncation scheme for the GBH in this regime (see the discussion at the end of Appendix 3.6), so going beyond $x = 30$ requires a more efficient numerical implementation.

Notwithstanding, and as pointed out before, on the small scales neutrinos free-stream, so it no longer becomes important that neutrino transfer functions are obtained very accurately (nonlinearities also start to kick in). In fact, we find that when switching to a viscous

fluid approximation once a given mode is sufficiently inside the horizon (GBH+FA), we produce the density contrast neutrino transfer function as accurately as CLASS in its default precision settings (CLASS-DPS, where a similar switch to a viscous fluid approximation is also employed), over all scales of interest, and at redshift $z = 0$. We also find that CAMB in its default precision settings (CAMB-DPS) is as accurate as CLASS-DPS and the GBH+FA at intermediate and large scales, but more accurate on the small scales, where both the GBH+FA and CLASS-DPS are in the FA regime, and hence not producing neutrino transfer functions very accurately. Finally, CAMB in its high-precision settings (CAMB-HPS) [2] is found to closely agree with CLASS in its high-precision settings (CLASS-HPS), according to expectation. These results are illustrated in Fig. 3.4.

Finally, as derived in Appendix 3.6, T $\sim \sqrt{T_0/m}$ in the nonrelativistic regime. It implies that, and for a given fixed scale k, one needs higher $l_{\max}$ and $n_{\max}$ for smaller neutrino masses (assuming that m is big enough for the transition to the nonrelativistic regime to happen before today). In Fig. 3.5, we plot both $l_{\max}$ and $n_{\max}$, as given by Eq. (3.12), as a function of scale for varying mass.

3.5 Conclusion

We introduced the so-called generalized Boltzmann hierarchy (GBH) for noncold dark matter cosmological perturbations in a flat universe, an alternative to the usual Boltzmann hierarchy for accurately producing neutrino transfer functions in the linear regime. It was determined that the GBH agrees with Boltzmann solvers in high precision settings, to a $\lesssim 0.5\%$ level accuracy, in both large and intermediate scales compared to the neutrino free-streaming scale.

On the small scales one needs to choose a very high $n_{\max}$ in order to produce accurate neutrino transfer functions, and the numerical integration of the GBH becomes computationally expensive. However, one should keep in mind that on small scales free-streaming

[2] CAMB-HPS is defined by the following choice of precision settings: massive_nu_approx=0, accurate_massive_neutrino_transfers = T, accuracy_boost=3 and l_accuracy_boost=3.

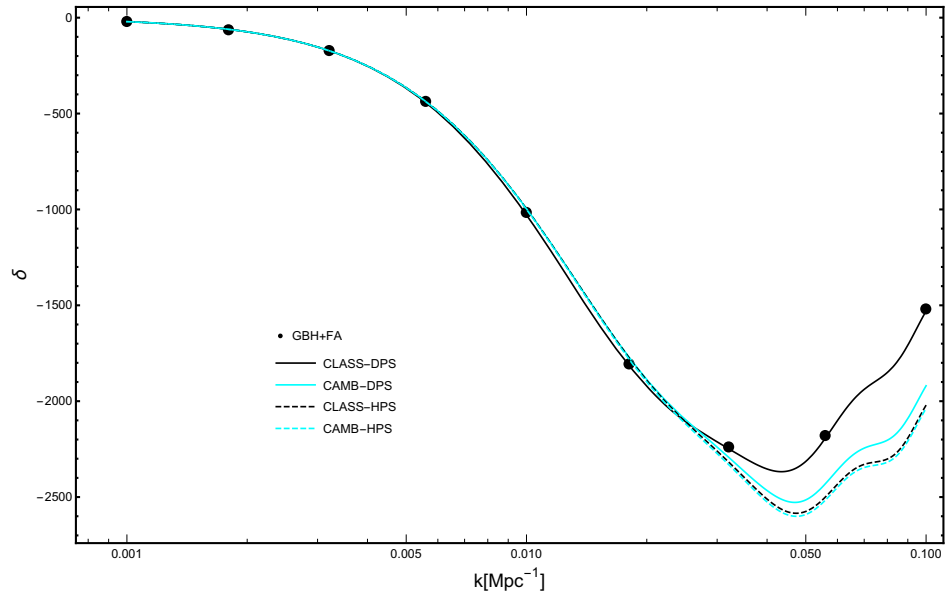

Figure 3.4: Density contrast neutrino transfer functions, as a function of scale and at redshift $z = 0$, from the GBH while switching to a fluid approximation at $x_{\max} = 15$ (GBH + FA), CLASS in both default (CLASS-DPS) and high (CLASS-HPS) precision settings, and CAMB in both default (CAMB-DPS) and high (CAMB-HPS) precision settings as well. Here we set the neutrino mass to $m = 0.1$eV. We verified that choosing a value of $x_{\max} = 30$ for the turning point between the GBH and the FA produced very similar results for the black dots. Also, the neutrino transfer function from CAMB is originally in the synchronous gauge, so we had to perform a gauge transformation to the Newtonian gauge to produce the curves in cyan.

effects, and nonlinearities, start to kick in, and hence the less important it becomes that the neutrino transfer functions are produced very accurately, which enables a switch to a fluid approximation once a given mode becomes smaller than the free-streaming scale (GBH+FA).

For a given scale and accuracy goal, the GBH approach involves solving at least roughly the same number of equations as a standard Boltzmann solver would, but completely avoids the inconvenience, and computational challenges, associated with solving the hierarchy to

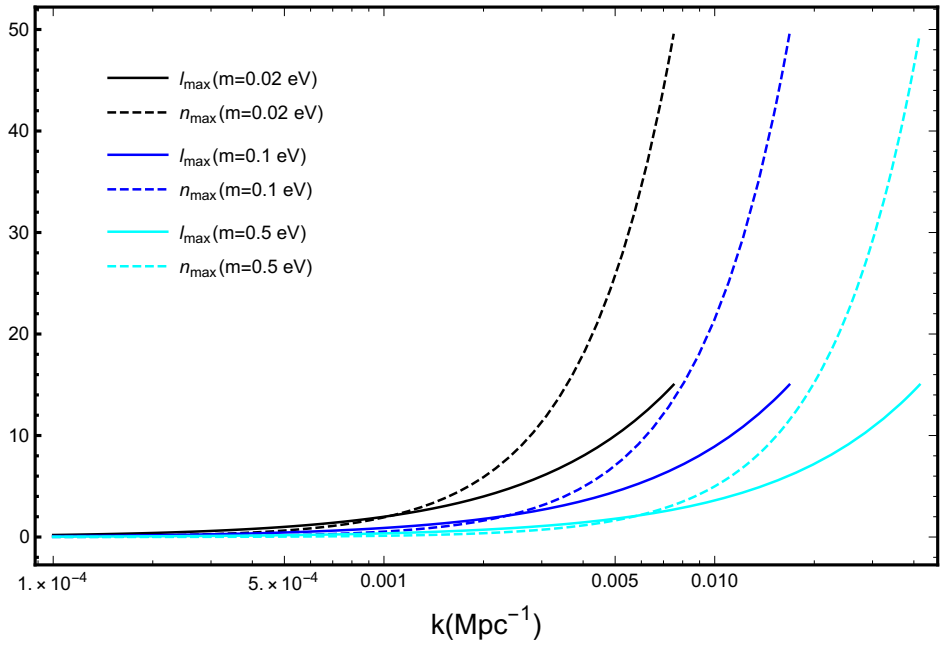

Figure 3.5: Choices of $l_{\max}$ and $n_{\max}$, as given by Eq. (3.12) to approximately ensure convergence of the GBH, as a function of scale for varying neutrino mass, and up to $x = 30$. For a given fixed scale, the dimension of the GBH increases for smaller neutrino mass. In fact, since T $\sim 1/\sqrt{m}$, we can use Eq. (3.12) to conclude that $l_{\max} \sim 1/\sqrt{m}$, while $n_{\max} \sim 1/m^{0.8}$.

later integrate over momentum space, a procedure which is known to be computationally slow, since the momentum integrals are performed to compute neutrino fluid properties, which in turn are coupled to the perturbations to the other components in the universe via the Einstein equations. This feature makes it plausible that the GBH may be faster than the standard approach.

In this work, we considered the numerical implementation of the GBH in a simplified scenario, where the sources $\phi(a)$ and $\psi(a)$ are obtained directly from CLASS. In this way, we do not need to solve the Einstein equations, and can evolve neutrino perturbations alone, with the goal of providing a proof-of-principle for the GBH. The full problem of implementing the GBH in a Boltzmann solver (allowing for a comparison of performance between the GBH

and the standard approach), is left for future work.

Being a novel approach to the inclusion of massive neutrinos in the computation of cosmological perturbations in linear theory, the GBH is interesting in its own right. It can provide some insight, and serve as cross check for the standard approach. Moreover, while both CLASS and CAMB are now greatly optimized for speed and performance, it is always worthwhile to explore new numerical approaches. [3] This is because cosmological analyses often employ Markov chain Monte Carlo (MCMC) methods to map the likelihood in a multidimensional parameter space, requiring Boltzmann codes to be run multiple times, and thus demanding significant computational resources.

In future work, we plan to implement the GBH+FA in the Boltzmann solver CLASS. Based in our findings, we expect to obtain similar accuracy as in the standard approach in default precision settings, and, we surmise, reduced computational time. Furthermore, with a more efficient numerical implementation of the GBH, we will be able to assess its applicability to produce accurate neutrino transfer functions on the small scales, i.e. for $x > 30$. This is a current limitation of the GBH, that stops it from producing neutrino transfer functions as accurately as existing Boltzmann solvers in high precision settings, on the small scales (see Fig. 3.4). However, we reinforce that because of neutrino free-streaming, the accuracy currently achieved with the GBH+FA is good enough for the vast majority of projects that involve massive neutrinos.

3.6 Appendix: Truncation scheme

We now look for a well-defined truncation scheme for the GBH: For a given neutrino mass m and scale k, we must be able to find values of $l_{\rm max}$ and $n_{\rm max}$ for which the GBH accurately produces the neutrino transfer functions, and becomes insensitive to a further increase in these parameters. This is, of course, just the statement of convergence. From our experience with the Boltzmann hierarchy, we expect that convergence with respect to $l_{\rm max}$ is not hard

[3] This is especially true for the implementation of massive neutrinos, one of the most time-consuming tasks of Boltzmann codes.

to achieve. In fact, we found that the following truncation, suggested in [19]:

$$\Psi_{l_{\max}+1} \approx \frac{(2l_{\max}+1)\epsilon}{qk\tau}\Psi_{l_{\max}} - \Psi_{l_{\max}-1} \tag{3.14}$$

is compatible with the structure of the GBH, and produces good results. Substitution of Eq. (3.14) into Eq. (3.6) yields

$$f_{l_{\max}+1,n} \approx (l_{\max}+1)\Big(\frac{1}{k\tau}f_{l_{\max},n} - \frac{l_{\max}}{4l_{\max}^2-1}f_{l_{\max}-1,n+1}\Big) \tag{3.15}$$

for $l_{\max} > 1$. Next we move on to the truncation with respect to $n_{\max}$. First notice that

$$f_{l,n_{\max}+1} - f_{l,n_{\max}} \sim \int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n_{\max}+l}\Big[1-\Big(\frac{q}{\epsilon}\Big)^2\Big]\Psi_l \tag{3.16}$$

This integrand contains a term with the form $f(y) = y^k(1-y^2)$, for $y \to q/\epsilon$. This goes to zero in both the relativistic and nonrelativistic regimes, with a peak in between that goes as $1/k$ for $k \gg 1$. If our truncation scheme is based on finding an approximate expression for Eq. (3.16), it seems reasonable to assume that convergence will be achieved for high enough $n_{\max}$. We proceed in analogy to what is done in [127]: Find an approximate expression to

$$\Lambda_0 \overset{\text{def}}{=} \frac{\delta_{n_{\max}+1}}{\delta_{n_{\max}}} = \frac{\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n_{\max}+2}\Psi_0}{\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n_{\max}}\Psi_0} \tag{3.17a}$$

$$\Lambda_l \overset{\text{def}}{=} \frac{f_{l,n_{\max}+1}}{f_{l,n_{\max}}} = \frac{\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n_{\max}+l+2}\Psi_l}{\int_0^\infty dqq^2 f_0(q)\epsilon\Big(\frac{q}{\epsilon}\Big)^{2n_{\max}+l}\Psi_l} \quad \forall\ l \geq 1 \tag{3.17b}$$

based on an educated guess on the q/ϵ dependence of the multipoles Ψ_l. In order to investigate this carefully, let us go back to the Boltzmann hierarchy in Eq. (3.7). After setting

$$\Psi_l = -\frac{d\ln f_0}{d\ln q}\tilde{\Psi}_l \tag{3.18}$$

and introducing a new (q-dependent) time variable,

$$\mathrm{T} = \int_i d\tau\frac{q}{\epsilon} \tag{3.19}$$

along with $x = k\mathrm{T}$, the Boltzmann hierarchy reads

$$\frac{d\tilde{\Psi}_0}{dx} = -\tilde{\Psi}_1 + \frac{d\phi}{dx} \tag{3.20a}$$

$$\frac{d\tilde{\Psi}_1}{dx} = \frac{1}{3}(\tilde{\Psi}_0 - 2\tilde{\Psi}_2) + \frac{1}{3}\tilde{\psi} \tag{3.20b}$$

$$\frac{d\tilde{\Psi}_l}{dx} = \frac{1}{2l+1}[l\tilde{\Psi}_{l-1} - (l+1)\tilde{\Psi}_{l+1}] \tag{3.20c}$$

This is the same set of equations one would find for radiation, but in terms of the time parameter T, and a (q-dependent) effective gravitational potential $\tilde{\psi} = (\epsilon/q)^2\psi$. This has two important consequences: First, all dependence on scales is actually encoded in $x = k\mathrm{T}$, i.e. horizon crossing is effectively defined by the condition that $k\mathrm{T} \sim 1$. Second, the mass dependence is encoded in x, but also in the effective gravitational potential, and in the nonrelativistic limit it dominates the right-hand side of the evolution equation for $\tilde{\Psi}_1$: This is just the well-known decoupling of $l < 2$ from higher multipoles in the nonrelativistic regime. It is then true that

$$\tilde{\Psi}_1 \sim \int dx\ \tilde{\psi} \sim \int d\tau \frac{\epsilon}{q}\psi \tag{3.21}$$

and hence $\tilde{\Psi}_1 \propto \epsilon/q$ to leading order, where we think of expanding $\tilde{\Psi}_l$ in a power series on q/ϵ around the nonrelativistic regime. Substitution of this into Eq. (3.20) now implies that $\tilde{\Psi}_0 \propto 1$ and $\tilde{\Psi}_l \propto (q/\epsilon)^{l-2}$ for $l \geq 1$, to leading order. This, combined with Eq. (3.18), are used on Eq. (3.11)

$$\Lambda_0 \approx \frac{\int_0^\infty dq q^2 f_0(q)\epsilon \frac{d\ln f_0}{d\ln q}\left(\frac{q}{\epsilon}\right)^{2(n_{\max}+1)}}{\int_0^\infty dq q^2 f_0(q)\epsilon \frac{d\ln f_0}{d\ln q}\left(\frac{q}{\epsilon}\right)^{2n_{\max}}} \tag{3.22a}$$

$$\Lambda_l \approx \frac{\int_0^\infty dq q^2 f_0(q)\epsilon \frac{d\ln f_0}{d\ln q}\left(\frac{q}{\epsilon}\right)^{2(n_{\max}+l)}}{\int_0^\infty dq q^2 f_0(q)\epsilon \frac{d\ln f_0}{d\ln q}\left(\frac{q}{\epsilon}\right)^{2(n_{\max}+l-1)}} \quad l \geq 1. \tag{3.22b}$$

After integration by parts, this can be written solely in terms of the background pressures

P_n (or equation of state parameters ω_n) as follows:

$$\Lambda_0 \approx \frac{2n_{\max}+5}{2n_{\max}+3}\frac{\omega_{n_{\max}+1}}{\omega_{n_{\max}}}\frac{1-\frac{2n_{\max}+1}{2n_{\max}+5}\frac{\omega_{n_{\max}+2}}{\omega_{n_{\max}+1}}}{1-\frac{2n_{\max}-1}{2n_{\max}+3}\frac{\omega_{n_{\max}+1}}{\omega_{n_{\max}}}} \tag{3.23a}$$

$$\Lambda_l \approx \frac{2(n_{\max}+l)+3}{2(n_{\max}+l)+1}\frac{\omega_{n_{\max}+l}}{\omega_{n_{\max}+l-1}}\frac{1-\frac{2(n_{\max}+l)-1}{2(n_{\max}+l)+3}\frac{\omega_{n_{\max}+l+1}}{\omega_{n_{\max}+l}}}{1-\frac{2(n_{\max}+l)-3}{2(n_{\max}+l)+1}\frac{\omega_{n_{\max}+l}}{\omega_{n_{\max}+l-1}}} \tag{3.23b}$$

There is only one final piece of information that needs to be specified in order to complete the truncation scheme: How to choose the values of $l_{\max}$ and $n_{\max}$. We know that higher multipoles and higher velocity weight fluid properties contribute as small scale effects, acting as viscosity, since on large scales a simple fluid approximation suffices. Based on this, and the observations made following Eq. (3.20), we expect that higher values of $l_{\max}$ and $n_{\max}$ are needed as $x = k\mathrm{T}$ increases.

Due to its importance, let us stop for a moment to study the time variable T, defined in Eq. (3.19). During the relativistic regime, it is identical to τ. Let us now see what happens in the nonrelativistic regime, assuming that the time of transition $a_{\mathrm{tr}} \sim q/m$, happens during matter domination, as is the case for massive neutrinos. We may then write the following approximation:

$$\mathrm{T} \approx \tau_{\mathrm{tr}} + \int_{\mathrm{tr}} \frac{da}{a}\frac{1}{aH}\frac{q}{ma} \tag{3.24}$$

where we split the integral from the initial time to the transition, and from the transition to the final time, use $d\tau = \frac{da}{a}\frac{1}{aH}$, with $H = \frac{\mathcal{H}}{a}$ is the Hubble rate, and approximate $\epsilon \approx ma$ in the nonrelativistic regime, along with $\epsilon \approx q$ up to the transition. Further using $H \sim a^{-3/2}$ during matter domination, one obtains for the integral in the right-hand side of Eq. (3.24)

$$\int_{\mathrm{tr}} \frac{da}{a}\frac{1}{aH}\frac{q}{ma} \sim \frac{q}{m}\int_{\mathrm{tr}} \frac{da}{a}a^{-\frac{1}{2}} \sim \frac{q}{m}a_{\mathrm{tr}}^{-\frac{1}{2}} \sim \sqrt{\frac{q}{m}}, \tag{3.25}$$

where we use the fact that the integral is dominated by its lower limit, and use $a_{\mathrm{tr}} \sim q/m$. Further notice that during matter domination $\tau_{\mathrm{tr}} \sim a_{\mathrm{tr}}^{1/2} \sim \sqrt{\frac{q}{m}}$ as well, so $\mathrm{T} \sim \sqrt{\frac{T_0}{m}}$ approaches a time-independent constant, as opposed to τ, which grows indefinitely. In other words, T grows like τ up to the time of transition, effectively freezing as one approaches the nonrelativistic regime. Indeed, since $q/\epsilon = v$ is the neutrino velocity, T is the comoving

distance traveled by a neutrino particle through cosmic history. When evaluated at the peak of the Fermi-Dirac distribution (say $q/T_0 = 3$), this is roughly the neutrino horizon, or the free-streaming scale (integrated over e-folds). From this point forward, when used as a time variable, it is implicitly assumed that T is evaluated at $q/T_0 = 3$, and hence corresponds to the neutrino horizon (plotted in Fig. 3.1).

On the large scales, i.e. $x = k\mathrm{T} \lesssim 1$, a simple viscous fluid approximation with $l_{\max} = 2$ and $n_{\max} = 0$ suffices, and we discuss this in detail in Appendix 3.7.

As $x \gg 1$, we expect that higher values of both $l_{\max}$ and $n_{\max}$ are needed. A naive assumption would then be that $l_{\max}, n_{\max} \propto x$. Indeed, our experience with the GBH indicates that it converges for $l_{\max} \approx x/2$, with the difference between T and τ explaining why one needs a much higher $l_{\max}$ for radiation than for massive neutrinos [19]. Unfortunately, the same cannot be said about the n expansion, with $n_{\max}$, and hence the dimensionality of the system, growing very rapidly for modes inside the horizon. This can be explained as follows: The higher velocity weight variables are effectively accounting for the dynamics of neutrinos with higher particle velocities. This means that T, as defined by Eq. (3.19), should not be evaluated at the peak of the Fermi-Dirac distribution $q/T_0 \approx 3$, but rather at a larger value $q = q_{n_{\max}}$, that we choose to be the peak of the integrand in Eq. (3.2), for the corresponding value of $n = n_{\max}$. This is to make sure that $\mathrm{T}(q_{n_{\max}})$ is the comoving distance traveled by the neutrino particles that are actually being probed by the higher-velocity weight fluid properties. We then expect that

$$\frac{n_{\max}}{x} \sim \frac{k\mathrm{T}(q_{n_{\max}})}{x} = \frac{\mathrm{T}(q_{n_{\max}})}{\mathrm{T}} \tag{3.26}$$

We evaluated Eq. (3.26) numerically for many values of $n_{\max}$, to find that $n_{\max} \sim x^{1.6}$ up to $x = 30$. From our experience with the GBH, $n_{\max} \approx x^{1.6}/5$ is approximately sufficient for convergence.

Before moving on to the viscous fluid approximation, we should point out that our truncation scheme in Eq. 3.23 requires accurate numerical evaluation of quotients involving higher-velocity weight equations of state, which according to Eq. (3.2), become really small numbers

in the nonrelativistic regime. The numerical computation of Λ_l can then be very time consuming, especially for high $n_{\rm max}$. However, as it is clear from Eq. (3.2), ω_n is a function of just a single variable $y = ma/T_0$, and hence can be easily tabulated.

3.7 Appendix: Fluid approximation

On large scales for which $x = k\mathrm{T} \lesssim 1$, a simple viscous fluid approximation with $l_{\rm max} = 2$ and $n_{\rm max} = 0$ should suffice. Since in the nonrelativistic regime T basically freezes at $\tau_{\rm tr}$, the value of conformal time evaluated at the transition $a_{\rm tr} \sim \frac{q}{m} \sim \frac{T_0}{m}$, this can be rephrased as to say that the mode has to be superhorizon at the transition, i.e. in enters the horizon during the nonrelativistic regime. This is exactly what was found in [129], from comparing the exact solution with a simple fluid approximation.

Furthermore, on the small scales $x \gg 1$, free-streaming and nonlinear effects start to kick in and it no longer becomes important that the neutrino transfer function is produced very accurately. One can then choose a $x_{\rm max}$ above which a simple fluid approximation can be used once again. Let us then stop for a moment to carefully study the case $l_{\rm max} = 2$ and $n_{\rm max} = 0$. The fluid equations are given by Eq. (3.9). Our truncation scheme (TS), developed in Appendix 3.6, provides approximations for the quantities $\delta P/\delta\rho$, Σ, Θ and kf_3, in terms of the dynamical variables in the system.

Setting $l_{\rm max} = 2$ and $n_{\rm max} = 0$ in Eq. (3.15) (which is known not to be a particularly good approximation for $l_{\rm max} = 2$ [148]), one obtains

$$kf_3 \approx \frac{3}{\tau}\sigma - \frac{2}{5}\Theta \tag{3.27}$$

Both sides of this equation can be accurately determined at intermediate scales, from the GBH. A comparison can be found in Fig. 3.6.

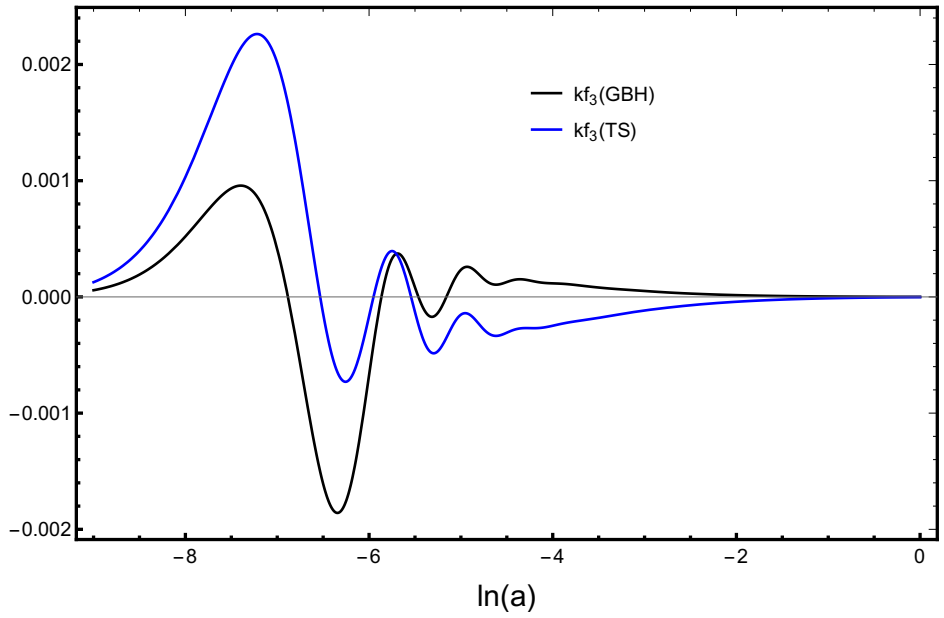

Figure 3.6: Left (GBH) and right-hand (TS) sides of Eq. (3.27) as obtained from the GBH. Here $x = 15$, or $k = 0.008\text{Mpc}^{-1}$, $m = 0.1\text{eV}$, $l_{\max} = 8$, and $n_{\max} = 16$. The approximation reproduces the right features, but is not accurate.

Furthermore, set $n_{\max} = 0$ into Eq. (3.23) to arrive at

$$\frac{1}{3}\Lambda_0 = \frac{\delta P}{\delta \rho} \approx \frac{5}{3}\frac{\omega}{1+\omega}\left(1 - \frac{1}{5}\frac{\omega_2}{\omega_1}\right) = c_g^2 \tag{3.28a}$$

$$\Lambda_1 = \frac{\Theta}{\theta} \approx 3c_g^2 \iff c_{\text{vis}}^2 = \frac{3}{4}(1+\omega)c_g^2 \tag{3.28b}$$

$$\Lambda_2 = \frac{\Sigma}{\sigma} \approx \frac{7}{5}\frac{\omega_2}{\omega_1}\frac{1 - \frac{1}{7}\frac{\omega_3}{\omega_2}}{1 - \frac{1}{5}\frac{\omega_2}{\omega_1}} \tag{3.28c}$$

where a fluid viscosity speed c_{vis}, was introduced as a different parametrization to Λ_1, using notation from [143]

$$c_{\text{vis}}^2 = \frac{1}{4}(1+\omega)\frac{\Theta}{\theta} \tag{3.29}$$

In the CLASS ncdm fluid approximation, the truncation in the multipole is done in the exact same way as in Eq. (3.27), while for the higher velocity weight quantities, the authors

of [127] apply a bit of trial and error to arrive at the following *ad hoc* approximations

$$\frac{1}{3}\Lambda_0 = \frac{\delta P}{\delta \rho} \approx \frac{5}{3}\frac{\omega}{1+\omega}\left(1 - \frac{1}{5}\frac{\omega_2}{\omega_1}\right) = c_g^2 \tag{3.30a}$$

$$\Lambda_1 = \frac{\Theta}{\theta} \approx 12\frac{\omega}{1+\omega}c_g^2 \iff c_{\text{vis}}^2 = 3\omega c_g^2 \tag{3.30b}$$

$$\Lambda_2 = \frac{\Sigma}{\sigma} \approx \frac{\omega_2}{\omega_1} \tag{3.30c}$$

The Eqs. (3.28) and (3.30) differ slightly on the expressions for Θ/θ and Σ/σ.

In Fig. 3.7 we compare the adiabatic sound speed squared to the exact solutions coming from both CLASS and the GBH: The GBH and CLASS agree to a subpercent level, with both differing from the assumption of adiabaticity when approaching the nonrelativistic regime.

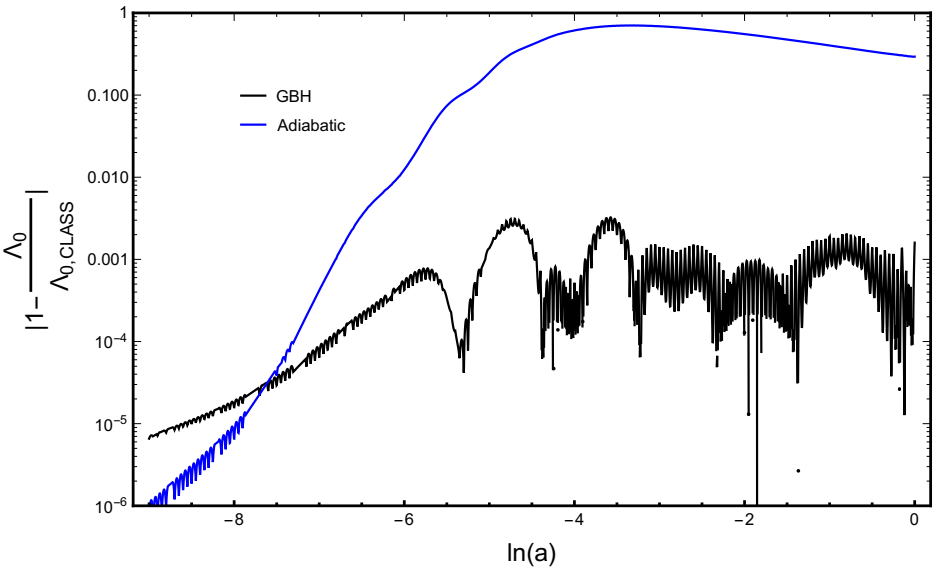

Figure 3.7: Relative difference between sound speeds squared, coming from the GBH and the assumption of adiabaticity, when compared to the exact solution from CLASS. Here $x = 15$, or $k = 0.008\text{Mpc}^{-1}$, $m = 0.1\text{eV}$, $l_{\text{max}} = 8$, and $n_{\text{max}} = 16$.

In Fig. 3.8, we compare the Λ_1's from CLASS and GBH truncation schemes for the FA with the exact solution from the GBH: There is an overall 10% level error in both cases, but the GBH truncation scheme is an order of magnitude better in the nonrelativistic regime.

In Fig. 3.9, we compare Λ_2's from CLASS and GBH truncation schemes for the FA with the exact solution from the GBH: there is a 10% level error in both cases, with the CLASS truncation scheme being an order of magnitude better in the nonrelativistic regime. However, we found that the accuracy of the FA is insensitive to the specific choice of Λ_2.

Finally, in Fig. 3.10 we compare fluid approximations, with GBH and CLASS truncation schemes, with the exact solution from CLASS: there is also an overall 10% level error in the neutrino density contrast transfer function, but the CLASS fluid approximation works better at late times. Indeed, this is possible because the CLASS truncation scheme is tuned to produce the best outcome, even though all the individual approximations are in a similar level of accuracy as in the truncation scheme for the GBH, which was motivated from first principles. Because of this, the CLASS fluid approximation works better overall, and should be the one used in the regime $x > x_{\rm max}$, as discussed.

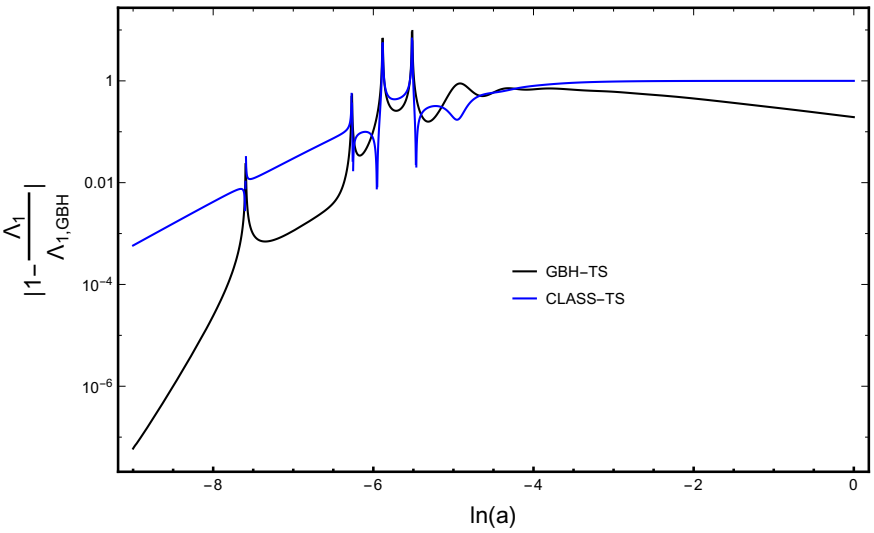

Figure 3.8: Relative difference between Λ_1's, coming from CLASS and GBH truncation schemes, when compared to the exact solution from the GBH. Here $x = 15$, or $k = 0.008\text{Mpc}^{-1}$, $m = 0.1\text{eV}$, $l_{\rm max} = 8$, and $n_{\rm max} = 16$.

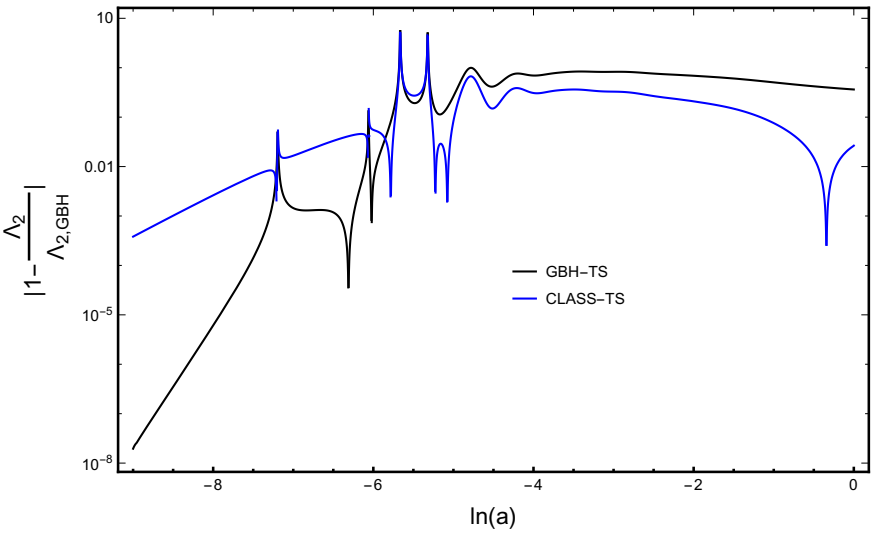

Figure 3.9: Relative difference between Λ_2's, coming from CLASS and GBH truncation schemes, when compared to the exact solution from the GBH. Here $x = 15$, or $k = 0.008\mathrm{Mpc}^{-1}$, $m = 0.1\mathrm{eV}$, $l_{\max} = 8$, and $n_{\max} = 16$.

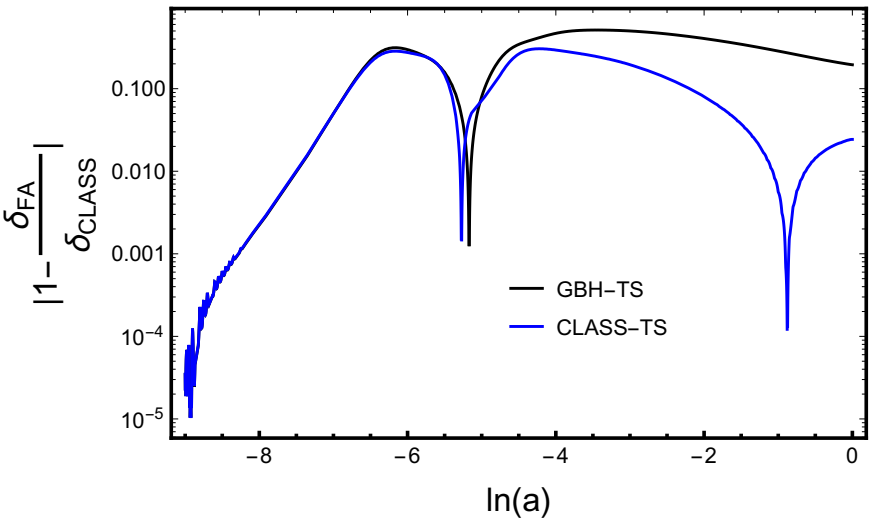

Figure 3.10: Relative difference between density contrast neutrino transfer functions coming from the fluid approximation with GBH and CLASS truncation schemes, when compared to the exact solution from CLASS. Here $x = 15$, or $k = 0.008\mathrm{Mpc}^{-1}$, and $m = 0.1\mathrm{eV}$.

Chapter 4

ACCURATE FLUID APPROXIMATION FOR MASSIVE NEUTRINOS IN COSMOLOGY

4.1 Abstract

A measurement of the neutrino mass scale will be achieved with cosmological probes in the upcoming decade. On one hand, the inclusion of massive neutrinos in the linear perturbation theory of cosmological structure formation is well understood and can be done accurately with state of the art Boltzmann solvers. On the other hand, the numerical implementation of the Boltzmann equation is computationally expensive and is a bottleneck in those codes. This has motivated the development of more efficient fluid approximations, despite their limited accuracy over all scales of interest, $k \sim (10^{-3} - 10)\mathrm{Mpc}^{-1}$. In this work we account for the dispersive nature of the neutrino fluid, i.e., the scale dependence in the sound speed, leading to an improved fluid approximation. We show that overall $\lesssim 5\%$ errors can be achieved for the neutrino density and velocity transfer functions at redshift $z \lesssim 5$, which corresponds to an order of magnitude improvement over previous approximation schemes that can be discrepant by as much as a factor of two.

4.2 Introduction

The observation of neutrino oscillations has established that at least two of the neutrino mass eigenstates have a non-zero mass, with an associated lower bound on the sum of the masses of $\sum_\nu m_\nu \gtrsim 0.06$ and 0.1eV for the normal and inverted hierarchies respectively [149–151]. Complementary information comes from beta decay experiments, which set an upper bound to a weighted sum of the masses $m_{\nu,\beta} < 0.8\mathrm{eV}$ [152]. Additionally, massive neutrinos suppress cosmological structure formation at small scales [64], leading to the most

stringent upper bound on the sum of neutrino masses to date, i.e., $\sum_\nu m_\nu < 0.12\text{eV}$ [11]. [1] It is expected that future cosmological surveys in the upcoming decade will be sensitive to the lower bound from oscillation experiments and hence will allow for a detection of the neutrino mass scale [156–158]. This is a crucial measurement since it sets a clear target for laboratory experiments and serves as a cross-check of the consistency between particle physics and cosmology [159].

The inclusion of massive neutrinos in linear cosmological perturbation theory has a long history (see [19] and references therein). Due to the large velocity dispersion of massive neutrinos, one must go beyond a simple fluid treatment and solve a hierarchy of Boltzmann equations in phase space. At each time step the neutrino distribution function is then integrated over momenta to obtain the neutrino stress-energy tensor, which in turn contributes to the right hand side of Einstein's equations and sets the coupling of neutrinos to the other species in the universe. This is a cumbersome procedure and a computational bottleneck in state of the art Boltzmann solvers, such as the Cosmic Linear Anisotropy Solving System (CLASS) [127] and the Code for Anisotropies in the Microwave Background (CAMB) [160].

As a consequence, the search for more efficient alternative approaches to the inclusion of massive neutrinos in linear cosmological perturbation theory remains a well motivated direction of research since modern cosmological parameter inference techniques require these codes to be run tens or hundreds of thousands of times. For instance, [161] formulates the problem as an integral equation and proposes an iterative solution, while [79] integrates out the momentum dependence at the cost of a significant increase in the dimensionality of the resulting system of ordinary differential equations.

An alternative approach consists of a simple fluid approximation for the exact neutrino dynamics. This is a viable option whenever the neutrinos only give a small contribution to the total matter energy density, since we can then afford for some inaccuracies in the neutrino density provided we are only interested in the total matter (or cold dark matter)

[1] This upper bound can be relaxed with nonstandard scenarios such as an unstable neutrino species and dynamical dark energy [153–155].

field. This is especially true on small scales where the neutrino density is suppressed due to free-streaming and the cold dark matter evolution basically decouples from the neutrinos, the same circumstances in which the Boltzmann hierarchy needs to be truncated at a large multipole and becomes computationally expensive. Indeed, a fluid approximation is used in CLASS to evolve the neutrino component on scales that are much smaller than the cosmological horizon [127]. Another fluid approximation for massive neutrinos follows from the generalized dark matter approach of [143].

What all fluid approximations have in common is that they become inaccurate at sufficiently small scales [129], exactly in the regime where the approximation is the most useful since the exact dynamics is more (computationally) expansive as we discussed above. In this work we show that this failure of the fluid approximation at small scales is mostly a result of *not* accounting for the dispersive nature of the neutrino fluid, i.e., the sound speed is scale dependent [162]. In previous works much of the focus was directed towards modeling the evolution of the neutrino shear stress implicitly presuming that the assumption of an adiabatic sound speed does not dominate the total error [127]. Instead, we find that the assumption of an adiabatic sound speed leads to a significant overestimation of the sound speed on small scales that dominates the error in the fluid approximation.

We obtain a simple analytic expression for the sound speed at small scales and use it to introduce a scale dependent approximation to this quantity that interpolates between the small and large scale regimes. This, in combination with a scale dependent approximate expression to the anisotropic stress, leads to a resulting fluid approximation for massive neutrinos with $\lesssim 5\%$ errors for the neutrino density and velocity transfer functions at redshift $z \lesssim 5$ and over scales $k = (10^{-3} - 10)\mathrm{Mpc}^{-1}$, which corresponds to an order of magnitude improvement over previous approximation schemes that can be as much as a factor of two wrong.

We consider neutrino masses in the range $0.02\mathrm{eV} \leq m_\nu \leq 0.5\mathrm{eV}$, for which a $\lesssim 5\%$ error in the neutrino component is sufficient to produce the total matter power spectrum to sub-percent level accuracy. This fluid approximation is then a powerful alternative to the

full Boltzmann hierarchy for most projects, allowing for a significant reduction in computing time.

The paper is organized as follows: In Section 4.3 we introduce the fluid equations and the approximate expressions for the sound speed and anisotropic stress in the Newtonian gauge. In Section 4.4 we compare our fluid approximation with the results from CLASS in both high and default precision settings, along with the fluid approximation used in CLASS. In Section 4.5 we summarize our results. Details of calculations that motivate the approximations employed can be found in Appendix 4.6, and in Appendix 4.7 we extend our fluid approximation to alternative gauges (other than the Newtonian gauge), showing explicit expressions in the synchronous gauge.

4.3 Fluid equations

The fluid equations satisfied by massive neutrinos are quite generic as they follow from energy-momentum conservation laws. In this section we first introduce the relevant equations (referring the reader to [19] for further details). Then we briefly motivate and write down formulas that approximate the scale dependent sound speed and anisotropic stress. Derivations and details can be found in Appendix 4.6.

We consider scalar perturbations to the Friedmann-Lemaître-Robertson-Walker (FLRW) universe in the (conformal) Newtonian gauge, where the metric reads

$$ds^2 = a^2(\tau)\left[-(1+2\psi)d\tau^2 + (1-2\phi)d\vec{x}^2\right], \tag{4.1}$$

and where $a(\tau)$ is the scale factor, τ is the conformal time (related to the cosmic time t via the expression $dt = ad\tau$), $\vec{x}$ are comoving spatial coordinates, $\phi(\tau,\vec{x})$ and $\psi(\tau,\vec{x})$ are gravitational potentials that we treat as (small) linear perturbations. We also define the conformal Hubble rate $\mathcal{H} = a'/a$, where throughout a prime denotes derivative with respect to conformal time τ.

At the level of background the neutrinos are distributed in phase-space with the rela-

tivistic Fermi-Dirac profile, $f_0(q) = f_{\rm FD}(q/T_{\nu,0})$, where

$$f_{FD}(x) = \frac{g_\nu}{e^x + 1}\,, \tag{4.2}$$

with q the magnitude of the comoving momentum $\vec{q}$, $T_{\nu,0} \approx 1.95\text{K} \approx 1.7 \times 10^{-4}\text{eV}$ is the neutrino temperature today, $x = q/T_{\nu,0}$ and $g_\nu = 2$ to account for both left-handed neutrinos and right-handed antineutrinos. From the condition of isotropy the only non-vanishing components of the neutrino stress-energy tensor are its energy density and pressure. They can be obtained from Eq. (4.2) as follows:

$$\rho(a) = a^{-4}\int_0^\infty \frac{dq}{2\pi^2} q^2 \epsilon(q,a) f_0(q)\,, \tag{4.3}$$

$$P(a) = \frac{1}{3}a^{-4}\int_0^\infty \frac{dq}{2\pi^2} q^2 \epsilon(q,a) \left[\frac{q}{\epsilon(q,a)}\right]^2 f_0(q)\,. \tag{4.4}$$

Here $\epsilon(q,a) = \sqrt{q^2 + m_\nu^2 a^2}$ is the comoving energy. Note that these are not the total energy density and pressure of the Universe, as they refer only to the neutrino species.

It will also prove useful to define the equation of state as:

$$w(a) = \frac{P(a)}{\rho(a)}\,. \tag{4.5}$$

In the presence of non-vanishing gravitational potentials in Eq. (4.1), the neutrino density and pressure also acquire perturbations, $\delta\rho(\tau,\vec{x})$ and $\delta P(\tau,\vec{x})$, that are time and position dependent. Additionally, there is a net bulk flow that we parameterize by the divergence in the velocity field $\theta(\tau,\vec{x})$, and an anisotropic (shear) stress $\sigma(\tau,\vec{x})$. We follow the standard practice of working in terms of a density contrast $\delta(\tau,\vec{x}) = \delta\rho(\tau,\vec{x})/\rho(a)$, and define the sound speed as

$$c_{\rm s}^2(\tau,\vec{x}) = \frac{\delta P(\tau,\vec{x})}{\delta\rho(\tau,\vec{x})}\,. \tag{4.6}$$

We now have all the ingredients to write down the fluid equations which are exact and

follow from the conservation of the neutrino stress-energy tensor [2],

$$\delta' = -(1+w)(\theta - 3\phi') - 3\mathcal{H}(c_{\rm s}^2 - w)\delta\,, \tag{4.7}$$

$$\theta' = -\mathcal{H}(1-3w)\theta - \frac{w'}{1+w}\theta + \frac{c_{\rm s}^2}{1+w}k^2\delta - k^2\sigma + k^2\psi\,. \tag{4.8}$$

In order to close the system of equations we need approximate expressions for both the sound speed $c_{\rm s}^2$ and the anisotropic stress $k^2\sigma$ (in the full Boltzmann hierarchy they can be obtained from the distribution function after an integration over momentum). A complete discussion on the motivations for our approximations (with relevant derivations) can be found in Appendix 4.6. Here we will just introduce the main ideas. At sufficiently large scales $c_{\rm s}^2$ approaches the so-called adiabatic sound speed,

$$c_{\rm g}^2(a) = \frac{P'(a)}{\rho'(a)}\,. \tag{4.9}$$

This follows from separate universe arguments: At sufficiently large scales the neutrino anisotropies can be absorbed into a local shift of the neutrino temperature, $T_\nu \to T_\nu(1+\mathcal{N}_\nu)$, with $\mathcal{N}_\nu = \delta T_\nu/T_\nu$ a constant. It then follows from Eq. (4.2) that the total distribution function reads,

$$f(q) = f_{\rm FD}\left(\frac{q}{T_{\nu,0}(1+\mathcal{N}_\nu)}\right) = f_0(q)\left(1 - \frac{d\ln f_0}{d\ln q}\mathcal{N}_\nu\right) \implies \delta f = -q\frac{df_0}{dq}\mathcal{N}_\nu\,, \tag{4.10}$$

where from the second to the third line we expanded to leading order in $\mathcal{N}_\nu = \delta T_\nu/T_\nu$. Eq. (4.9) can now be obtained upon integration over the comoving momentum to produce the neutrino density and pressure perturbations. Similarly, at sufficiently small scales $c_{\rm s}^2$ approaches what we call the asymptotic (asp) sound speed,

$$c_{\rm asp}^2(a) = \frac{1}{3}\frac{1+w(a)}{1+\lambda(a)}\,, \tag{4.11}$$

where the quantity $\lambda(a)$ is defined by,

$$\rho(a)\lambda(a) = \frac{1}{3}a^{-4}\int_0^\infty \frac{dq}{2\pi^2}q^2\epsilon(q,a)\left[\frac{\epsilon(q,a)}{q}\right]^2 f_0(q)\,. \tag{4.12}$$

[2]Moving forward we work in Fourier space $\vec{\nabla} \to i\vec{k}$ and often omit time and scale dependences of fluid properties for simplicity of notation.

Eq. (4.11) can be extracted from the static limit of the Boltzmann equation leveraging on the following observation: Neutrinos have a velocity $v \equiv q/\epsilon$ so that neutrino fluctuations with a (comoving) wavenumber k have a characteristic time scale, $t \sim a/kv$, which is much smaller than a Hubble time at scales $k \gg \mathcal{H}/v$. We can then consider the static limit where the expansion of the universe can be taken as slow and the general solution to the Boltzmann equation is an arbitrary function of the (adiabatic invariant) total comoving energy. At the background level this is $\epsilon_{\rm tot} = \epsilon = \sqrt{q^2 + m_\nu^2 a^2}$, and the distribution function is $f(q) = f_0(\sqrt{\epsilon^2 - m_\nu^2 a^2})$ as given by Eq. (4.2). However, in the presence of a gravitational potential ψ the total comoving energy reads $\epsilon_{\rm tot} = \epsilon(1+\psi)$. The total distribution function then becomes,

$$f = f_0(\sqrt{\epsilon^2(1+\psi)^2 - m_\nu^2 a^2}) = f_0(q)\left[1 + \frac{d\ln f_0}{d\ln q}\left(\frac{\epsilon}{q}\right)^2 \psi\right] \implies \delta f = q\frac{df_0}{dq}\left(\frac{\epsilon}{q}\right)^2 \psi \quad (4.13)$$

where we expand to leading order in ψ. Eq. (4.11) can then be derived by integrating over the comoving momenta to obtain the neutrino fluid properties. More details can be found in Appendix 4.6, where we systematically derive the Eqs. (4.9) and (4.13) as the large and small scale limits of the Boltzmann equation.

Our strategy will now be to interpolate between these two regimes in order to write an approximate expression for the sound speed that accounts for its scale dependence. We similarly also want to introduce a scale dependent approximate expression to the anisotropic stress. To accomplish these goals we first need to understand what are the characteristic scales associated to the neutrino thermal motion. Indeed, there are two time dependent scales. One is the (instantaneous) free-streaming scale $k_{\rm fs}(a)$ defined by,

$$k_{\rm fs}(a) = \sqrt{\frac{3}{2}\Omega_{\rm m}(a)}\frac{\mathcal{H}(a)}{c_{\rm asp}(a)}\,, \quad (4.14)$$

which is proportional to the (comoving) distance that neutrinos travel over the course of one expansion time $\tau \sim 1/\mathcal{H}$, i.e., $\lambda_{\rm fs} = 2\pi/k_{\rm fs} \sim c_{\rm asp}/\mathcal{H} \sim c_{\rm asp}t$, with $\Omega_{\rm m}(a)$ the fractional contribution of matter (including neutrinos) to the total energy budget. Notice from Eqs. (4.3), (4.4), (4.5), (4.11) and (4.12) that in the nonrelativistic regime, where $\epsilon \approx ma \gg q$, we have

that $\lambda \gg 1$ and $w \ll 1$ such that,

$$c_{\rm asp}(a) \approx \frac{1}{\sqrt{3\lambda(a)}} \approx \sigma_\nu(a) \tag{4.15}$$

where we have used Eq. (4.2), and introduce the neutrino velocity dispersion, [3]

$$\sigma_\nu(a) = \sqrt{\frac{3\zeta(3)}{\ln 4}}\frac{T_{\nu,0}}{m_\nu a}\,. \tag{4.16}$$

The other scale associated to the neutrino thermal motion is the neutrino horizon, $\lambda_{\rm hor}(a)$, defined by:

$$\lambda_{\rm hor}(a) = \int_0^a d\ln a' \lambda_{\rm fs}(a')\,, \tag{4.17}$$

or its wavenumber $k_{\rm hor}(a) = 2\pi/\lambda_{\rm hor}(a)$, where similarly $\lambda_{\rm fs}(a) = 2\pi/k_{\rm fs}(a)$. The neutrino horizon is proportional to the total distance traveled by neutrinos over the entire expansion history. In fact, at late times $k_{\rm fs} \gg k_{\rm hor}$ and both scales play a role in the dynamics of massive neutrinos: $k_{\rm hor}$ is the scale below which ($k \leq k_{\rm hor}$) neutrino velocities can be ignored, and hence neutrinos cluster like cold dark matter, while $k_{\rm fs}$ is the scale above which ($k \geq k_{\rm fs}$) the neutrino pressure dominates over the gravitational potential leading to the suppression of neutrino structure. This is illustrated in Fig.4.1 for an individual neutrino mass of $m_\nu = 0.1$eV. While the neutrino horizon grows with the expansion of the universe, the free-streaming scale peaks when the neutrinos first become nonrelativistic due to the subsequent decrease in the thermal velocity. This produces a large separation of scales at late times, which is the reason why galaxy surveys cannot directly probe the scale dependence of the neutrino suppression.

We are finally ready to write down the approximations we use, after which we compare to previous approximation schemes and explain the different terms involved. More details can be found in Appendix 4.6. The approximations are:

$$c_{\rm s}^2(a,k) = c_{\rm g}^2(a) + \left[c_{\rm asp}^2(a) - c_{\rm g}^2(a)\right] e^{-\frac{4}{3}\frac{k_{\rm fs}(a)}{k}}\,, \tag{4.18}$$

$$k^2\sigma(a,k) = -\frac{2}{5}\frac{k_{\rm hor}(a)}{k}e^{-\frac{k_{\rm hor}(a)}{k}}\frac{c_{\rm s}^2(a,k)}{1+w(a)}k^2\delta(a,k) + \frac{k}{k_{\rm fs}(a)}e^{-5\frac{k_{\rm fs}(a)}{k}}w^2(a)\theta(a,k)\,. \tag{4.19}$$

[3] In Appendix 4.6 we motivate our choice of Eq. (4.16) for the neutrino velocity dispersion, and hence the appearance of $c_{\rm asp}(a)$ in the definition of the free-streaming scale as in Eq. (4.14).

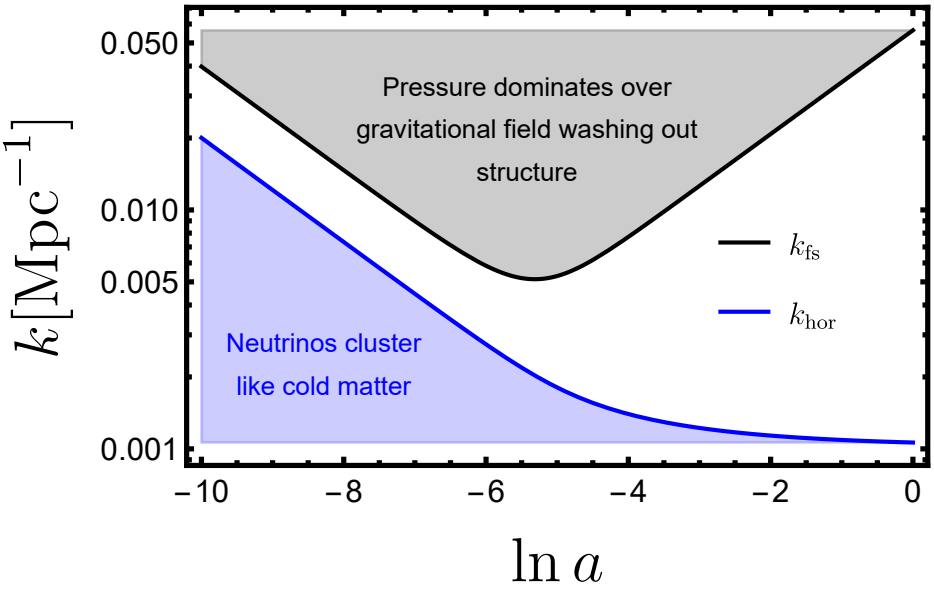

Figure 4.1: Neutrino free-streaming (black curve) and neutrino horizon (blue curve) scales as a function of the scale factor, for an individual neutrino mass of $m_\nu = 0.1\mathrm{eV}$. The black dashed region corresponds to sub free-streaming scales where pressure dominates over the gravitational field washing out structure, and the blue shaded region corresponds to scales above the neutrino horizon where neutrinos cluster like cold matter.

Eq. (4.18) is a simple interpolation between the two regimes given by Eqs. (4.9) and (4.11) [4]. It improves on previous approximation schemes in the literature that generally assume an adiabatic sound speed, since we account for deviations from adiabaticity on the small scales (which can be phrased as the presence of an entropy perturbation). This is illustrated by Fig. 4.2, where we compare both the adiabatic and asymptotic expressions [Eqs. (4.9) and (4.11) respectively] to the exact sound speed extracted from the Boltzmann code CLASS in high precision settings [5] at three different scales and for a neutrino mass $m_\nu = 0.1\mathrm{eV}$. As

[4]The precise numerical factors in the exponents of Eqs. (4.18) and (4.19) are adjusted in such a way as to optimize the fluid approximation.

[5]These are the high precision settings we employed in CLASS: `ncdm_fluid_approximation = 3` (this turns off the CLASS fluid approximation), `Quadrature strategy = 3`, `Maximum q = 15`, `Number of`

we move from larger to smaller scales the exact sound speed shifts from the adiabatic to the asymptotic formulas. Note that the adiabatic sound speed overestimates the exact sound speed (a result that holds true in both Newtonian and synchronous gauges).

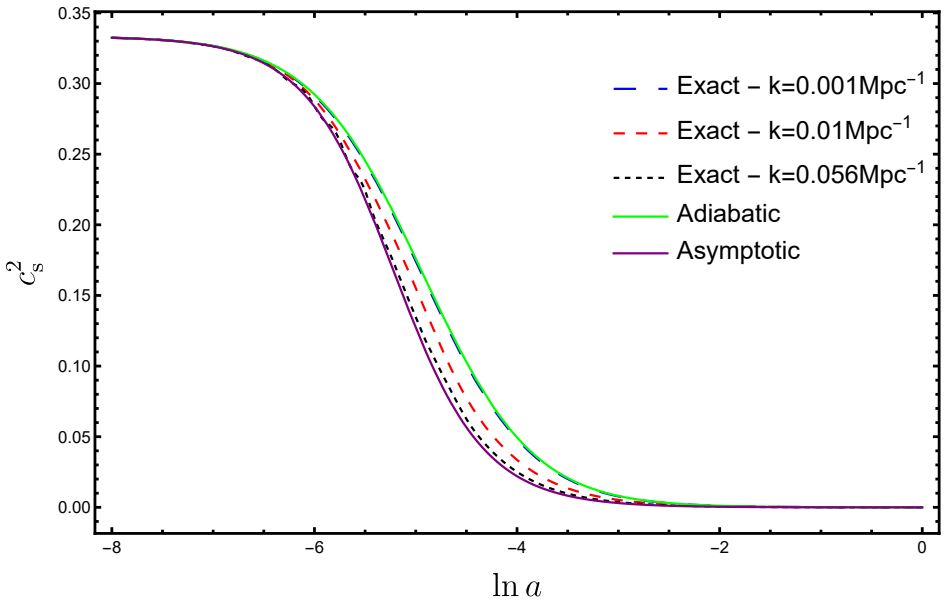

Figure 4.2: A comparison of both adiabatic and asymptotic sound speeds, as given by Eqs. (4.9) and (4.11), with the exact sound speed from the Boltzmann code CLASS in high precision settings (and in the Newtonian gauge) at three different scales and as a function of the scale factor. Here we choose $m_\nu = 0.1$eV. The green and purple solid curves correspond to the adiabatic and asymptotic sound speeds, respectively. The blue, red and black dashed curves correspond to the exact solutions for $k = 10^{-3}\text{Mpc}^{-1}$, $k = 10^{-2}\text{Mpc}^{-1}$, and $k = 5.6 \times 10^{-2}\text{Mpc}^{-1}$ respectively.

Next we move on to the anisotropic stress. As we argue in Appendix 4.6, we expect it to mostly give an additional contribution to the sound speed at scales that are around the neutrino horizon $k \sim k_{\text{hor}}$. This is accomplished by the first term in the right-hand side of Eq. (4.19). However, we generally also expect it to give a viscosity type contribution

---

`momentum bins = 30, l_max_ncdm = 30.`

proportional to the divergence of the velocity at small sub-free streaming scales, that turns out to be important for the numerical stability of the fluid equations. This leads to the second term in the right-hand size of Eq. (4.19); the precise time dependence is not important, but we choose the equation of state squared because it produces good results. The standard approach in the literature consists in modeling the evolution equation for the shear stress (see [127] for a summary) while we introduce an algebraic relation to directly approximate the shear stress in terms of the neutrino density and velocity fields.

It is important to point out that the approximation in Eq. (4.19) is heuristic in the sense that it is only loosely motivated and we do not expect it to accurately reproduce the anisotropic stress. However, the approximation in Eq. (4.18) is much more robust and plays a central role in the fluid equations, while Eq. (4.19) at least qualitatively accounts for the subtle effects of shear stress on the scales where they are needed. In other words, the precise modeling of the shear stress is not important as long as one is only interested in the neutrino density and velocity fields in the nonrelativistic regime. As we will see in the next section, the fluid approximation benefits from a dramatic increase in accuracy when the dispersive nature of the neutrino fluid is accounted for.

The approximations in Eqs. (4.18) and (4.19) are tailored to the Newtonian gauge. An extension to alternative gauges is presented in Appendix 4.7, where we show explicit formulas in the synchronous gauge.

4.4 Numerical results

We consider the standard fluid equations for the evolution of cosmological perturbations in massive neutrinos species, i.e., Eqs. (4.7) and (4.8), but now with novel scale dependent approximate expressions for the sound speed and shear stress, Eqs. (4.18) and (4.19) respectively. We are left with a simple closed system of two ordinary differential equations that we refer to as the modified fluid approximation (Modified FA), and solve numerically.

We extract the Hubble expansion rate and gravitational potentials directly from CLASS

so we can just focus on the neutrino species [6]. We consider three distinct values for the individual neutrino mass, $m_\nu = 0.02$eV, $m_\nu = 0.1$eV and $m_\nu = 0.5$eV, and solve the evolution equations for 17 wavenumbers ranging from $k_{\rm min} = 10^{-3}$ Mpc^{-1} to $k_{\rm max} = 10$ Mpc^{-1}, equally spaced logarithmically. We set the standard adiabatic initial conditions at super horizon scales, extracting the initial values of neutrino transfer functions directly from CLASS. We similarly evolve the neutrino transfer functions with the CLASS fluid approximation for comparison [127].

In Fig.4.3 we plot the neutrino density contrast as a function of the scale factor for varying neutrino mass and scale. As one can see from the plot, the Modified FA produces the late-time neutrino growth at intermediate and small scales much more accurately than the CLASS FA. Also, the errors in both fluid approximations can be large in the relativistic regime where the effects of shear stress are significant. In Fig.4.4 we plot the relative differences in the neutrino density contrast, comparing the exact solution to the fluid approximations, as a function of redshift for $z \leq 30$. At large scales the two fluid approximations have a similar performance, but errors are much smaller in the Modified FA when compared to the CLASS FA at intermediate and small scales. Note that, for the Modified FA, the errors are always below a $\lesssim 30\%$ for redshift $z \lesssim 30$ and $\lesssim 5\%$ for redshift $z \lesssim 5$. On the other hand, the CLASS FA can be as much as a factor of two wrong even at $z = 0$. We also point out that the fluid approximations are more accurate for larger neutrino masses, which is to be expected since a larger mass implies neutrinos are deeper in the nonrelativistic regime where the contributions from the shear stress can be neglected.

In Fig.4.5 we plot the divergence of the neutrino velocity as a function of the scale factor for varying neutrino mass and scale, and in Fig.4.6 we plot the relative differences in the divergence of the neutrino velocity, comparing the exact solution to the fluid approximations, as a function of redshift for $z \leq 30$. The divergence of the neutrino velocity displays fast

[6] In a Boltzmann solver the neutrino species is coupled to all the other species in the universe via the Einstein equations, and so there are additional evolution equations for the Hubble expansion rate (the Friedmann equation) and the gravitational potentials.

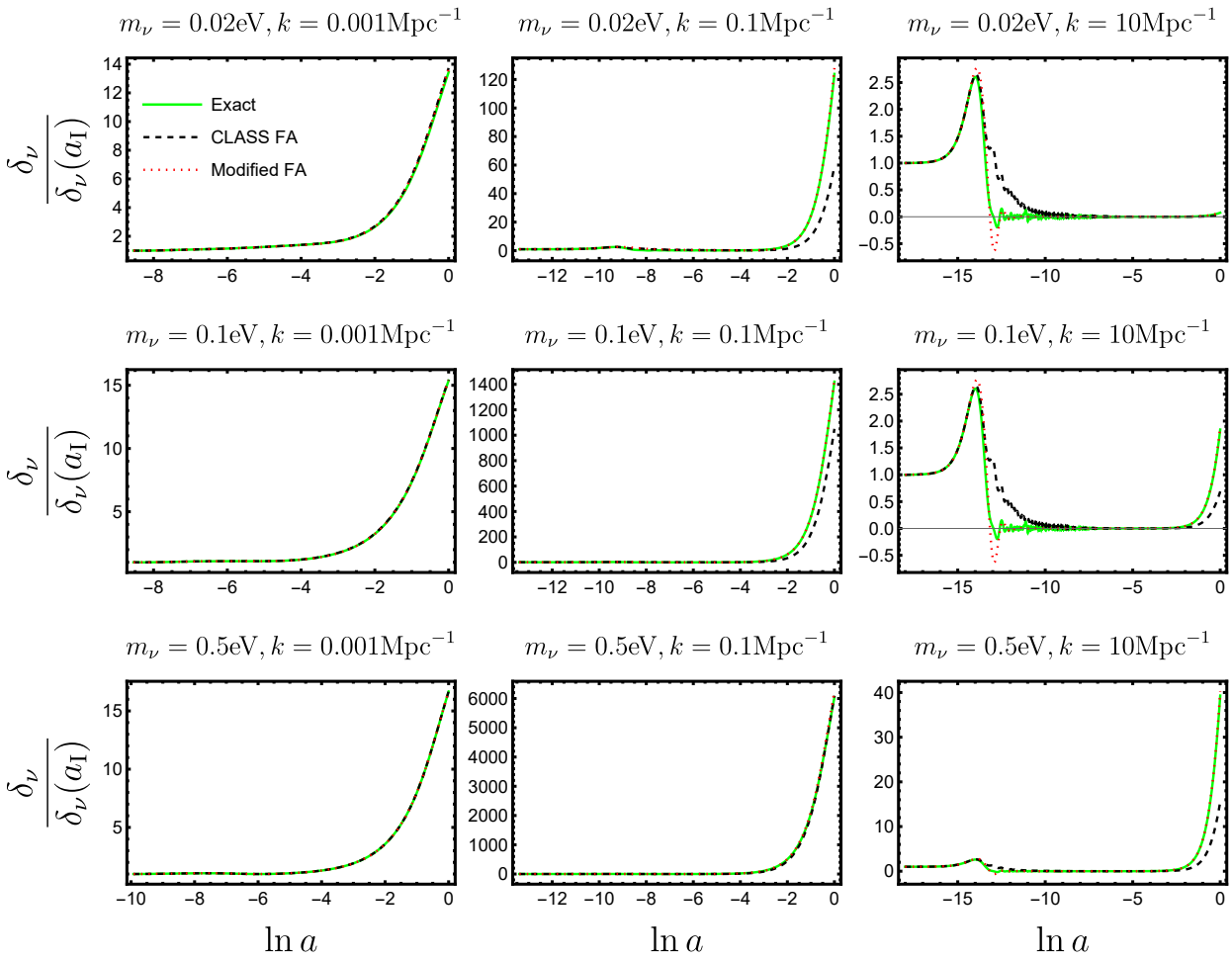

Figure 4.3: Neutrino density contrast (normalized by its initial value $\delta_\nu(a_\text{I})$, and in the Newtonian gauge) as a function of the scale factor and for varying neutrino mass and scale. The plot compares the modified and CLASS fluid approximations (FA) with the exact solution from CLASS in high precision settings. The solid green curves corresponds to CLASS in high precision settings, black dashed curves to the CLASS FA, and red dotted curves to the Modified FA.

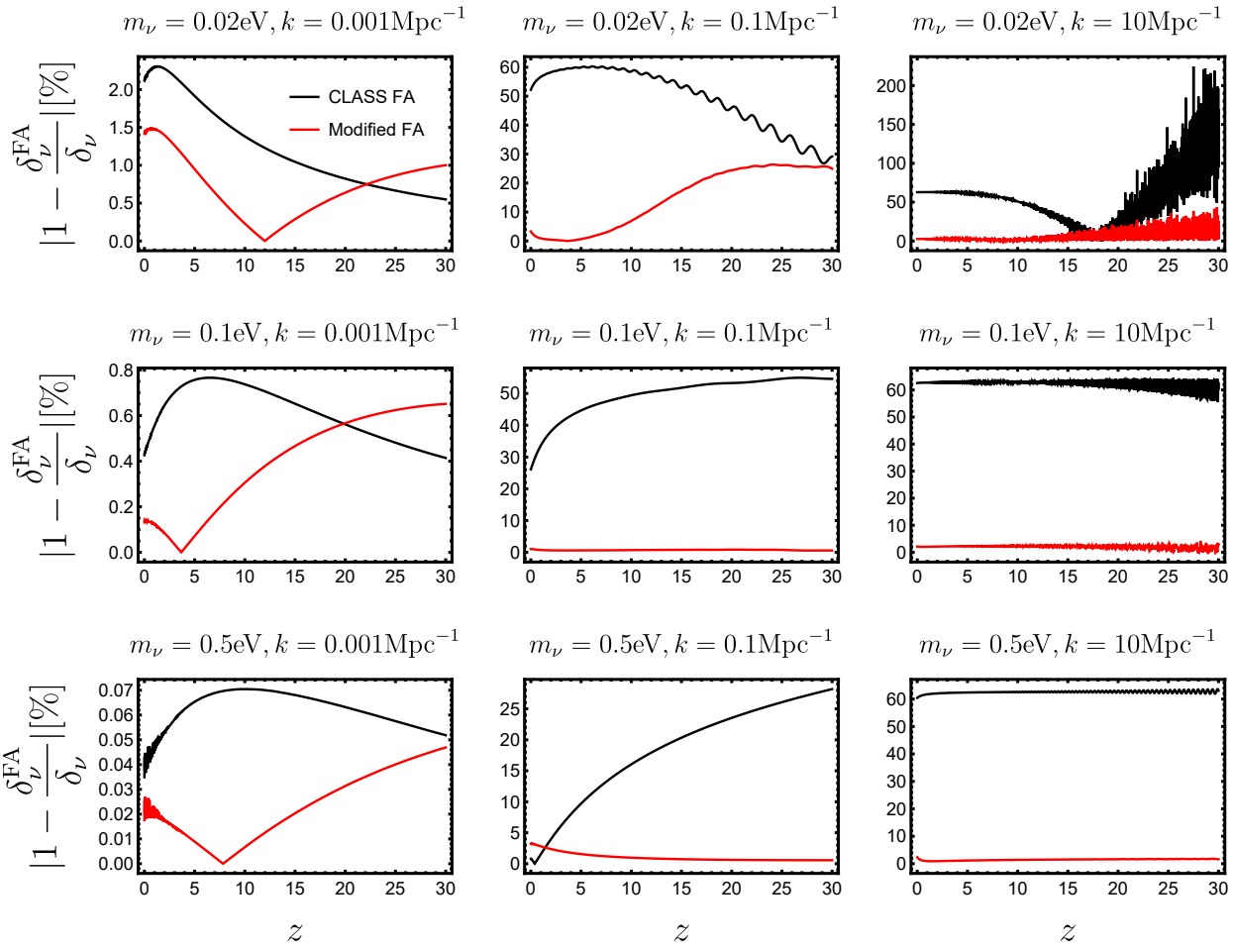

Figure 4.4: Per cent relative differences in the neutrino density contrast between the exact solution and fluid approximations (in the Newtonian gauge) as a function of redshift (with $z \leq 30$) for varying neutrino mass and scale. The plot compares the modified and CLASS fluid approximations (FA) with the exact solution from CLASS in high precision settings. The solid black curves corresponds to the CLASS FA and red solid curves to the Modified FA.

oscillations around zero at the smallest scale and at late times, as can be seen in Fig.4.5, in which case small phase shifts lead to large relative differences that are insignificant. This is why we choose not to include the $k = 10\text{Mpc}^{-1}$ plots in Fig.4.6. Once again the Modified FA is significantly more accurate than the CLASS FA at intermediate and small scales, and overall the errors in the divergence of the neutrino velocity are of the same size as the errors in the neutrino density contrast.

Finally, in Fig.4.7 we plot the neutrino density contrast as a function of scale for varying neutrino mass and for two values of redshift, $z = 0$ and $z = 3$. The neutrino density contrast from the Modified FA is in very good agreement with the exact solution from CLASS in high precision settings, with a $\lesssim 5\%$ agreement at redshifts $z \lesssim 5$ at all scales. We expect linear perturbation theory to break down at sufficiently small scales, $k \gtrsim k_{\rm NL}$, with $k_{\rm NL}$ the scale of nonlinearities, [7] which is included as a vertical dashed line in Fig.4.7. One can then see from the plot that, and specially for smaller neutrino masses, the Modified FA leads to significant improvements when compared to previous approximations schemes at scales where the linear perturbation theory can be safely applied.

Since in high precision settings CLASS and CAMB agree to a percent level [147], we can also conclude that the Modified FA agrees with the Boltzmann solver CAMB as well. Additionally, in Appendix 4.7 we argue that the fluid approximation is as accurate in the synchronous gauge as it is in the Newtonian gauge. Furthermore, from Fig.4.7 the Modified FA is visibly superior to both the CLASS FA and CLASS in default precision settings, specially at scales comparable to, and smaller than, the neutrino free-streaming scale [8]. The Modified FA is then a simple system of two ordinary differential equations that can accurately

[7] We adopt the definition for the scale of nonlinearities based on the root-mean-square linear theory displacement, i.e., $k_{\rm NL}^{-2} = f^2 \int (dk/2\pi^2) P(k)$, with $f = d\log D_{\rm L}/d\log a$ the liner growth rate [in terms of the linear growth factor $D_{\rm L}(a)$] and $P(k)$ the matter power spectrum. This leads to the numerical values of $k_{\rm NL} \approx 0.12\text{Mpc}^{-1}$ at $z = 0$, and $k_{\rm NL} \approx 0.21\text{Mpc}^{-1}$ at $z = 3$, in our reference Planck 2018 cosmology [11]. An alternative definition is given by the scale where the dimensionless power spectrum becomes unity, which leads to similar numerical values.

[8] In Fig.4.7, the fact that the CLASS FA aligns with CLASS in default precision settings is not a coincidence, since in default precision settings CLASS switches from the Boltzmann hierarchy to the CLASS FA after horizon crossing (at $k\tau = 15$).

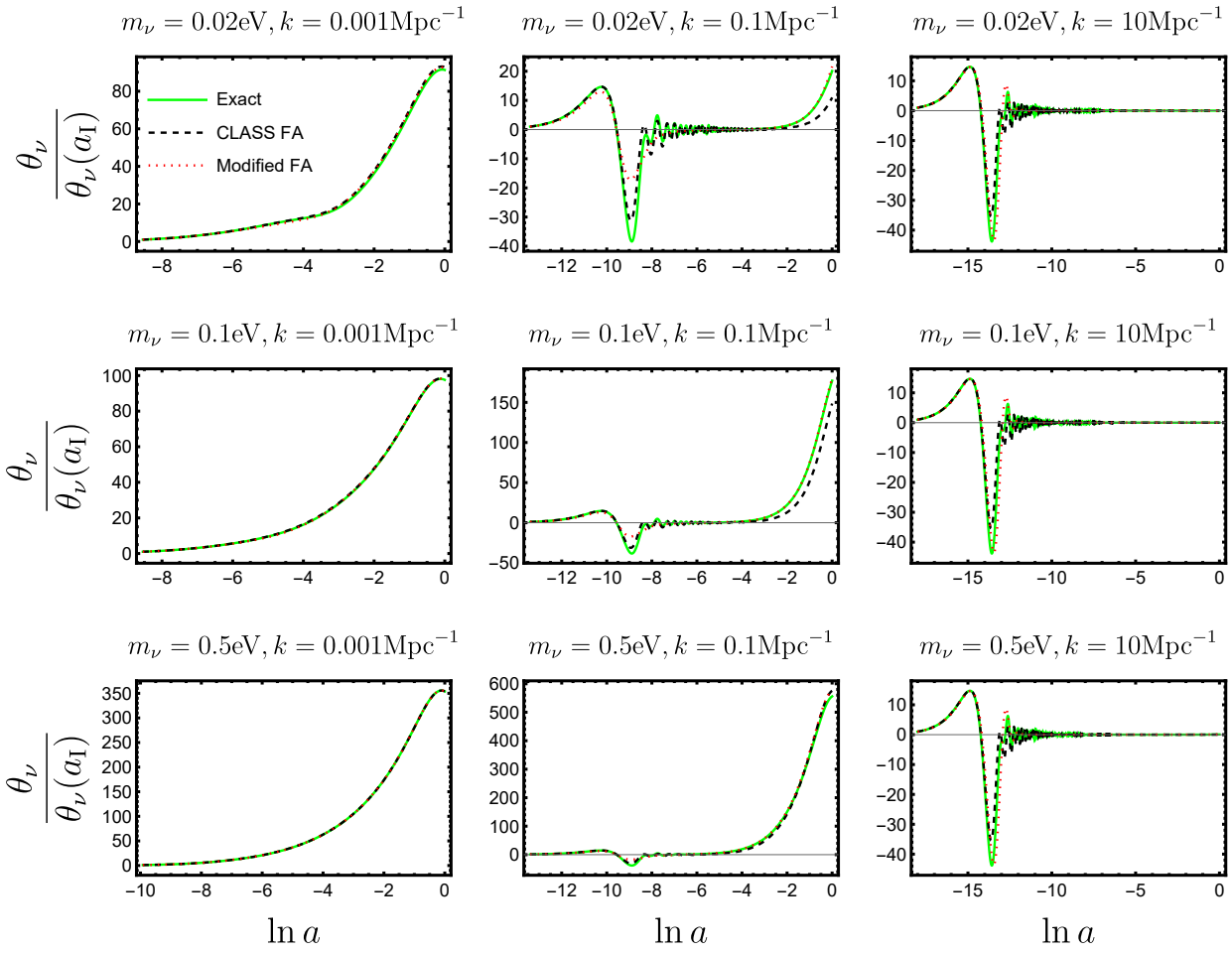

Figure 4.5: Divergence of the neutrino velocity (normalized by its initial value $\theta_\nu(a_\text{I})$, and in the Newtonian gauge) as a function of the scale factor and for varying neutrino mass and scale. The plot compares the modified and CLASS fluid approximations (FA) with the exact solution from CLASS in high precision settings. The solid green curves corresponds to CLASS in high precision settings, black dashed curves to the CLASS FA, and red dotted curves to the Modified FA.

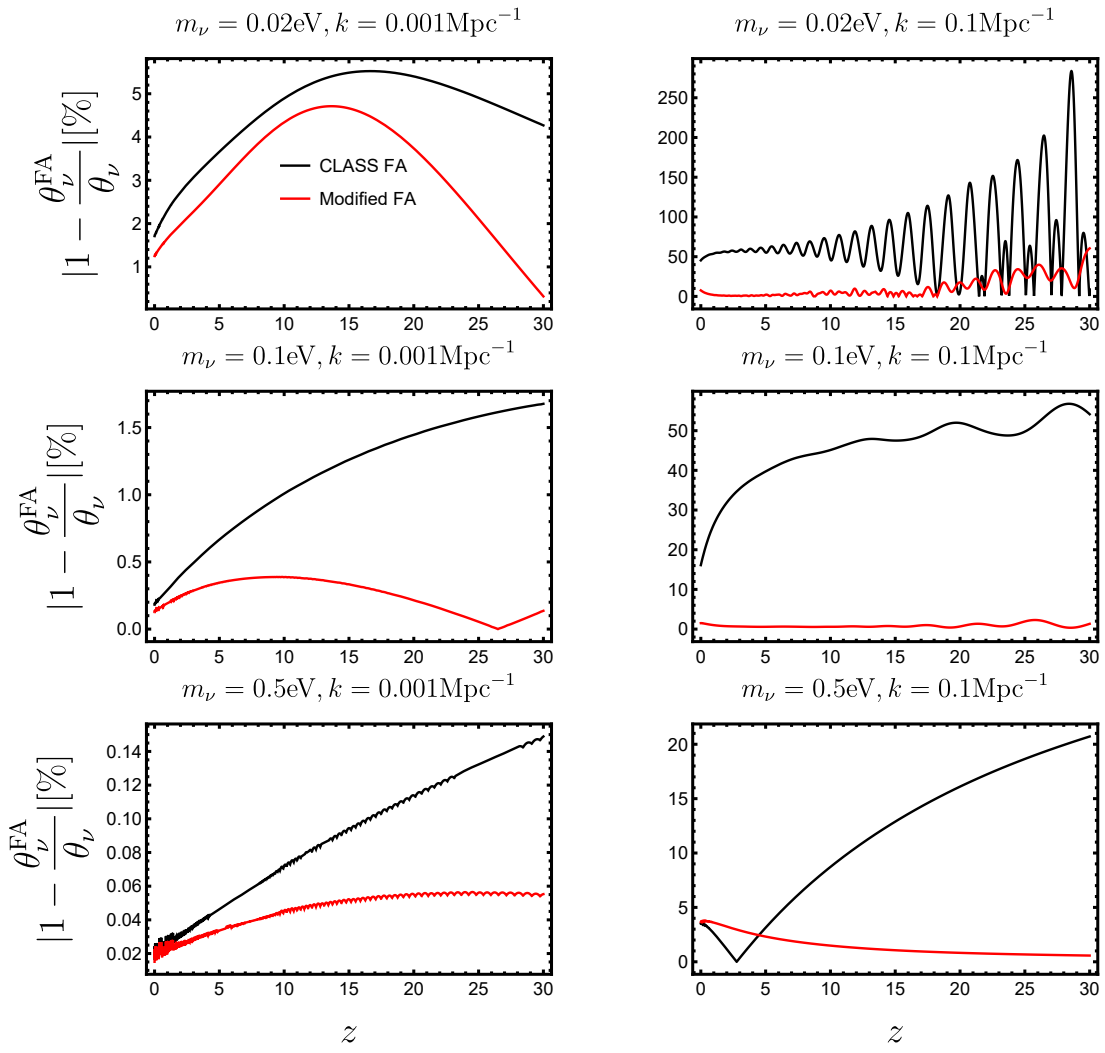

Figure 4.6: Per cent relative differences in the divergence of the neutrino velocity between the exact solution and fluid approximations (in the Newtonian gauge) as a function of redshift (with $z \leq 30$) for varying neutrino mass and scale. The plot compares the modified and CLASS fluid approximations (FA) with the exact solution from CLASS in high precision settings. The solid black curves corresponds to the CLASS FA and red solid curves to the Modified FA.

predict the evolution of linear cosmological neutrino anisotropies at late times.

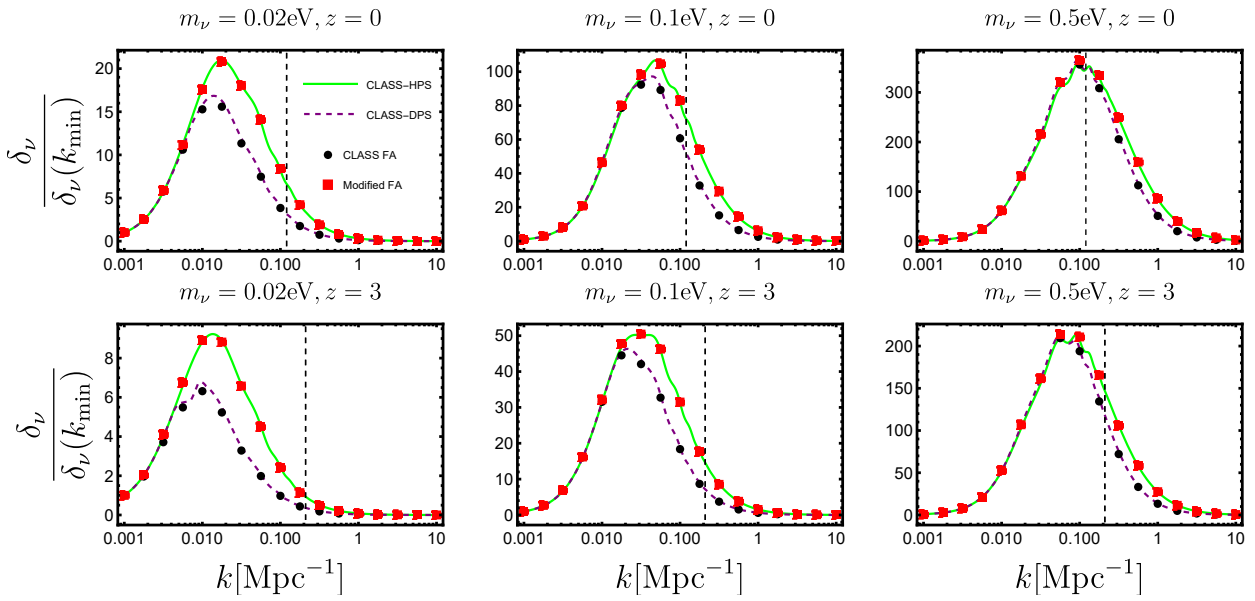

Figure 4.7: Neutrino density contrast (normalized by its value at the largest scale we consider, $k_{\min} = 10^{-3}\text{Mpc}^{-1}$, and in the Newtonian gauge) as a function of scale and for varying neutrino mass and redshift. The plot compares the modified and CLASS fluid approximations (FA) with the solution from CLASS in both high and default precision settings. The green solid curves corresponds to CLASS in high precision settings, and the dashed purple curves to CLASS in default precision settings. The black round data points are obtained with the CLASS FA and red square data points are obtained with the Modified FA. The black dashed vertical line sits at the scale of nonlinearities $k_{\text{NL}} \approx 0.12\text{Mpc}^{-1}$, beyond which the linear perturbation theory is expected to break down.

4.5 Conclusion

We revisited the fluid approximation for massive neutrinos in linear cosmological perturbation theory, but now accounting for the dispersive nature of the neutrino fluid. In Sec.4.3 we introduced the analytic expressions in Eqs. (4.9) and (4.11) for the large and small scales

limits of the sound speed respectively (see Fig.4.2), leading to the novel approximations in Eqs. (4.18) and (4.19) that can be used to close the fluid Eqs. (4.7) and (4.8).

In Sec. 4.4 we showed that the resulting modified fluid approximation produces a neutrino transfer function that is in very good agreement with the exact solution from CLASS in high precision settings, achieving a $\lesssim 5\%$ errors for redshifts $z \lesssim 5$, and over scales (at least) in the range $k = (10^{-3} - 10)\mathrm{Mpc}^{-1}$. Furthermore, Figs. 4.4, 4.6 and 4.7 show the superiority of the modified fluid approximation when compared to both the Boltzmann solver CLASS in default precision settings and the CLASS fluid approximation, which corresponds to an order of magnitude improvement over previous approximation schemes that can be as much as a factor of two wrong even at $z = 0$.

The modified fluid approximation we propose then offers a simple implementation of massive neutrinos in linear cosmological perturbation theory, being much faster and more versatile than the full Boltzmann hierarchy while delivering accurate neutrino transfer functions at late times. In terms of the sum of neutrino masses $M_\nu = \sum_\nu m_\nu$, the contribution of nonrelativistic neutrinos to the energy density in the Universe is $\Omega_\nu h^2 \approx M_\nu/93.14\mathrm{eV}$ [64]. For $M_\nu \lesssim 1\mathrm{eV}$ and $\Omega_\mathrm{m} h^2 = 0.1424$ [11], this leads to a fractional contribution of neutrinos to the total matter of $f_\nu \lesssim 7\%$. As a consequence, $\lesssim 5\%$ errors in the neutrino clustering lead to a matter power spectrum that is accurate to the sub-percent level. This estimate is conservative as the effect of neutrino masses in the matter power spectrum comes mostly from the absence of neutrino perturbations at small scales, and not from the clustering of neutrinos itself. We conclude that the Modified FA is sufficient for most applications, with the Boltzmann hierarchy being needed only if one is interested in the neutrino anisotropies at redshift $z > 5$, or if the warm dark matter component (neutrinos) has a significant contribution to the total matter. We tested the modified fluid approximation for neutrino masses in the range $m_\nu = (0.02 - 0.5)\mathrm{eV}$, but we expect it to remain accurate outside of this range (provided the neutrino mass is large enough for it to become non-relativistic at sufficiently early times and otherwise can be treated as a radiation component).

In a future work we plan on investigating a reformulation of how massive neutrinos are

implemented in the Boltzmann solver CLASS, based on the generalized Boltzmann hierarchy [79], a novel alternative approach to the exact massive neutrino dynamics that can reach sub-percent level accuracy at large and intermediate scales, coupled to our fluid approximation for the small scale dynamics. We will then perform a thorough comparison with standard methods and approximation schemes, and we expect to achieve significant improvements in both accuracy and computation time. Finally, another possible direction of future research is to investigate potential implications of this work to the clustering of neutrinos in the nonlinear regime on the basis of a fluid approach.

4.6 Appendix: Analytic calculations and approximations

We investigate analytically the evolution of cosmological linear perturbations in a massive neutrino species, with the goal of extracting both the large and small scale limits of the sound speed and the qualitative behavior of the anisotropic stress. This study motivates the approximations in Eqs. (4.18) and (4.19).

We follow [129] and write evolution equations for the phase-space distribution of massive neutrinos. The first step is to split it into a background and perturbation components (working in Fourier space),

$$f(k,q,\mu,\tau) = f_0(q)[1+\Psi(k,q,\mu,\tau)]\,, \tag{4.20}$$

where k the magnitude of the wavevector $\vec{k}$, q the magnitude of the comoving momentum $\vec{q}$, $\mu = \hat{k}\cdot\hat{q}$ is the cosine of the angle between these two vectors, and τ is the conformal time. Furthermore, the background is a relativistic Fermi-Dirac distribution:

$$f_0(q) = \frac{g_\nu}{e^{\frac{q}{T_{\nu,0}}}+1}\,, \tag{4.21}$$

with $g_\nu = 2$ to account for both left-handed neutrinos and right-handed antineutrinos, and $T_{\nu,0} \approx 1.95\text{K} \approx 1.7\times10^{-4}\text{eV}$ is the neutrino temperature today. We remind the reader that we are working in the Newtonian gauge:

$$ds^2 = a^2(\tau)\left[-(1+2\psi)d\tau^2 + (1-2\phi)d\vec{x}^2\right]\,. \tag{4.22}$$

The evolution equation for the perturbation to the phase-space distribution follows from the collisionless Boltzmann equation, and reads [9],

$$\Psi' + i\frac{kq}{\epsilon}\mu\Psi = \frac{d\ln f_0}{d\ln q}\left(i\frac{k\epsilon}{q}\mu\psi - \phi'\right) , \tag{4.23}$$

where $\epsilon(q,\tau) = \sqrt{q^2 + m_\nu^2 a(\tau)^2}$ is the comoving energy. It is convenient to perform a decomposition of $\Psi(k,q,\mu,\tau)$ into Legendre polynomials, $P_l(\mu)$, as follows:

$$\Psi(k,q,\mu,\tau) = \sum_{l=0}^{\infty}(-i)^l(2l+1)\Psi_l(k,q,\tau)P_l(\mu)\,, \tag{4.24}$$

where the $\Psi_l(k,q,\tau)$ are called the multipoles of the distribution function. The substitution of Eq. (4.24) into Eq. (4.23) yields, after using Legendre polynomial identities,

$$\Psi_0' = -\frac{q}{\epsilon}\Psi_1 - \frac{d\ln f_0}{d\ln q}\phi'\,, \tag{4.25}$$

$$\Psi_1' = \frac{q}{3\epsilon}(\Psi_0 - 2\Psi_2) - \frac{\epsilon}{3q}\frac{d\ln f_0}{d\ln q}\psi\,, \tag{4.26}$$

$$\Psi_l' = \frac{q}{(2l+1)\epsilon}[l\Psi_{l-1} - (l+1)\Psi_{l+1}] \;\forall l \geq 2\,. \tag{4.27}$$

Once the solution to this set of equations is obtained, one can integrate over momenta to get the neutrino stress-energy tensor, and hence the fluid properties involved in the fluid equations (see Eqs. (4.7) and (4.8)), as follows [10],

$$\delta\rho(k,\tau) = a(\tau)^{-4}\int_0^\infty \frac{dq}{2\pi^2}q^2\epsilon(q,\tau)f_0(q)\Psi_0(k,q,\tau)\;, \tag{4.28}$$

$$\delta P(k,\tau) = \frac{1}{3}a(\tau)^{-4}\int_0^\infty \frac{dq}{2\pi^2}q^2\epsilon(q,\tau)\left[\frac{q}{\epsilon(q,\tau)}\right]^2 f_0(q)\Psi_0(k,q,\tau)\;, \tag{4.29}$$

$$\rho(1+w)\theta(k,\tau) = ka(\tau)^{-4}\int_0^\infty \frac{dq}{2\pi^2}q^3 f_0(q)\Psi_1(k,q,\tau)\;, \tag{4.30}$$

$$\rho(1+w)\sigma(k,\tau) = \frac{2}{3}a(\tau)^{-4}\int_0^\infty \frac{dq}{2\pi^2}q^2\epsilon(q,\tau)\left[\frac{q}{\epsilon(q,\tau)}\right]^2 f_0(q)\Psi_2(k,q,\tau)\;. \tag{4.31}$$

---

[9] From this point forward we omit scale, momentum and time dependences when it is convenient to do so. Also, prime denotes a derivative with respect to conformal time.

[10] We remind the reader that we do not employ a ν subscript when referring to neutrino fluid properties.

Here $\rho = \rho(a)$ is the neutrino background energy density as given by Eq. (4.3), and $w = w(a)$ is the equation of state parameter from Eq. (4.5). Furthermore, $\delta\rho(k,\tau)$ is the perturbation to the neutrino energy density, $\delta P(k,\tau)$ the perturbation to the pressure, $\theta(k,\tau)$ is the divergence of the velocity field, and $\sigma(k,\tau)$ is the anisotropic (shear) stress. This sets the stage for an analytic investigation of the sound speed and anisotropic stress.

4.6.1 The sound speed

The sound speed is defined by:

$$c_{\rm s}^2(k,\tau) = \frac{\delta P(k,\tau)}{\delta\rho(k,\tau)}\,, \tag{4.32}$$

and it is both time and scale dependent. From Eqs. (4.28) and (4.29), one finds that it can be determined from the momentum dependence in the zeroth multipole Ψ_0, so this is what we investigate next.

From the structure of Eq. (4.23) it is convenient to introduce a momentum dependent time variable [11],

$$dz = \frac{q}{\epsilon}d\tau \implies z(\tau,q) = \int_0^\tau \frac{q}{\epsilon(q,\tau')}d\tau'\,. \tag{4.33}$$

Since $v \equiv q/\epsilon$ is the peculiar velocity, this is just the distance traveled by the neutrino particles over the entire expansion history. When evaluated at around the peak of the Fermi-Dirac distribution, $q = 3T_{\nu,0}$, we denote it by $\bar{z} = z(q = 3T_{\nu,0})$. This is analogous to the neutrino horizon scale introduced in Eq. (4.17). We further define $y = kz$, in terms of which the Eq. (4.23) reads,

$$\frac{\partial\Psi}{\partial y} + i\mu\Psi = \frac{d\ln f_0}{d\ln q}\left(ik\frac{\epsilon}{q}\mu\psi - \phi'\right)\,, \tag{4.34}$$

is a first-order ordinary differential equation, whose solution is straightforward to write down:

$$\Psi = \frac{d\ln f_0}{d\ln q}\int_0^y dy'\frac{\hat{\epsilon}}{kq}\left(ik\frac{\hat{\epsilon}}{q}\mu\hat{\psi} - \hat{\phi}'\right)e^{-i\mu(y-y')}\,, \tag{4.35}$$

[11] Not to be confused with redshift. Everywhere in the Appendix 4.6, the variable z will stand for the distance traveled by neutrino particles over the expansion history as given by Eq. (4.33).

where we use the hat notation (exemplified by $\hat{\phi}$) to denote time-dependent quantities evaluated at the intermediate time $y' = kz'$ begin integrated. Additionally, we drop the initial condition terms and send the initial time to zero, i.e., $y_i \to 0$, for simplicity. In the regime $y \gg 1$ this is inconsequential as the solution is dominated by the source term and becomes insensitive to the initial conditions. However, this is no longer true in the regime $y \ll 1$ which corresponds to the largest scales. This does not pose a significant problem for us since we expect the adiabatic sound speed to be recovered on the largest scales (and we will see why that is). A much more important issue is to accurately obtain the small scale limit $y \gg 1$ of the sound speed for which we can safely drop the initial values of the multipoles.

From the multipole expansion in Eq. (4.24) and the orthogonality relations of Legendre polynomlias, we can write:

$$\Psi_l = \frac{1}{2(-i)^l} \int_{-1}^{1} d\mu \; \Psi P_l(\mu) \,. \tag{4.36}$$

This can be evaluated using Eq. (4.35), the identity

$$e^{-i\mu(y-y')} = \sum_{l=0}^{\infty} (-i)^l (2l+1) j_l(y-y') P_l(\mu) \,, \tag{4.37}$$

where $j_l(y)$ are the spherical Bessel functions of the first kind, and the orthogonality relations of Legendre polynomials, to yield:

$$\Psi_l = \frac{d\ln f_0}{d\ln q} \int_0^y dy' \left\{ \left(\frac{\hat{\epsilon}}{q}\right)^2 \hat{\psi} \left[\frac{l+1}{2l+1} j_{l+1}(y-y') - \frac{l}{2l+1} j_{l-1}(y-y') \right] - \frac{\hat{\epsilon}}{kq} \hat{\phi}' j_l(y-y') \right\} . \tag{4.38}$$

We now use the recurrence relation

$$l f_{l-1}(y) - (l+1) f_{l+1}(y) = (2l+1) f_l'(y) \,, \tag{4.39}$$

and integrate the first term in the right-hand side of Eq. (4.38) by parts to arrive at:

$$\Psi_l = \frac{d\ln f_0}{d\ln q} \left\{ j_l(0) \left(\frac{\epsilon}{q}\right)^2 \psi - j_l(y)\psi_i - \int_0^y dy' \frac{\hat{\epsilon}}{kq} \left[\left(\frac{\hat{\epsilon}}{q}\right)^2 \hat{\psi} + \hat{\phi} \right]' j_l(y-y') \right\} . \tag{4.40}$$

Now the crucial observation is that this greatly simplifies in the limit $y \gg 1$ for $l = 0$, since then $j_0(y-y')$ peaks at $\Delta y = y - y' = 0$, and goes to zero in the limit $\Delta y = k(z-z') = k\Delta z = y\Delta z/z \gg 1$, and hence $\Delta z/z \gg 1/y$ is a small number. As a consequence, simply evaluating the integrand at the final time should give a very good approximation and indeed eventually approach the exact solution in the asymptotic regime $x \to \infty$. Indeed, this has been previously investigated and exploited to generate a fluid approximation, in the special case of the massless limit, i.e., for a radiation component [47]. Using $j_0(0) = 1$ and $j_0(y) \approx 0$ for $y \gg 1$ leads to,

$$\begin{aligned}
\Psi_0 &\approx \frac{d\ln f_0}{d\ln q}\left\{\left(\frac{\epsilon}{q}\right)^2\psi - \frac{\epsilon}{kq}\left[\left(\frac{\epsilon}{q}\right)^2\psi + \phi\right]'\int_0^y dy' j_0(y-y')\right\} \\
&\approx \frac{d\ln f_0}{d\ln q}\left\{\left(\frac{\epsilon}{q}\right)^2\psi + \frac{\pi}{2}\frac{\epsilon}{kq}\left[\left(\frac{\epsilon}{q}\right)^2\psi + \phi\right]'\right\} \\
&\approx \frac{d\ln f_0}{d\ln q}\left(\frac{\epsilon}{q}\right)^2\psi\,,
\end{aligned} \tag{4.41}$$

where we used,

$$\int_0^y dy' j_0(y-y') \approx -\int_0^\infty dy j_0(y) = -\frac{\pi}{2}\,, \tag{4.42}$$

and from the second to the third line in Eq. (4.41) we dropped a term that scales like $1/k$ and hence becomes negligible in the limit $y \to \infty$. Indeed, that term is the next to leading order correction to the asymptotic formula.

The final formula in the third line of Eq. (4.41) has a straightforward interpretation. Neutrino fluctuations with (comoving) wavenumber k have a characteristic time scale, $t \sim a/k\sigma_\nu \sim (k_{\rm fs}/k)H^{-1}$, which is much smaller than a Hubble time $\sim 1/H$, at sub-free streaming scales ($k \gg k_{\rm fs}$). We can then consider a static limit of the Boltzmann equation where the expansion of the universe can be taken as slow. In fact, one can check that simply dropping the derivative terms in Eq. (4.23) immediately leads to our final formula in the third line of Eq. (4.41).

It is now straightforward to obtain the asymptotic formula for the sound speed, using Eqs. (4.28),(4.29) and (4.32). We first explicitly evaluate the perturbation to the energy

density, using Eq. (4.41),

$$\begin{aligned}\delta\rho &\approx a^{-4}\psi\int_0^\infty \frac{dq}{2\pi^2}q^3\epsilon(q,a)\left[\frac{\epsilon(q,a)}{q}\right]^2\frac{df_0(q)}{dq}\\ &= -3\rho(a)[1+\lambda(a)]\psi \implies \delta \approx -3[1+\lambda(a)]\psi\,,\end{aligned} \tag{4.43}$$

where from the first to the second line we integrate by parts and introduce the quantity $\lambda(a)$ as:

$$\rho(a)\lambda(a) = \frac{1}{3}a^{-4}\int_0^\infty \frac{dq}{2\pi^2}q^2\epsilon(q,a)\left[\frac{\epsilon(q,a)}{q}\right]^2 f_0(q)\,. \tag{4.44}$$

We can obtain the asymptotic expression for the neutrino pressure in a way that is completely analogous to the density and leads to a sound speed of, using the Eq. (4.32),

$$c^2_{\rm asp}(a) = \frac{1}{3}\frac{1+w(a)}{1+\lambda(a)}\,. \tag{4.45}$$

Before moving on to the opposite large scale regime of the sound speed, it is instructive to investigate the non-relativistic limit of Eq. (4.43) where $\epsilon \approx ma$ and $\lambda \gg 1$, such that,

$$\delta \approx -3\lambda(a)\psi \approx -\frac{1}{\rho}\frac{\ln 4}{2\pi^2}\frac{m_\nu^3 T_{\nu,0}}{a}\psi\,, \tag{4.46}$$

after using the Eq. (4.21). In a universe with matter background energy density $\rho_{\rm m}(a)$ and matter density contrast $\delta_{\rm m}(a)$, the Poisson equation reads:

$$k^2\psi = -4\pi G a^2\rho_{\rm m}\delta_{\rm m}\,, \tag{4.47}$$

where G is Newton's gravitational constant. Using this, the Friedmann equation,

$$\mathcal{H}^2 = \frac{8\pi G}{3}a^2\frac{\rho_{\rm m}}{\Omega_{\rm m}}\,, \tag{4.48}$$

with $\Omega_{\rm m}(a)$ the fractional contribution of matter to the total energy density and also evaluating explicitly the Eq. (4.3) for $\rho(a)$ in the non-relativistic regime, we can write the Eq. (4.46) as:

$$\delta = \left(\frac{k_{\rm fs}}{k}\right)^2\delta_{\rm m}\,, \tag{4.49}$$

where,

$$k_{\rm fs}(a) = \sqrt{\frac{3}{2}\Omega_{\rm m}(a)}\frac{\mathcal{H}(a)}{\sigma_\nu(a)}\,, \tag{4.50}$$

is (proportional to) the distance that neutrinos travel over the course of one expansion time $t \sim 1/\mathcal{H}$, i.e., $\lambda_{\rm fs} \sim 1/k_{\rm fs} \sim \sigma_\nu t \sim \sigma_\nu/\mathcal{H}$, with,

$$\sigma_\nu(a) = \sqrt{\frac{3\zeta(3)}{\ln 4}}\frac{T_{\nu,0}}{m_\nu a}\,, \tag{4.51}$$

the expression for the neutrino velocity dispersion, chosen in such a way as to produce the Eq. (4.49) with no additional coefficients. Indeed, the small scale result in Eq. (4.49) is well-known in the literature [163] and the next to leading order correction in the second line of Eq. (4.41) leads to a contribution that scales as $\sim (k_{\rm fs}/k)^3$.

We are now ready to obtain an expression for the sound speed in the large scale regime $k \ll k_{\rm fs}$. For this we go back to the Eq. (4.40) in the case of $l = 0$, but now assume $\Delta y \leq y \ll 1$ such that $j_0(\Delta y) \approx 1$ and we obtain,

$$\int_0^y dy' \frac{\hat{\epsilon}}{kq}\left[\left(\frac{\hat{\epsilon}}{q}\right)^2\hat{\psi} + \hat{\phi}\right]' j_0(y-y') \approx \left(\frac{\epsilon}{q}\right)^2\psi + \phi - \psi_i - \phi_i\,, \tag{4.52}$$

upon using $d\tau' = dy'(\hat{\epsilon}/kq)$ and assuming the neutrinos to be relativistic at the initial time. This yields:

$$\Psi_0 \approx (\phi_i - \phi)\frac{d\ln f_0}{d\ln q}\,. \tag{4.53}$$

As we argued previously, we do not necessarily expect the Eq. (4.53) to give an accurate approximation to the zeroth multipole on the large scales, since we dropped the initial values of the multipoles that play a role in this regime. However, the statement that the momentum dependence of Ψ_0 is set by the log derivative of the background distribution function (the so-called separable ansatz), still holds true at sufficiently large scales as it follows from separate universe arguments (see [164, 165]). We may now proceed to obtain an expression for the sound speed as before, using Eqs. (4.28), (4.29), (4.32) and (4.53). We get the so-called adiabatic sound speed,

$$c_{\rm g}^2(a) = \frac{P'(a)}{\rho'(a)}\,, \tag{4.54}$$

after using the Eqs. (4.3) and (4.4) for the neutrino energy density $\rho(a)$, and pressure $P(a)$, respectively. Also, the derivatives with respect to conformal time can be obtained via the relation $\epsilon' = \mathcal{H}\epsilon[1-(q/\epsilon)^2]$.

Now equipped with both $k \gg k_{\rm fs}$ and $k \ll k_{\rm fs}$ limits of the sound speed in the Newtonian gauge, as given by Eqs. (4.45) and (4.54) respectively, we see that an interpolation such as the one provided by the Eq. (4.18) should yield a good approximation to the scale-dependent sound speed.

4.6.2 Anisotropic stress

After obtaining an expression for the sound speed in both the $k/k_{\rm fs} \gg 1$ and $k/k_{\rm fs} \ll 1$ regimes, we would like to develop a qualitative understanding on the behavior of the anisotropic stress following the discussion in [129].

Combining the Eqs. (4.25) and (4.27) for $l = 2$, we obtain,

$$\left(\Psi_2 + \frac{2}{5}\Psi_0 + \frac{2}{5}\phi\frac{d\ln f_0}{d\ln q}\right)' = -\frac{3}{5}\frac{q}{\epsilon}\Psi_3\,. \tag{4.55}$$

Since our goal is to build intuition about the anisotropic stress we can make some simplifying approximations. The first is to set $\phi' \approx 0$, which is always a good approximation on small scales, and exact in a matter dominated universe. The second will be to drop the initial values of the multipoles, which is the same approximation we applied in the last subsection. This leads to,

$$\Psi_2 \approx -\frac{2}{5}\Psi_0 - \frac{3}{5}\int_0^y dy'\frac{\hat{\Psi}_3}{k}\ , \tag{4.56}$$

where we used the Eq. (4.33), $y = kz$, and the hat notation to denote time-dependent quantities evaluated at the intermediate time y' to be integrated. Using the Eqs. (4.29) and (4.31), we conclude that the first term in the right-hand side of Eq. (4.56) generates the following contribution to the anisotropic stress,

$$\sigma \supset -\frac{4}{5}\frac{c_{\rm s}^2}{1+w}\delta\ , \tag{4.57}$$

which in the fluid equations looks like an additional contribution to the sound speed. However, we expect that setting $\Psi_3 = 0$, which is equivalent to dropping the second term in the right-hand side of Eq. (4.56), is only a reasonable approximation when $y = kz \lesssim 3$ such that the shear acts like a contribution to the sound speed at scales that are around the horizon

$k \sim 1/\bar{z} \sim k_{\rm hor}$. In the regime $y \gg 1$ we expect the second term in the right-hand side of Eq. (4.56) to dominate, generally giving a viscosity-type contribution to the shear stress at scales that are comparable to, or smaller than, the free-streaming scale ($k \gtrsim k_{\rm fs}$). These considerations motivate the heuristic approximate expression in Eq. (4.19).

4.7 Appendix: Alternative gauges

In the main text we work solely in the conformal Newtonian gauge for simplicity. In this section we extend our fluid approximation to a general alternative gauge under the example of the synchronous gauge due to its usage in Boltzmann solvers, for concreteness. We follow [19] where more details can be found.

Small scalar perturbations to the FLRW universe are given by, in the synchronous gauge:

$$ds^2 = a^2(\tau)[-d\tau^2 + (\delta_{ij} + h_{ij})dx^i dx^j]\,, \tag{4.58}$$

where δ_{ij} is the Kronecker symbol, and $h_{ij}(\tau, \vec{x})$ can be decomposed (in Fourier space), in terms of two metric perturbations $h(\tau, \vec{k})$ and $\eta(\tau, \vec{k})$, as follows:

$$h_{ij}(\tau, \vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \left[\hat{k}_i\hat{k}_j h(\tau, \vec{k}) + (\hat{k}_i\hat{k}_j - \frac{1}{3}\delta_{ij})6\eta(\tau, \vec{k})\right]\,, \tag{4.59}$$

with $\hat{k}_i = \vec{k}_i/k_i$ normalized to unit length. In other gauges the metric will similarly be written in terms of two other scalar metric perturbations (such as ψ and ϕ in the Newtonian gauge).

In the synchronous gauge the fluid equations that follow from stress-energy conservation read,

$$\tilde{\delta}' = -(1+w)\left(\tilde{\theta} + \frac{h'}{2}\right) - 3\mathcal{H}(\tilde{c}_s^2 - w)\tilde{\delta}\,, \tag{4.60}$$

$$\tilde{\theta}' = -\mathcal{H}(1-3w)\tilde{\theta} - \frac{w'}{1+w}\tilde{\theta} + \frac{\tilde{c}_s^2}{1+w}k^2\tilde{\delta} - k^2\tilde{\sigma}\,. \tag{4.61}$$

Here $\tilde{\delta}$ is the neutrino density contrast, $\tilde{\theta}$ is the divergence of the velocity field, $\tilde{\sigma}$ is the anisotropic stress, and the pressure perturbation term $\delta\tilde{P}$ is parameterized by the sound

speed $\tilde{c}_s^2$ as in the main text,

$$\frac{\delta\tilde{P}}{\rho} = \frac{\delta\tilde{P}}{\delta\tilde{\rho}}\tilde{\delta} = \tilde{c}_s^2\tilde{\delta}\,, \tag{4.62}$$

where from now on tilded quantities are in the alternative gauge (which in our example is the synchronous gauge) and untilded quantities are in the Newtonian gauge. These are the neutrino fluid properties that are involved in the fluid equations. Also $\rho(a)$ and $P(a)$ are the background density and pressure respectively, and $w = P/\rho$ is the equation of state.

Starting from Newtonian gauge coordinates x^μ, we can apply a gauge transformation $x^\mu \to \tilde{x}^\mu = x^\mu + d^\mu(x^\nu)$ in order to arrive at an arbitrary gauge. For scalar perturbations, the d^μ can be decomposed into time and spatial components as,

$$d^0(x^0, \vec{x}) = \alpha(\tau, \vec{x})\ , \tag{4.63}$$

$$\vec{d}(x^0, \vec{x}) = \vec{\nabla}\beta(\tau, \vec{x})\ . \tag{4.64}$$

In the example of the synchronous gauge, these are given by (in Fourier space),

$$\beta = \frac{1}{2k^2}(h + 6\eta)\,, \tag{4.65}$$

$$\alpha = \beta'\,. \tag{4.66}$$

In general, α and β will be given in terms of the two metric perturbations in the alternative gauge. The transformation laws for the fluid properties in the Eqs. (4.60) and (4.61) (in any gauge, here illustrated in the synchronous gauge) follow from the covariant transformation law satisfied by the energy-momentum tensor, and read

$$\tilde{\delta} = \delta - \alpha\frac{\rho'}{\rho}\,, \tag{4.67}$$

$$\tilde{\theta} = \theta - \beta' k^2\ , \tag{4.68}$$

$$\delta\tilde{P} = \delta P - \alpha P'\ , \tag{4.69}$$

$$\tilde{\sigma} = \sigma\,. \tag{4.70}$$

Our approximations in Eqs. (4.18) and (4.19), suitable to the Newtonian gauge, can then be

straightforwardly mapped into an arbitrary alternative gauge as follows:

$$\frac{\delta\tilde{P}}{\rho} = c_{\rm s}^2\tilde{\delta} + \alpha(c_{\rm s}^2 - c_{\rm g}^2)\frac{\rho'}{\rho}\,, \tag{4.71}$$

$$k^2\tilde{\sigma} = -\frac{2}{5}\frac{k_{\rm hor}}{k}e^{-\frac{k_{\rm hor}}{k}}\frac{c_{\rm s}^2}{1+w}k^2\left(\tilde{\delta} + \alpha\frac{\rho'}{\rho}\right) + \frac{k}{k_{\rm fs}}e^{-5\frac{k_{\rm fs}}{k}}w^2\left(\tilde{\theta} + \beta' k^2\right)\,, \tag{4.72}$$

where,

$$c_{\rm s}^2 = c_{\rm g}^2 + (c_{\rm asp}^2 - c_{\rm g}^2)e^{-\frac{4}{3}\frac{k_{\rm fs}}{k}}\,, \tag{4.73}$$

is the approximate expression for the scale-dependent sound speed in the Newtonian gauge. Also, the adiabatic $c_{\rm g}^2$ and asymptotic $c_{\rm asp}^2$ sound speeds are given by Eqs. (4.9) and (4.11) respectively, and the neutrino free-streaming $k_{\rm fs}$ and horizon $k_{\rm hor}$ scales are given by Eqs. (4.14) and (4.17), respectively.

We were not able to directly test the fluid Eqs. (4.60) and (4.61), with the approximations in Eqs. (4.71) and (4.72) in the synchronous gauge, due to the fact that the Boltzmann solver CLASS does not output the synchronous gauge metric perturbations h and η as a function of the scale factor for a given scale k. However, we were able to extract the function α, applying the Eq. (4.67) with both the Newtonian and synchronous gauge exact solutions (obtained from CLASS in high precision settings). This allowed us to use the fluid equations to obtain the approximate solution in the Newtonian gauge and then transform that into the synchronous gauge, which is mathematically equivalent to solving the fluid equations in the synchronous gauge. When comparing to the exact synchronous gauge transfer functions we obtain a level of accuracy which is the same as observed in the Newtonian gauge, hence verifying that the fluid approximation works as well in the synchronous gauge as it does in the Newtonian gauge.

Chapter 5

NEUTRINO WINDS ON THE SKY

5.1 Abstract

We develop a first-principles formalism to compute the distortion to the relic neutrino density field caused by the peculiar motions of large-scale structures. This distortion slows halos down due to dynamical friction, causes a local anisotropy in the neutrino-CDM cross-correlation, and reduces the global cross-correlation between neutrinos and CDM. The local anisotropy in the neutrino-CDM cross-spectrum is imprinted in the three point cross-correlations of matter and galaxies, or the bispectrum in Fourier space, producing a signal peaking at squeezed triangle configurations. This bispectrum signature of neutrino masses is not limited by cosmic variance or potential inaccuracies in the modeling of complicated nonlinear and galaxy formation physics, and it is not degenerate with the optical depth to reionization. We show that future surveys have the potential to detect the distortion bispectrum.

5.2 Introduction

Neutrino oscillation experiments have established that at least two of the neutrino mass eigenstates have a non-zero mass. Two squared mass differences are known from atmospheric and solar neutrino oscillations, which leads to a lower bound on the sum of the masses of $\sum_\nu m_\nu \gtrsim 0.06, 0.1\text{eV}$ for the normal and inverted hierarchies respectively [149, 151, 166]. Additionally, terrestrial beta decay experiments set an upper bound to a weighted sum of the masses of $m_{\nu,\beta} < 0.8\text{eV}$ at 90% confidence level [152]. Tighter constraints are obtained from the neutrino-induced suppression of cosmological structure formation at small scales, leading to the bound $\sum_\nu m_\nu \lesssim 0.12\text{eV}$ at 95% confidence [11], depending on the dataset.

Future cosmological surveys will reach the required sensitivity to detect the neutrino mass scale at the lower limits imposed by oscillation data. This highly anticipated measurement will place tight constraints on sub-percent components of the matter density that can alter the power spectrum, and put the synergy between particle physics and cosmology to the test [159].

Relic neutrinos are nonrelativistic at late times but even today still have a nonnegligible (thermal) velocity dispersion of $\sigma_\nu \sim 3T_{\nu,0}/m_\nu \approx 1500$ km/s for a neutrino mass $m_\nu = 0.1$eV, and temperature $T_{\nu,0} \approx 1.95\text{K} \approx 1.7 \times 10^{-4}$eV. We can define the free-streaming scale as the distance traveled by neutrinos over the course of one expansion time, i.e., $\lambda_{\rm fs} \sim \sigma_\nu/H_0 \sim 20$ Mpc when evaluated at the present time, where $H_0 \approx 70$ km/s/Mpc is the Hubble expansion rate. At scales below the free-streaming length ($k > k_{\rm fs}$) pressure dominates over the gravitational potential and neutrino anisotropies are washed out creating a homogeneous sea of neutrinos with a characteristic size $L \sim \lambda_{\rm fs}$. At the present time cold dark matter (CDM) halos have a typical peculiar velocity $v_{\rm H} \sim 500$ km/s. We will show that the peculiar motion of halos causes neutrino particles to accumulate behind the moving halo, generating wakes that slow the halo down due to dynamical friction, and reducing the cross-correlation between neutrinos and CDM. Some previous studies of neutrino wake effects can be found in [167–169] (also see [170] for cosmic string wakes), we will comment on similarities and differences between our approaches and those throughout this text. As we will see, the neutrino wake effects are also directly related to the CDM-ν dipole distortion produced by the relative velocity between CDM and neutrinos [171–173].

In this work we develop an analytic approach to offsets in neutrino and CDM clustering on small scales. To do this, we leverage the fact that the physics of neutrino clustering simplifies in the limit that the neutrino free-streaming scale is large in comparison with the scale of nonlinearities. The CDM bulk flow produces a distortion in the neutrino density field that naturally emerges in nonlinear structure formation with massive neutrinos. This distortion can be explicitly extracted from the solution to the Boltzmann equation for the neutrino distribution function in the background of a nonlinear dark matter structure, while

remaining completely agnostic about the nonlinear dynamics of CDM. In contrast, previous approaches to neutrino wakes are anchored on the introduction of an ad-hoc effective relative displacement between CDM and neutrino fluids, from which the observational signals are obtained.[1] This framework is effective on scales below a relative velocity coherence length but it breaks down at sufficiently large scales. Our approach improves on the analytical understanding of neutrino wakes in an expanding universe and allows for a more complete treatment of observables.

In general, the problem of including massive neutrinos in nonlinear structure formation is rather challenging and has been investigated in the literature using different approaches. That includes simulations that directly implement the neutrinos as N-body particles [65, 104, 106, 111, 113–116, 175, 176], but also (the computationally less expensive) simulations that assume massive neutrinos remain linearly clustered up to late times [122, 123, 126, 132, 148, 177, 178], and analytic approaches as well [136, 179–183]. The peculiar motion of halos and the associated distortion to the neutrino density field should, in principle, be included in some of these methods and indeed the existence of a CDM-ν dipole distortion has been explicitly verified in the TianNu simulations using a real space three point function correlating the CDM and neutrino densities with their relative velocity field [184]. Additionally, signatures of neutrino wakes were also observed in simulations that directly integrate the six-dimensional Boltzmann (Vlasov)-Poisson equations in phase space [185].

Our novel theoretical framework enables a throughout investigation of potential observational signatures of this dipole distortion effect; it slows halos down due to dynamical friction, decreases the cross correlation between CDM and neutrinos, and leaves an imprint in three point cross correlations of matter and galaxies, or the bispectrum in Fourier space, peaking at squeezed triangle configurations. As we shall see, this observable benefits from cosmic variance cancellation: it is proportional to the particular realization of the CDM power spectrum on small scales. Furthermore, it is not limited by the (lack of) knowledge of

[1] This is analogous to the moving background perturbation theory approach to modeling the effects of the relative bulk motion between CDM and baryons on the formation of the first structures [174].

optical depth to reionization,[2] and only has a small degeneracy with additional contributions to the bispectrum from standard nonlinear structure formation. This is then a promising additional independent signature of neutrino masses in the large-scale structure that can be probed in future surveys and complement the standard method based on the small-scale matter power suppression in the presence of massive neutrinos. While the suppression in the matter power spectrum from neutrinos is expected to provide the most stringent limits on the neutrino mass, the power spectrum method is limited by the knowledge of optical depth to reionization as probed by the cosmic microwave background (CMB), uncertainties associated with modeling complex nonlinear structure and galaxy formation physics, and degeneracies with other new physics such as a dynamical dark energy component (see e.g. [98, 191–193]). The neutrino wake effects discussed in this paper are free from some of these limitations.

Our analysis of wakes only assumes generic features of a thermally produced collisionless non-cold dark species and could therefore apply to other sources of hot dark matter. Nevertheless, we work with neutrinos for definiteness. In numerical calculations throughout the paper we assume a reference flat ΛCDM cosmology, with $\Omega_{\rm m,0} = \Omega_{\rm c,0} + \Omega_{\rm b,0} + \Omega_{\nu,0} = 0.32$, $\Omega_{\rm b,0} = 0.05$, $\Omega_{\Lambda,0} = 0.68$ and $h = 0.67$. We emphasize that the neutrino effects considered in this paper scale with powers of the individual neutrino masses, rather than the total energy density in neutrinos. This is in contrast to the linear-theory suppression of the matter power spectrum, which scales with the total energy density in neutrinos as $\sum_i m_{\nu i}$. For simplicity we will work with a single massive species with mass m_ν and consider the dynamical friction and neutrino distortions produced by that mass state. The net effects of three massive neutrinos can be computed straightforwardly. For a given neutrino mass m_ν we adjust the fractional contribution of CDM to the energy density $\Omega_{\rm c,0}$ in such a way as to keep the total matter fraction $\Omega_{\rm m,0}$ fixed. Linear transfer functions are obtained with the Cosmic Linear Anisotropy Solving System (CLASS) [47], and we compute power spectra assuming an almost

[2]Indeed, both of these qualities are common to multi-tracer approaches to measuring the neutrino mass with large-scale structure [186–190].

scale invariant primordial power defined by a tilt $n_s = 0.96$ and amplitude $A_s = 2.2 \times 10^{-9}$ at the pivot scale $k_{\rm piv} = 0.05{\rm Mpc}^{-1}$. Throughout this paper we treat baryons and CDM identically and henceforth CDM refers both.

The paper is organized as follows: In Section 5.3 we review the formalism necessary to study the distortion in the neutrino density field produced by the peculiar motion of halos. In Section 5.4 we apply it to the simple case of a moving point mass halo that approximates the distribution of CDM structure at sufficiently small scales. We also derive the 1-halo contribution to the dynamical friction effect, and discuss the limitations associated to neglecting the large-scale gravitational potential. In Section 5.5 we apply our formalism to a general nonlinear CDM distribution where we only impose kinematic relations and hence remain agnostic about the nonlinear dynamics of CDM. We then revisit the dynamical friction effect and comment on large-scale structure observables. In Section 5.6 we carry out some numerical calculations using the halo model to derive the 2-halo contribution to the dynamical friction effect, and standard perturbation theory (SPT) to compute two and three point cross correlations of neutrinos and CDM. The latter are used to forecast the detectability of neutrino wakes effect in future surveys combining two tracers of the matter distribution. In 5.7 we summarize our findings. Finally, in Appendix 5.8 we present the details of signal-to-noise calculations.

5.3 Basic formalism

We follow the standard approach to study how neutrinos are distributed around CDM, i.e., to solve for the neutrino phase space distribution function $f_\nu(t, \vec{x}, \vec{q})$, where t is the Friedmann-Lemaître-Robertson-Walker (FLRW) cosmic time, $\vec{x}$ the comoving position, and $\vec{q} = m_\nu a^2(t) d\vec{x}/dt$ the comoving momentum in the cosmic frame. Here m_ν is the neutrino mass and $a(t)$ is the scale factor. We work with one neutrino mass eigenstate at a time. The collisionless Boltzmann equation reads [19],

$$\frac{\partial f_\nu}{\partial t} + \frac{d\vec{x}}{dt} \cdot \frac{\partial f_\nu}{\partial \vec{x}} + \frac{d\vec{q}}{dt} \cdot \frac{\partial f_\nu}{\partial \vec{q}} = 0 \,. \tag{5.1}$$

The neutrinos move in the gravitational potential $\phi(t, \vec{x})$ of CDM, from where it follows that,

$$\frac{d\vec{q}}{dt} = -m_\nu \frac{\partial \phi}{\partial \vec{x}} \,. \tag{5.2}$$

Three approximations are being made here. First, we consider the motion of neutrinos in the external gravitational potential of CDM, ignoring the back reaction of neutrinos on the gravitational potential. This is a good approximation as long as the fractional contribution of neutrinos to the matter density is small. We also neglect large-scale general relativistic corrections, such as dilation effects via ϕ' terms. These do not contribute to the physics on the sub-horizon scales we consider. Finally, we assume that neutrinos are nonrelativistic, so our equations apply at sufficiently late times. Note that for a neutrino mass $m_\nu \sim 0.1$eV, the transition between relativistic and nonrelativistic regimes happens at a redshift $z_{\rm tr} \sim m_\nu/3T_{\nu,0} \sim 200$, much earlier than the onset of nonlinear structure formation.

The Boltzmann Eq. (5.1) can then be written as:

$$\frac{\partial f_\nu}{\partial t} + \frac{\vec{q}}{m_\nu a^2} \cdot \frac{\partial f_\nu}{\partial \vec{x}} - m_\nu \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f_\nu}{\partial \vec{q}} = 0 \,. \tag{5.3}$$

We split the neutrino phase-space distribution function as follows: $f_\nu = f_{\nu,0} + f_{\nu,1}$, where $f_{\nu,0}$ is a solution to the case $\phi = 0$, and $f_{\nu,1}$ represents a response to the gravitational field of dark matter. Setting $\phi = 0$ in Eq. (5.3) gives,

$$\frac{\partial f_{\nu,0}}{\partial t} + \frac{\vec{q}}{m_\nu a^2} \cdot \frac{\partial f_{\nu,0}}{\partial \vec{x}} = 0 \,. \tag{5.4}$$

A general solution to Eq. (5.4), which is also isotropic in the cosmic frame, is an arbitrary function of the magnitude of the comoving momentum $f_{\nu,0} = f_{\nu,0}(q)$. Since we are considering the distribution of a single relic neutrino species, we set this to be the relativistic Fermi-Dirac distribution $f_{\nu,0}(q) = f_{\rm FD}(q/T_{\nu,0})$, with $T_{\nu,0} \approx 1.95$K the neutrino temperature today, and:

$$f_{\rm FD}(x) = \frac{g_\nu}{e^x + 1} \,, \tag{5.5}$$

with $g_\nu = 2$ to account for both left-handed neutrinos and right-handed antineutrinos. Going back to the general case of a nonvanishing CDM potential, we arrive at

$$\frac{\partial f_{\nu,1}}{\partial t} + \frac{\vec{q}}{m_\nu a^2} \cdot \frac{\partial f_{\nu,1}}{\partial \vec{x}} = m_\nu \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f_{\nu,0}}{\partial \vec{q}} + m_\nu \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f_{\nu,1}}{\partial \vec{q}} \,. \tag{5.6}$$

We now make the Brandenberger-Kaiser-Turok (BKT) approximation which consists in dropping the second term in the right-hand side of Eq. (5.6) [194]. This can be phrased as the statement that we compute the linear response of the neutrino distribution to the (nonlinear) gravitational potential of CDM. We expect the approximation to break down at sufficiently small scales, or in close encounters of neutrinos with very massive halos, in which case one can proceed to solve the Eq. (5.6) iteratively [183]. We are interested in the perturbative quasi-linear regime where the BKT approximation should be adequate.

The Boltzmann Eq. (5.6) then greatly simplifies to:

$$\frac{\partial f_{\nu,1}}{\partial \eta} + \frac{\vec{q}}{m_\nu} \cdot \frac{\partial f_{\nu,1}}{\partial \vec{x}} = m_\nu a^2(\eta) \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f_{\nu,0}}{\partial \vec{q}} \,, \tag{5.7}$$

where we also introduced a new time coordinate η via $dt/d\eta = a^2(\eta)$, proportional to the comoving distance traveled by a nonrelativistic particle throughout cosmic history. In the remainder of this manuscript we will study how the solution to Eq. (5.7) captures the distortion in the neutrino density field induced by the peculiar motion of halos. As a warm-up, in Sec. 5.4 we will first solve it for the simple case that ϕ is the potential for a single moving point mass halo that works as a toy model for the distribution of CDM at sufficiently small scales. The single-halo approach neglects the clustering of nearby halos and the overall incoherence of the velocity field of nearby structures. In Sec. 5.5 we will consider Eq. (5.7) in the context of the full large-scale CDM density field. Our approach will be to use the Poisson equation and the fully nonlinear continuity equation to develop a systematic expansion for perturbations to the neutrino density in terms of derivatives of the CDM density field. In this sense, our predictions for the neutrino wakes do not rely on any particular implementation of non-linear gravitational evolution.

5.4 A warm-up exercise: Neutrino wakes around a single moving halo

Our goal in this section is to determine the late-time clustering of relic neutrinos around a moving point mass halo with a given mass M, following the comoving trajectory $\vec{x}_{\rm H}(\eta)$. We first compute the neutrino energy density and then proceed to determine the resulting

dynamical friction force. The section ends with a discussion on the limitations associated to neglecting the large-scale gravitational potential which provides intuition for some of the calculations that are carried out in Section 5.5.

5.4.1 Anisotropic neutrino density profile

As a first step we solve Eq. (5.7) for the gravitational potential of a moving point mass halo. In general, the solution to Eq. (5.7) reads

$$f_{\nu,1}(\eta,\vec{x},\vec{q}) = \int d\eta' \int d^3\vec{x}'\; G(\eta-\eta',\vec{x}-\vec{x}')m_\nu a^2(\eta')\frac{\partial\phi}{\partial\vec{x}'}\cdot\frac{\partial f_{\nu,0}}{\partial\vec{q}}\,, \tag{5.8}$$

where:

$$G(\eta,\vec{x}) = \Theta(\eta)\delta^{(3)}\left(\vec{x}-\frac{\vec{q}}{m_\nu}\eta\right)\,, \tag{5.9}$$

is the causal Green's function of Eq. (5.7), with $\Theta(\eta)$ the Heaviside step function. The substitution of Eq. (5.9) into Eq. (5.8) gives, after some algebra,

$$f_{\nu,1}(\eta,\vec{x},\vec{q}) = m_\nu\int_0^\eta d\eta' a^2(\eta')\frac{\partial\phi}{\partial\vec{y}}\cdot\frac{\partial f_{\nu,0}}{\partial\vec{q}}\,, \tag{5.10}$$

where we set $\eta = 0$ corresponding to the arbitrary initial integration time and

$$\vec{y} = \vec{x} - \frac{\vec{q}}{m_\nu}(\eta-\eta')\,. \tag{5.11}$$

It is now straightforward to compute the perturbation to the neutrino energy density as follows:

$$\begin{aligned}\delta\rho_\nu(\eta,\vec{x}) &= m_\nu a^{-3}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}\; f_{\nu,1}(\eta,\vec{x},\vec{q})\,,\\ &= m_\nu^2 a^{-3}(\eta)\int_0^\eta d\eta' a^2(\eta')\int\frac{d^3\vec{q}}{(2\pi)^3}\;\frac{\partial\phi}{\partial\vec{y}}\cdot\frac{\partial f_{\nu,0}}{\partial\vec{q}}\,.\end{aligned} \tag{5.12}$$

After integration by parts, and using Eq. (5.11), this becomes

$$\delta\rho_\nu(\eta,\vec{x}) = m_\nu a^{-3}(\eta)\int_0^\eta d\eta'(\eta-\eta')a^2(\eta')\int\frac{d^3\vec{q}}{(2\pi)^3}\; f_{\nu,0}(q)\nabla_y^2\phi(\eta',\vec{y})\,. \tag{5.13}$$

For a point mass halo with a given mass M and following the comoving trajectory $\vec{x}_{\rm H}(\eta)$,

$$\nabla_y^2\phi(\eta',\vec{y}) = \frac{4\pi GM}{a(\eta')}\delta^{(3)}(\vec{y}-\vec{x}_{\rm H}(\eta'))\,, \tag{5.14}$$

with Newton's gravitational constant G. The substitution of Eq. (5.14) into Eq. (5.13) yields, after integrating over the comoving momentum to eliminate the delta function,

$$\delta\rho_\nu(\eta,\vec{x}) = \frac{GMm_\nu^4}{2\pi^2}a^{-3}(\eta)\int_0^\eta d\eta' \frac{a(\eta')}{(\eta-\eta')^2} f_{\rm FD}\left(\frac{m_\nu}{T_{\nu,0}}\frac{|\vec{x}-\vec{x}_{\rm H}(\eta')|}{\eta-\eta'}\right), \tag{5.15}$$

where we used the relation $f_{\nu,0}(q) = f_{\rm FD}(q/T_{\nu,0})$, along with Eq. (5.5) (with $g_\nu = 2$). This is the expression for the mass density perturbation of relic neutrinos in the presence of a moving point mass halo, that we will apply in the next section in order to study the dynamical friction effect.

For now, let us build some intuition on the physics at hand by considering the case of a halo moving with constant velocity $\vec{v}_{\rm H}$ in a static universe, i.e., $a(t) = 1$ and $\vec{x}_{\rm H}(t) = \vec{x}_{\rm H}(0) + \vec{v}_{\rm H}t$ with $\eta = t$. In this case, the Eq. (5.15) reads in the halo frame

$$\delta\rho_\nu^{\rm H}(t,\vec{x}) = \delta\rho_\nu(t,\vec{x}+\vec{x}_{\rm H}(t)) \underset{t\to\infty}{=} \frac{GMm_\nu^3 T_{\nu,0}}{2\pi^2}\int_0^\infty du\ f_{\rm FD}(|u\vec{x}+\vec{\alpha}|), \tag{5.16}$$

where we introduce the vector

$$\vec{\alpha} = \frac{m_\nu a}{T_{\nu,0}}\vec{v}_{\rm H}, \tag{5.17}$$

which points in the direction of the halo motion and whose magnitude is proportional to the ratio of halo to thermal neutrino velocities, and also changed the integration variable to $u = (m_\nu/T_{\nu,0})(1/t - t')$ before taking the static limit $t \to \infty$. Equation (5.16) is plotted in Fig. 5.1 for a neutrino mass $m_\nu = 0.1$eV, and a halo with mass $M = 10^{13}M_\odot$ moving with a velocity $v_{\rm H} = 500$km/s along the x-axis. Note that the halo peculiar velocity creates a distortion in the neutrino density field, which clusters anisotropically behind the moving halo.

5.4.2 Dynamical friction on a single halo

We now have an expression for the anisotropic clustering of neutrinos around moving point mass halos in an expanding universe, i.e., Eq. (5.15). We expect this to generate dynamical friction that slows the halo down and this is what we compute next. Using Newton's third

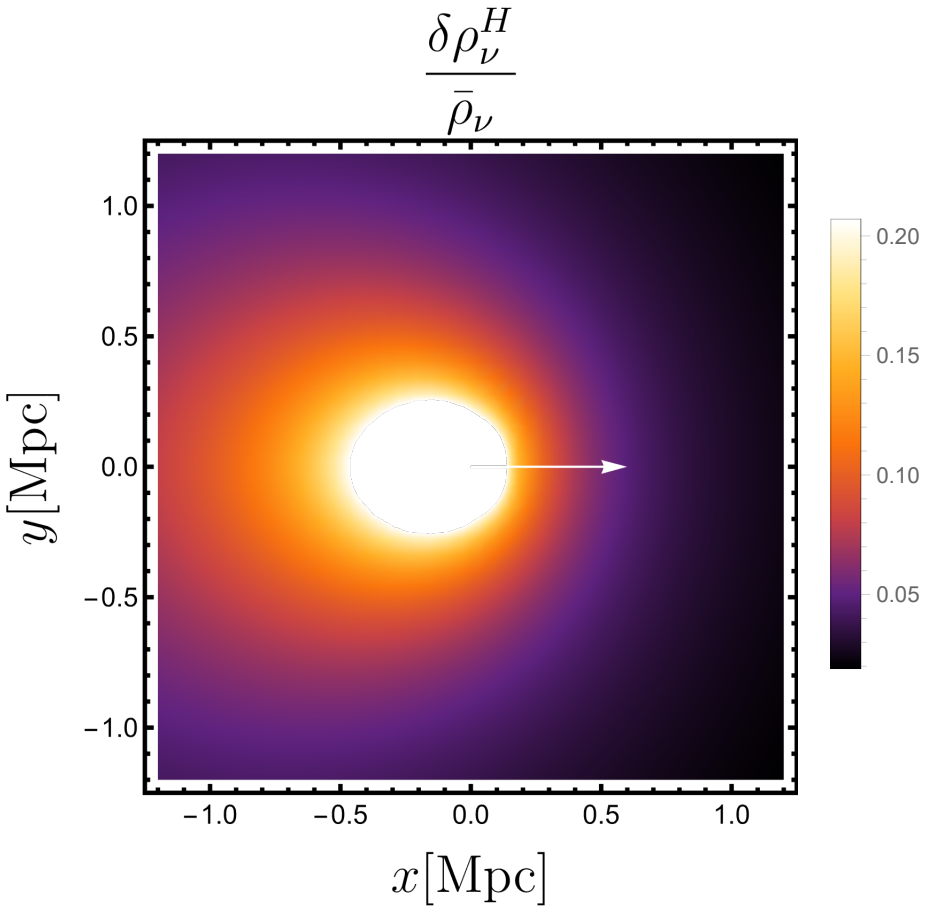

Figure 5.1: Anisotropic clustering of massive neutrinos behind a point mass halo with a constant velocity in a static universe. Here $m_\nu = 0.1$eV is the neutrino mass, $M = 10^{13} M_\odot$ is the halo mass and $v_{\rm H} = 500$km/s is the halo velocity along the x-axis as illustrated by the white arrow. The plot shows the neutrino density contrast in a frame of reference comoving with the halo, which diverges at the origin (the point mass halo position) and hence we introduce a cutoff at an arbitrary maximum density contrast which here corresponds to the value of $\delta^H_{\nu,\rm max} = 0.2$.

law, the force acting on the halo can be obtained by integrating the neutrino mass density

over the gravitational field of the halo

$$\vec{F}(\eta) = -a^3(\eta) \int d^3\vec{x}\ \delta\rho_\nu(\eta,\vec{x})\vec{g}_{\rm H}(\eta,\vec{x})\,, \tag{5.18}$$

where for a moving point mass halo in an expanding universe the gravitational field reads

$$\vec{g}_{\rm H}(\eta,\vec{x}) = -\frac{\vec{\nabla}_x\phi(\eta,\vec{x})}{a(\eta)} = -\frac{GM}{a^2(\eta)}\frac{\vec{x}-\vec{x}_{\rm H}(\eta)}{|\vec{x}-\vec{x}_{\rm H}(\eta)|^3}\,. \tag{5.19}$$

Substituting Eqs. (5.15) and (5.19) into Eq. (5.18) gives

$$\vec{F}(\eta) = \frac{G^2M^2m_\nu^4}{2\pi^2}\int_0^\eta \frac{d\eta'}{(\eta-\eta')^2}\frac{a(\eta')}{a^2(\eta)}\int d^3\vec{x}\ f_{\rm FD}\left(\frac{m_\nu}{T_{\nu,0}}\frac{|\vec{x}-\vec{x}_{\rm H}(\eta')|}{\eta-\eta'}\right)\frac{\vec{x}-\vec{x}_{\rm H}(\eta)}{|\vec{x}-\vec{x}_{\rm H}(\eta)|^3}\,, \tag{5.20}$$

we now change the integration variable in the second integral in the right hand side of Eq. (5.20). In terms of $\vec{y} = (m_\nu/T_{\nu,0})(\vec{x}-\vec{x}_H(\eta'))/(\eta-\eta')$ the integral is formally equivalent to the problem of calculating the gravitational field generated by the isotropic mass density $f_{\rm FD}(y)$. Applying Gauss's law we find

$$\vec{F}(\eta) = -\frac{2}{\pi}G^2M^2m_\nu^4\left(\frac{T_{\nu,0}}{m_\nu}\right)^3\int_0^\eta\frac{d\eta'}{\eta-\eta'}\frac{a(\eta')}{a^2(\eta)}\frac{\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}}{|\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}|^3}\int_0^{\frac{m_\nu}{T_{\nu,0}}|\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}|} dy\ y^2 f_{\rm FD}(y)\,. \tag{5.21}$$

In the limit $\eta'\to\eta$ we have that $(\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta'))/(\eta-\eta')\to a(\eta)\vec{v}_H(\eta)$, where $\vec{v}_{\rm H} = ad\vec{x}_{\rm H}/dt$ is the halo peculiar velocity, and the integral is singular due to the factor of $(\eta-\eta')$ in the denominator. To regularize it we introduce the dimensionless time integration variable $\Delta = \eta/(\eta-\eta')$ in terms of which the Eq. (5.21) becomes

$$\vec{F}(\eta) = -\frac{2}{\pi}G^2M^2m_\nu^4\left(\frac{T_{\nu,0}}{m_\nu}\right)^3\int_1^\infty\frac{d\Delta}{\Delta}\frac{a(\eta')}{a^2(\eta)}\frac{\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}}{|\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}|^3}\int_0^{\frac{m_\nu}{T_{\nu,0}}|\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}|} dy\ y^2 f_{\rm FD}(y)\,, \tag{5.22}$$

with $\eta' = \eta(1-1/\Delta)$ implicit on Eq. (5.22). The limit $\eta'\to\eta$ is mapped into $\Delta\to\infty$ where the integrand approaches $\sim d\Delta/\Delta$, i.e., we encounter a logarithmic divergence. This is to be expected in our simple model of a point mass halo since gravity is a long range force and hence neutrinos coming towards the halo with arbitrarily large impact parameter give a nonvanishing contribution [195]. However, in realistic scenarios one has to account for

the large-scale gravitational potential which sets a cutoff Λ for the integral. We postpone a discussion of this to the next subsection and for now simply introduce the cutoff by hand. The expression in Eq. (5.22) greatly simplifies in the limit $\Lambda \gg 1$, which motivates the following regularization scheme:

$$\vec{F}(\eta) = -\frac{2}{\pi}\log\Lambda G^2 M^2 m_\nu^4 \left(\frac{T_{\nu,0}}{m_\nu}\right)^3 \lim_{\Lambda\to\infty}\frac{1}{\log\Lambda}\int_1^\Lambda \frac{d\Delta}{\Delta}\frac{a(\eta')}{a^2(\eta)}\frac{\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}}{|\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}|^3}\int_0^{\frac{m_\nu}{T_{\nu,0}}|\frac{\vec{x}_{\rm H}(\eta)-\vec{x}_{\rm H}(\eta')}{\eta-\eta'}|} dy\, y^2 f_{\rm FD}(y)\,. \tag{5.23}$$

the integral is now dominated by its contribution from $\Delta \gg 1$ so we arrive at,

$$\vec{F} = -\frac{2}{\pi}\log\Lambda G^2 M^2 m_\nu^4 \left(\frac{T_{\nu,0}}{m_\nu a}\right)^3 \frac{\vec{v}_{\rm H}}{v_{\rm H}^3}\int_0^\alpha dy\; y^2 f_{\rm FD}(y)\,, \tag{5.24}$$

where the time dependence is now implicit in Eq. (5.24), and $\alpha = |\vec{\alpha}|$ is the magnitude of the vector defined in Eq. (5.17). This is the Chandrasekhar dynamical friction formula [196] as used, for instance, in the pioneering paper [167]. Before moving forward, let us stop for a moment to investigate the explicit dependence of Eq. (5.22) on the cutoff Λ in the simple scenario where the halo is only subject to Hubble drag, that is, assuming $v_{\rm H} \sim a^{-1}$ and hence $d\vec{x}/d\eta = a\vec{v}_{\rm H}$ is constant. In this case when evaluating Eq. (5.22) at redshift $z = 0$ (or $a = 1$) we reproduce Eq. (5.24) but now with an additional factor of:

$$\frac{F}{F_{\Lambda\to\infty}} = \frac{1}{\log\Lambda}\int_1^\Lambda \frac{d\Delta}{\Delta} a\left(\eta\left(1-\frac{1}{\Delta}\right)\right)\,. \tag{5.25}$$

This can be calculated in our reference ΛCDM cosmology for any given initial redshift z_i associated to the halo formation time (where we set $\eta_i = 0$) and is plotted in Fig. 5.2. One can see the trend of $F/F_{\Lambda\to\infty} \to 1$ in the limit $\Lambda \to \infty$, with Chandrasekhar's formula in Eq. (5.24) being suppressed for any finite value of the cutoff Λ (also note that the suppression is larger for earlier forming halos which move faster at higher redshift and hence probe larger distance scales). We will give a physical interpretation for this result in the next subsection where we discuss the limitations associated to ignoring the large-scale gravitational potential in our simple single halo model.

We are primarily interested in the small halo velocity regime $\alpha \ll 1$ of Eq. (5.24), suitable for realistic neutrino masses as constrained by cosmology. In this case the dynamical friction

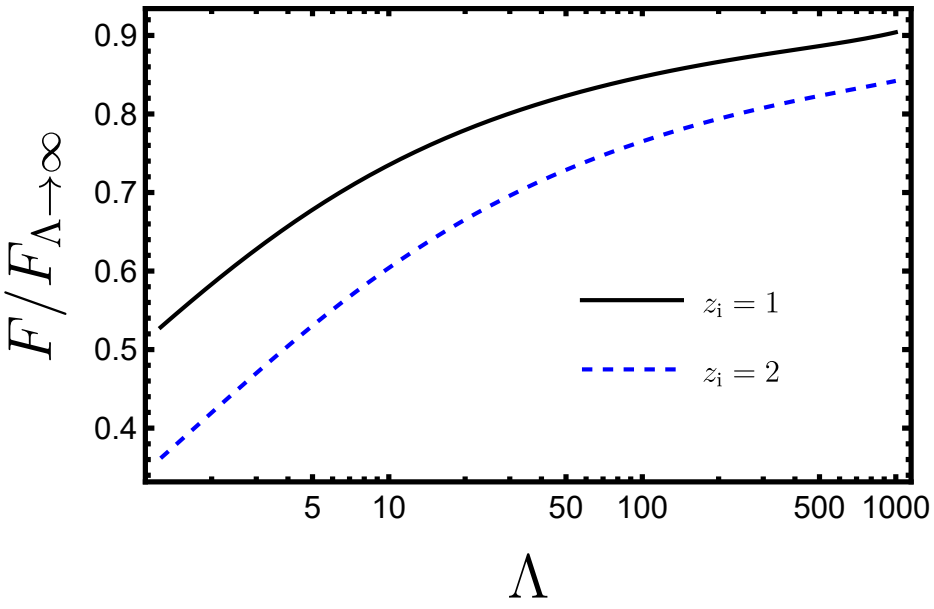

Figure 5.2: Suppression of Chandrasekhar's formula for the dynamical friction force $F/F_{\Lambda\to\infty}$ for finite values of the cutoff Λ in the simple case of a halo subject to Hubble drag. The solid black curve corresponds to a choice of halo formation redshift of $z_\mathrm{i} = 1$, and the dashed blue curve to $z_\mathrm{i} = 2$. In realistic applications we might expect $\Lambda \approx \lambda_\mathrm{coh}/R_\mathrm{halo} \approx 10 - 1000$ as we will see in Sec. 5.6.

force is proportional to the halo velocity,

$$\vec{F} \approx -\frac{2}{3\pi} \log \Lambda G^2 M^2 m_\nu^4 \vec{v}_\mathrm{H} \,, \tag{5.26}$$

and scales like m_ν^4 such that the dynamical friction effect is dominated by the most massive neutrino eigenstate. We may then assume a single neutrino species for simplicity, or explicitly write the total dynamical friction force as a sum of individual contribution from different eigenstates.

In order to connect Eq. (5.26) with the results of future sections, it is convenient to rewrite it in terms of a quantity with dimension of inverse time. Since $F = M dv_\mathrm{H}/dt$, we

can define:

$$\tau^{-1} = -\frac{\vec{F}\cdot\vec{v}_{\rm H}}{Mv_{\rm H}^2} = \frac{2}{3\pi}\log\Lambda G^2 M m_\nu^4 = 3.4\times10^{-5}\frac{\log\Lambda}{\log 100}\frac{M}{10^{13}M_\odot}\left(\frac{m_\nu}{0.1{\rm eV}}\right)^4 H_0\,, \qquad (5.27)$$

which is the characteristic time scale for an order one fractional decrease in the halo velocity due to the dynamical friction effect. Note that $1/\tau H_0 = \Delta v/v$ is the overall relative decrease in the halo velocity over the age of the Universe $t \sim 1/H_0$. We obtain a numerical value of $\Delta v/v = 3.4\times10^{-5}$ for a halo mass $M = 10^{13}M_\odot$ and individual neutrino mass $m_\nu = 0.1$eV, when also assuming $\Lambda = 100$. This already suggests that the dynamical friction effect is quite small, although it can pick up some significant contributions from the clustering of nearby halos as we will see in Sec. 5.6.

5.4.3 Limitations to the 1-halo approach

Thus far we have determined the anisotropic clustering of massive neutrinos behind moving point mass halos and the corresponding dynamical friction force. A more realistic calculation would have to account for both the finite extent of the halo and the presence of large-scale structure. Indeed, the Eq. (5.27) involves an unknown Coulomb logarithm, $\log\Lambda$, where in typical applications of the dynamical friction formula the cutoff Λ can be estimated as the ratio of maximum and minimum impact parameters, $\Lambda \sim b_{\rm max}/b_{\rm min}$ [195]. Here $b_{\rm min} \sim R_{\rm halo}$ is the halo radius, and $b_{\rm max} \sim \lambda_{\rm coh} \sim 0.1$ Mpc^{-1} is the CDM velocity coherence scale. The CDM bulk flow is only coherent over sufficiently small scales and hence our analysis based on a single moving halo is expected to break down at scales $\lambda \gtrsim \lambda_{\rm coh}$.[3] This point will be made more clear in the next section, where we also provide a precise definition for the velocity coherence scale.

[3]We should also impose a cutoff corresponding to the distance traveled by free-streaming neutrinos, which sets the scale where neutrino inhomogeneities are coherent with CDM. As we shall see the neutrino free-streaming scale is much larger than the CDM velocity coherence length $\lambda_{\rm coh}$. Relatedly, since our formalism assumes neutrinos are non-relativistic, one might also expect that we should impose a cutoff corresponding to the distance neutrinos have traveled while non-relativistic. Yet this scale is much larger than the free-streaming scale and, of course, the CDM velocity coherence length $\lambda_{\rm coh}$ and is therefore irrelevant.

When accounting for the large-scale structure, the halo velocity $v_{\rm H}$ is replaced by the root mean square CDM velocity dispersion σ given by,

$$\sigma^2 = \langle v^2 \rangle = \int \frac{dk}{2\pi^2} P_{\theta\theta}(z,k)\,, \tag{5.28}$$

with $P_{\theta\theta}(z,k)$ the power spectrum of the divergence of the CDM velocity field $\theta = \vec{\nabla}\cdot\vec{v}$ (note that the velocity dispersion depends on redshift). This can be used to define the scale associated to the CDM bulk flow, which is also known as the scale of nonlinearities[4]

$$k_{\rm NL} = \frac{aH}{\sigma}\,, \tag{5.29}$$

where a is the scale factor and $H = \dot{a}/a$ is the Hubble rate, with dot denoting a derivative with respect to cosmic time t. We evaluate the Eq. (5.28) numerically in our reference ΛCDM cosmology and in linear perturbation theory, using $P_{\theta\theta}(z,k) = (faH)^2[D_{\rm L}(z)/D_{\rm L}(z=0)]^2 P_{\delta\delta}(z=0,k)$, where $P_{\delta\delta}(z=0,k)$ is the CDM power spectrum evaluated at $z=0$. The linear growth factor $D_{\rm L}$ and linear growth rate $f = d\log D_{\rm L}/d\log a$ are obtained with the approximate formulas in [197].

In our calculations thus far we have implicitly assumed a large-scale homogeneous distribution of neutrinos. In general, neutrinos have a large velocity dispersion[5],

$$\sigma_\nu = \sqrt{\frac{3\zeta(3)}{\log 4}}\frac{T_{\nu,0}}{m_\nu a}\,, \tag{5.30}$$

and hence free-stream over cosmological distances with a characteristic scale $k_{\rm fs}$ of,

$$k_{\rm fs}^2 = \frac{4\pi G a^2 \bar{\rho}(a)}{\sigma_\nu^2} = \frac{3}{2}\Omega_{\rm m}(a)\left(\frac{aH}{\sigma_\nu}\right)^2\,, \tag{5.31}$$

[4] In the literature one often finds alternative definitions for the scale of nonlinearities, such as the scale at which the dimensionless power is unity, which we will here denote by $\tilde{k}_{\rm NL}$. In our reference ΛCDM cosmology one finds $\tilde{k}_{\rm NL} \sim 0.15{\rm Mpc}^{-1} \sim k_{\rm NL}$ at $z=0$ so the two definitions produce numerically equivalent results. However, their time evolution is different and this can produce significant differences at higher redshift. For example, $\tilde{k}_{\rm NL} \approx 0.5{\rm Mpc}^{-1}$ at $z=1$, while $k_{\rm NL}$ changes only slightly when compared to its $z=0$ value, as depicted in Fig. 5.3.

[5] In later sections we provide the motivation for this particular definition of the neutrino velocity dispersion.

where $\bar{\rho}(a)$ is the background matter energy density and we used the Friedmann equation. In Fig. 5.3 we plot the scale of nonlinearities, $k_{\rm NL}$, and the neutrino free-streaming scale, $k_{\rm fs}$, as a function of redshift in our reference ΛCDM cosmology for three values of the neutrino mass: $m_\nu = 0.05$eV, $m_\nu = 0.1$eV and $m_\nu = 0.15$eV.

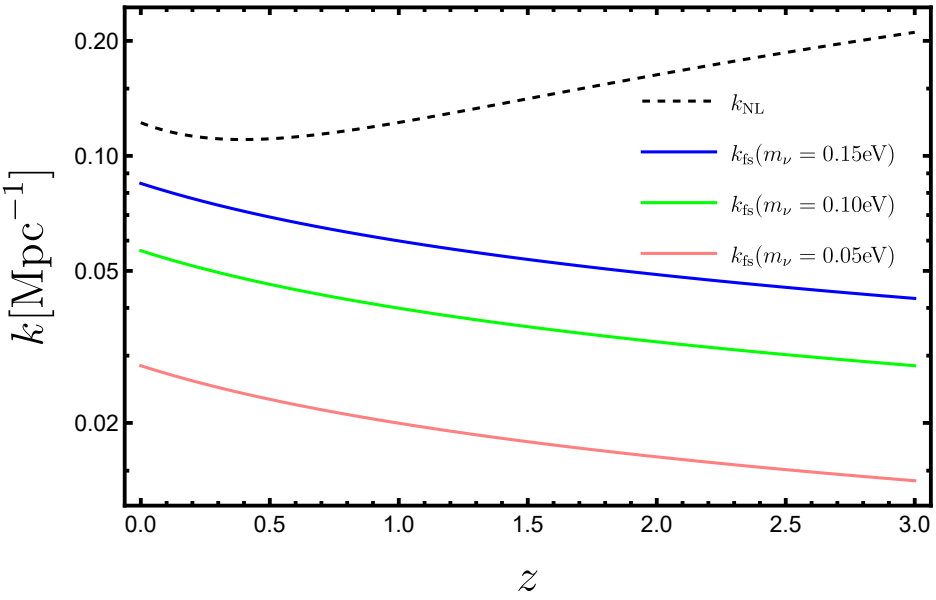

Figure 5.3: Scale of nonlinearities, $k_{\rm NL}$, and neutrino free-streaming scale, $k_{\rm fs}$, as a function of redshift and for varying neutrino mass. The black dashed curve is the scale of nonlinearities and blue, green and pink solid curves correspond to the free-streaming scale for neutrino masses of $m_\nu = 0.05$eV,$m_\nu = 0.1$eV and $m_\nu = 0.15$eV respectively.

At scales smaller than the free-streaming scale ($k \gtrsim k_{\rm fs}$), pressure dominates over the gravitational field and neutrino anisotropies are washed-out. On the other hand, at larger scales ($k \ll k_{\rm fs}$) neutrinos cluster like cold matter and there are significant anisotropies in the neutrino fluid.[6] As a consequence we expect our previous analysis, which assumes a

[6]In practice the transition at the free-streaming scale is rather smooth, with the neutrino species fully behaving like CDM only at scales above the neutrino horizon scale which is approximately set by the free-streaming scale when evaluated at the moment when neutrinos first became nonrelativistic and lies at much larger scales. See Fig. 1 in [80] for more details.

homogeneous distribution of neutrinos, to break down when the CDM flows over scales that are larger than $\lambda_{\rm fs} \sim 1/k_{\rm fs}$, i.e., when $k_{\rm fs} \gtrsim k_{\rm NL}$. From Fig. 5.3 we conclude that for the small neutrino masses as constrained by cosmology, $k_{\rm fs} < k_{\rm NL}$ and hence our assumptions are consistent.

Note that since $k_{\rm fs} \sim aH/\sigma_\nu$ and $k_{\rm NL} \sim aH/\sigma$, we have that $k_{\rm fs}/k_{\rm NL} \sim \sigma/\sigma_\nu$ is analogous to the parameter α introduced in Eq. (5.17) and this has two important consequences. First, as we showed before, the limit $\alpha \ll 1$ allowed for a significant simplification of the dynamical friction effect, leading to the simple formula in Eq. (5.27). In an approach that accounts for the large-scale structure, we expect a similar simplification to happen in the limit $k_{\rm fs}/k_{\rm NL} \ll 1$, and indeed we will use this as a guiding principle in future sections. Second, due to the breakdown of our framework in the regime $k_{\rm fs}/k_{\rm NL} \sim \sigma/\sigma_\nu \gtrsim 1$, the Chandrasekhar formula in Eq. (5.24) can no longer be applied when $\alpha \gtrsim 1$. If we try to ignore this and proceed to extract the $\alpha \gg 1$ limit in Eq. (5.24) we would (erroneously) arrive at,

$$\vec{F} \approx -\frac{6}{\pi}\zeta(3)\log\Lambda G^2 M^2 m_\nu^4 \left(\frac{T_{\nu,0}}{m_\nu a}\right)^3 \frac{\vec{v}_{\rm H}}{v_{\rm H}^3}\,, \tag{5.32}$$

where we used Eq. (5.5) with $g_\nu = 2$. For a given fixed neutrino temperature today $T_{\nu,0}$, one would then conclude that $F \sim m_\nu$ and hence the dynamical friction effect can be made arbitrarily large by increasing m_ν, which in turn would enable us to place an upper bound on the neutrino mass scale based solely on the observation that galaxies have some nonzero peculiar velocities today and hence the dynamical friction effect cannot be too large [167]. This argument completely ignores the cutoff at the free-streaming scale and is hence incorrect. Indeed, we can get some valuable intuition from considering the limit $k_{\rm fs}/k_{\rm NL} \sim \sigma/\sigma_\nu \gg 1$ (analogous to $\alpha \gg 1$), which corresponds to a neutrino mass so large that the neutrinos are no longer streaming over cosmological distances and hence effectively behave just like a cold matter component. Then neutrinos will cluster along with the CDM and there no longer exists a homogeneous sea of neutrinos around the halo. Another way to think about it is that in this limit, neutrinos are comoving with CDM and then, in the frame of the halo, the situation is equivalent to a halo at rest in a homogeneous sea of neutrinos. In that case,

neutrino clustering is isotropic in the halo frame and there are no wakes and no dynamical friction.

In later sections we will see precisely how the free-streaming scale comes into play when generalizing the approach of this section to a more realistic dark matter distribution. Another limitation to Eq. (5.27) is that it neglects the effects of clustering between nearby halos. In studying neutrino wakes and dynamical friction in large-scale structure, we then expect to see some corrections accounting for the fact that nearby halos will share neutrino wakes. The two-halo term was studied in detail in [168] under the assumption of a static universe, as seen from the frame of reference of a moving halo. This approach implicitly assumes that all of the CDM moves coherently, and hence breaks down at scales above the CDM velocity coherence scale. In the next section, we will develop a framework that incorporates the expansion of the Universe and also allows us to work with CDM structures that have a velocity dispersion, allowing us to push the regime of validity of the calculations all the way up to the free-streaming scale.

5.5 Neutrino wakes in large-scale structure

In the previous section we determined that the peculiar motion of halos generates anisotropic neutrino wakes that slow halos down due to dynamical friction. Our goal now is to extended the analysis of neutrino wakes to a general CDM distribution (in an expanding universe) in order to account for the large scale structure. This improves on several limitations of the previous approach: The cutoffs at both the free-streaming and CDM coherence scale will naturally emerge from the formalism, as opposed to being introduced by hand to regularize divergent integrals. We will also derive a 2-halo term to the dynamical friction formula and naturally account for the CDM velocity dispersion, as opposed to introducing an ad-hoc halo peculiar velocity. Finally, this approach allows us to derive corrections to the matter power spectrum from neutrino wakes and compute the matter-neutrino cross-bispectrum generated by the distortions to the neutrino field.

Based on our intuition from the previous section, we expect the problem to significantly

simplify in the limit $k_{\rm fs}/k_{\rm NL} \ll 1$. As we will see, by working in this regime we are able to address the problem analytically. Our strategy is to solve the Boltzmann Eq. (5.7) in Fourier space, integrate the solution over comoving momenta to obtain the neutrino density perturbation, and then consider a Taylor expansion in the small parameter $k_{\rm fs}/k_{\rm NL} \ll 1$. We will then calculate both the zeroth and first order terms. The latter will correspond to the distortion in the neutrino density field that we are looking for. In our calculations we remain agnostic about the nonlinear evolution of CDM perturbations, imposing only the continuity and Poisson equations. Our final expressions in this section will be written in terms of CDM power spectra and bispectra without assuming a particular model for their computation.

5.5.1 Solution to the Boltzmann equation for an arbitrary CDM density field

Going back to the problem at hand, we need to consider Eq. (5.7) in its full generality. In Fourier space (let $\vec{\nabla} \to i\vec{k}$), it reads:

$$\frac{\partial f_{\nu,1}}{\partial \eta} + i\frac{\vec{k}\cdot\vec{q}}{m_\nu} f_{\nu,1} = i m_\nu a^2(\eta)\phi(\eta,\vec{k})\vec{k}\cdot\frac{\partial f_{\nu,0}}{\partial \vec{q}}\,, \tag{5.33}$$

where we remind the reader that $f_{\nu,0}(q)$ corresponds to the background neutrino distribution function, while $f_{\nu,1}(\eta,\vec{k},\vec{q})$ is the linear response to the CDM gravitational potential $\phi(\eta,\vec{k})$. Furthermore, $\vec{q}$ is the comoving momentum, and we are working in the time coordinate η defined by $dt/d\eta = a^2(\eta)$. This is a first-order ordinary differential equation whose solution is straightforward to write down [7]

$$f_{\nu,1}(\eta,\vec{k},\vec{q}) = i m_\nu \vec{k}\cdot\frac{\partial f_{\nu,0}}{\partial \vec{q}}\int_0^\eta d\eta' a^2(\eta')\phi(\eta',\vec{k}) e^{-i\frac{\vec{k}\cdot\vec{q}}{m_\nu}(\eta-\eta')}\,, \tag{5.34}$$

and hence the perturbation to the neutrino mass density is,

$$\begin{aligned}\delta\rho_\nu(\eta,\vec{k}) &= m_\nu a^{-3}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3} f_{\nu,1}(\eta,\vec{k},\vec{q})\,,\\ &= i m_\nu^2 a^{-3}\vec{k}\cdot\int_0^\eta d\eta' a^2(\eta')\phi(\eta',\vec{k})\int\frac{d^3\vec{q}}{(2\pi)^3}\frac{\partial f_{\nu,0}}{\partial\vec{q}} e^{-i\frac{\vec{k}\cdot\vec{q}}{m_\nu}(\eta-\eta')}\,,\end{aligned} \tag{5.35}$$

[7] The distribution function quickly loses sensitivity to the initial conditions at sub-horizon scales, hence we set $\phi(0,\vec{k}) = 0$.

now integrate the second integral by parts, followed by integration over the solid angle, to arrive at:

$$\delta\rho_\nu(\eta,\vec{k}) = -\frac{m_\nu^2 k}{2\pi^2}a^{-3}(\eta)\int_0^\eta d\eta' a^2(\eta')\phi(\eta',\vec{k})\int_0^\infty dq\; q f_{\nu,0}(q)\sin\left[\frac{kq}{m_\nu}(\eta-\eta')\right]. \quad (5.36)$$

To proceed use Poisson's equation $k^2\phi(\eta,\vec{k}) = -4\pi G a^2\delta\rho(\eta,\vec{k})$ with $\delta\rho(\eta,\vec{k})$ the CDM density in excess of the average $\bar{\rho}(\eta)$, i.e., $\rho(\eta,\vec{x}) = \bar{\rho}(\eta)+\delta\rho(\eta,\vec{x}) \equiv \bar{\rho}(\eta)[1+\delta(\eta,\vec{x})]$ in real space, with $\delta(\eta,\vec{x})$ the CDM density contrast. We also change the integration variable in the integral over comoving momenta to $x = q/T_{\nu,0}$, using the relation $f_{\nu,0}(q) = f_{\rm FD}(q/T_{\nu,0}) = f_{\rm FD}(x)$. Equation (5.36) now reads,

$$\delta_\nu(\eta,\vec{k}) = \frac{\delta\rho_\nu(\eta,\vec{k})}{\bar{\rho}_\nu(\eta)} = \frac{4\pi G\bar{\rho}_0}{3\zeta(3)}\int_0^\infty dx\, x^2 f_{\rm FD}(x)\int_0^\eta d\eta'(\eta-\eta')a(\eta')\delta(\eta',\vec{k})j_0\left[k\frac{T_{\nu,0}}{m_\nu}x(\eta-\eta')\right], \quad (5.37)$$

with $\bar{\rho}_0$ the comoving background matter energy density (or the matter density when evaluated at the present time), $\delta_\nu(\eta,\vec{k})$ the neutrino density contrast, $j_0(x) = \sin x/x$ the spherical Bessel function of zeroth order and the background neutrino energy density $\bar{\rho}_\nu(\eta)$ is given by

$$\bar{\rho}_\nu(\eta) = m_\nu a^{-3}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}f_{\nu,0}(\vec{q}) = \frac{3\zeta(3)}{2\pi^2}a(\eta)^{-3}m_\nu T_{\nu,0}^3\,, \quad (5.38)$$

where we used $f_{\nu,0}(q) = f_{\rm FD}(q/T_{\nu,0})$ and Eq. (5.5) with $g_\nu = 2$. Also note that $3\zeta(3) = \int_0^\infty dx\, x^2 f_{\rm FD}(x)$ sets the numerical factor in the denominator of Eq. (5.37). Equation 5.37 is a key equation from which we will derive several observables. We emphasize that when referring to the CDM density or gravitational potential we add no subscripts, while the corresponding neutrino quantities always carry the subscript ν. In this work we always assume a small fractional contribution of neutrinos to the total matter, $f_\nu = \bar{\rho}_\nu/\bar{\rho} \ll 1$.

We first consider scales that are above the Jeans length, $kT_{\nu,0}\eta/m_\nu \ll 1$. In this case we can work under the approximation $j_0(x) \approx 1$ for $x \ll 1$, and assume a linear evolution for the CDM field (since we are interested in neutrino masses for which $k_{\rm NL} > k_{\rm fs}$)

$$\delta(\eta',\vec{k}) = \frac{D_{\rm L}(\eta')}{D_{\rm L}(\eta)}\delta(\eta,\vec{k})\,, \quad (5.39)$$

with $D_{\rm L}(\eta)$ the linear growth factor, which satisfies the following differential equation:

$$\frac{d^2}{d\eta^2}D_{\rm L}(\eta) = 4\pi G\bar{\rho}_0 a(\eta) D_{\rm L}(\eta)\,. \tag{5.40}$$

It then follows from Eq. (5.37) that, for $kT_{\nu,0}\eta/m_\nu \ll 1$,

$$\delta_\nu(\eta,\vec{k}) \underset{k\ll k_{\rm fs}}{=} 4\pi G\bar{\rho}_0 \frac{\delta(\eta,\vec{k})}{D_{\rm L}(\eta)} \int_0^\eta d\eta'(\eta-\eta')a(\eta')D_{\rm L}(\eta')\,, \tag{5.41}$$

where we used Eq. (5.39). Substituting Eq. (5.40) into Eq. (5.41) yields after integrating by parts,

$$\delta_\nu(\eta,\vec{k}) \underset{k\ll k_{\rm fs}}{=} \delta(\eta,\vec{k})\left[1-\eta\frac{D'_{\rm L}(0)}{D_{\rm L}(\eta)} - \frac{D_{\rm L}(0)}{D_{\rm L}(\eta)}\right] \approx \delta(\eta,\vec{k})\,, \tag{5.42}$$

where in the second equality we use the fact that the additional terms in the square brackets can be made negligible by setting the initial conditions at sufficiently early times (which is equivalent to just keeping the growing mode). This is the standard result that at scales above the neutrino horizon the neutrino species exactly traces the cold dark matter. In our derivation we had to assume a linear evolution for the CDM component at scales above the neutrino horizon, with the linear growth factor satisfying Eq. (5.40), in order to be consistent within our framework of a linear neutrino response. However, the final result is general as it is essentially the statement of causality and hence holds in any nonlinear approach, as long as one consistently accounts for the nonlinearities in both neutrinos and CDM.

Let us now investigate the opposite regime, $kT_{\nu,0}\eta/m_\nu \gg 1$, in Eq. (5.37). The spherical Bessel function peaks at $\Delta\eta = \eta - \eta' = 0$ and becomes vanishing when,

$$k\frac{T_{\nu,0}}{m_\nu}\eta\frac{\Delta\eta}{\eta} \gg 1 \iff \frac{\Delta\eta}{\eta} \gg \frac{1}{k\frac{T_{\nu,0}}{m_\nu}\eta}\,, \tag{5.43}$$

so the time integral in Eq. (5.37) becomes dominated by its contribution from $\Delta\eta/\eta \ll 1$, which suggests a Taylor series expansion of the integrand around $\eta' = \eta$. To leading order we can simply evaluate the integrand at $\eta' = \eta$, which implies that the clustering of neutrinos only depends on a snapshot of the CDM distribution at the final time. This can be interpreted as follows: Neutrino fluctuations with a (comoving) wavenumber k have a

characteristic time scale, $t \sim a/k\sigma_\nu \sim (k_{\rm fs}/k)H^{-1}$, which is much smaller than a Hubble time at sub-free streaming scales. We can then effectively consider a quasi-static space in terms of comoving quantities. However, since our goal in this paper is to study the effects of halo peculiar velocities, we need to go beyond the leading order term. Indeed, as long as $k_{\rm fs}/k_{\rm NL} \sim \sigma/\sigma_\nu \ll 1$, it suffices to consider the next to leading order term in the Taylor series expansion since the CDM bulk flow is slow when compared to the neutrino time scale. In what follows we compute the contributions from zeroth and first order terms in the expansion.

Zeroth order: The static limit

We begin with the zeroth-order term which corresponds to approximating $a(\eta')\delta(\eta', \vec{k}) \approx a(\eta)\delta(\eta, \vec{k})$ when evaluating the time integral in Eq. (5.37). As previously discussed, this is justified since the integral is dominated by its contribution from $\eta' = \eta$ in the limit where $kT_{\nu,0}\eta/m_\nu \gg 1$. This yields

$$\begin{aligned}
\delta_\nu^{(0)}(\eta, \vec{k}) &= \frac{4\pi G\bar{\rho}_0}{3\zeta(3)} a(\eta)\delta(\eta, \vec{k}) \int_0^\infty dx\, x^2 f_{\rm FD}(x) \int_0^\eta d\eta' (\eta - \eta') j_0\left[k\frac{T_{\nu,0}}{m_\nu} x(\eta - \eta')\right] \\
&= \frac{4\pi G\bar{\rho}_0}{3\zeta(3)} a(\eta)\delta(\eta, \vec{k}) \left(\frac{m_\nu}{kT_{\nu,0}}\right)^2 \int_0^\infty dx f_{\rm FD}(x) \left[1 - \cos\left(k\frac{T_{\nu,0}}{m_\nu} x\eta\right)\right] \\
&= \left(\frac{k_{\rm fs}}{k}\right)^2 \delta(\eta, \vec{k})\,,
\end{aligned} \tag{5.44}$$

where we used Eqs. (5.5), (5.30), (5.31), the Friedmann equation $\Omega_{\rm m}(a) = (8\pi G\bar{\rho}_0/3a^3H^2)$, and from the second to the third line we dropped the cosine term since $kT_{\nu,0}\eta/m_\nu \gg 1$. Note that our definition of the neutrino velocity dispersion, as given by Eq. (5.30), produces the result in the last line of Eq. (5.44) with no extra coefficients. Equation (5.44) is a well-known result at sub free-streaming scales, but also under the assumption that halos are static [163]. Ultimately, it is just the statement of neutrino free-streaming: The neutrino density perturbation is suppressed with respect to CDM at sub free-streaming scales.

In this work, however, we are interested in the effect produced by the peculiar motion of halos. That means we need to consider the leading order correction to Eq. (5.44), i.e., the first order term.

First order: Accounting for CDM bulk flows

Going back to Eq. (5.37), we now need to expand $a(\eta')\delta(\eta', \vec{k})$ to first order in $\Delta\eta = \eta - \eta'$. On one hand,

$$a(\eta') = a(\eta)\left[1 - a^2(\eta)H(\eta)(\eta - \eta') + \dots\right] . \tag{5.45}$$

Since only $\Delta\eta \lesssim m_\nu/kT_{\nu,0}$ contributes to the integral, and $k_{\rm fs} \sim (m_\nu/T_{\nu,0})a^2H$ from Eqs. (5.30) and (5.31), we find that the first order term in the right hand side of Eq. (5.45) scales like $\sim k_{\rm fs}/k$ with respect to the zeroth order term and can hence be neglected at sufficiently small scales. We also need to consider:

$$\delta(\eta', \vec{k}) = \delta(\eta, \vec{k}) - a^2(\eta)\dot{\delta}(\eta, \vec{k})(\eta - \eta') + \dots \tag{5.46}$$

where dot denotes a derivative with respect to cosmic time t. To proceed we assume the continuity equation for the CDM fluid in an expanding universe

$$\dot{\delta} = -\frac{1}{a}\vec{\nabla}\cdot\vec{P}\,, \tag{5.47}$$

when written in position space with $\vec{P} = (1+\delta)\vec{v}$ the CDM momentum (density) field, where $\vec{v}(\eta, \vec{x})$ is the CDM velocity. The ratio of first to zeroth order terms in Eq. (5.46) is then, after using Eq. (5.47) in Fourier space,

$$\frac{a(i\vec{k}\cdot\vec{P})\Delta\eta}{\delta} \lesssim \frac{1}{\sigma_\nu}\frac{P}{\delta}\,, \tag{5.48}$$

where once again only $\Delta\eta \lesssim m_\nu/kT_{\nu,0}$ contributes to the integral, $\sigma_\nu \sim T_{\nu,0}/m_\nu a$ from Eq. (5.30), and we introduce the notation $kP = i\vec{k}\cdot\vec{P}$ for the divergence of the momentum field. In the linear regime we have $P = v = (faH)\delta/k$ is the linear CDM velocity field, a tracer of the CDM density field [with $f = d\log D_{\rm L}/d\log a$ the linear growth rate]. The ratio of first to zeroth order terms in Eq. (5.48) then becomes $\sim aH/\sigma_\nu k \sim k_{\rm fs}/k$ once again. We reproduce a well-known important result: In linear perturbation theory the leading order correction to Eq. (5.44) is negligible at sufficiently small scales.

However, we now argue that the nonlinearities in the CDM momentum field, parameterized by $\vec{p} = \vec{P} - \vec{v}$, or $\vec{p} = \delta\vec{v}$ in position space (we also use the notation $kp = i\vec{k}\cdot\vec{p}$ for

its divergence) yield a nonnegligible contribution to next to leading order which has to be accounted for. The reason for this is as follows: At small enough scales the CDM moves coherently, with $v \sim \sigma$ the CDM velocity dispersion as defined in Eq. (5.28) in which case $p \sim \sigma\delta$. This gives a contribution to the ratio of first to zeroth order terms as given in Eq. (5.48) of size $\sim \sigma/\sigma_\nu \sim k_{\rm fs}/k_{\rm NL}$, which is independent of k. At scales $k \gtrsim k_{\rm NL}$ this becomes larger than the contributions scaling like $\sim k_{\rm fs}/k$ which we have neglected before, and it persists down to arbitrarily small scales. Such contributions cannot be neglected when studying nonlinear effects which turn on at scales $k \gtrsim k_{\rm NL}$.

The ratio $k_{\rm fs}/k_{\rm NL}$ is analogous to the parameter α we required to be small in Sec. 5.4.3 as a consistency condition for our framework in terms of halos moving through a homogeneous sea of neutrinos. Here, the smallness of $\sigma/\sigma_\nu \sim k_{\rm fs}/k_{\rm NL}$ is required in order to have a well-defined expansion in Eq. (5.46), with the physical interpretation that we can truncate the expansion at next to leading order when the typical halo motions are slow with respect to the characteristic time scale associated to the neutrino dynamics. We reinforce that this is consistent with current bounds on the sum of neutrino mass eigenstates from cosmological observations, as can be seen from Fig. 5.3.

We are now ready to compute the leading order correction to Eq. (5.44), i.e., to first order in $k_{\rm fs}/k_{\rm NL}$. We substitute Eq. (5.47) into the second term in the right hand side of Eq. (5.46). It then follows from Eq. (5.37), after dropping contributions that are negligible at sub-free streaming scale as discussed above:

$$\begin{aligned}
\delta_\nu^{(1)}(\eta,\vec{k}) &= \frac{4\pi G\bar{\rho}_0}{3\zeta(3)}a^2(\eta)kp(\eta,\vec{k})\int_0^\infty dx\, x^2 f_{\rm FD}(x)\int_0^\eta d\eta'(\eta-\eta')^2 j_0\left[k\frac{T_{\nu,0}}{m_\nu}x(\eta-\eta')\right] \\
&= \frac{4\pi G\bar{\rho}_0}{3\zeta(3)}a^2(\eta)kp(\eta,\vec{k})\int_0^\infty dx\, x^2 f_{\rm FD}(x)\left(\frac{m_\nu}{kT_{\nu,0}x}\right)^3\left[\sin\left(k\frac{T_{\nu,0}}{m_\nu}x\eta\right)-k\frac{T_{\nu,0}}{m_\nu}x\eta\cos\left(k\frac{T_{\nu,0}}{m_\nu}x\eta\right)\right] \\
&= \gamma\left(\frac{k_{\rm fs}}{k}\right)^2\frac{p(\eta,\vec{k})}{\sigma_\nu}\,,
\end{aligned} \tag{5.49}$$

where we used Eqs. (5.5), (5.30), (5.31), the Friedmann equation $\Omega_{\rm m}(a) = (8\pi G\bar{\rho}_0/3a^3H^2)$, and from the second to the third line we dropped the cosine term since $kT_{\nu,0}\eta/m_\nu \gg 1$. We

also introduced a numerical factor set by the (relativistic) Fermi-Dirac profile

$$\gamma = \frac{\pi}{6\zeta(3)} \left(\frac{3\zeta(3)}{\log 4} \right)^{\frac{3}{2}} \approx 1.83 \,. \tag{5.50}$$

We can now combine Eqs. (5.44) and (5.49) into an expression for the neutrino density contrast at sub-free streaming scales

$$\delta_\nu(\vec{k}) \underset{k \gg k_{\rm fs}}{=} \left(\frac{k_{\rm fs}}{k} \right)^2 \left[\delta(\vec{k}) + \gamma \frac{p(\vec{k})}{\sigma_\nu} + \mathcal{O}\left(\frac{k_{\rm fs}}{k_{\rm NL}} \right)^2 \right] . \tag{5.51}$$

Where the reader is reminded that $\vec{p}$ is the non-linear contribution to the CDM momentum, $\vec{p} = \vec{P} - \vec{v}$, or $\vec{p} = \delta\vec{v}$. Henceforth we omit the time dependence for simplicity of notation. At scales much larger than the free-streaming length the neutrinos exactly trace the cold dark matter as given by Eq. (5.42). On the other hand, at scales below the free-streaming length the neutrino density contrast is mostly proportional to the CDM density contrast with a scale dependent factor, but it also traces the CDM (nonlinear) momentum through a term that scales like $\sim k_{\rm fs}/k_{\rm NL} \ll 1$ which produces a distortion in the neutrino density field due to the CDM bulk flow, i.e., the peculiar motion of halos. This is a nonlinear effect at sub-free streaming scales and hence is not accounted for in linear perturbation theory. However, it is principle accounted for in implementations of massive neutrinos in nonlinear structure formation that consistently evolves neutrino perturbations in the nonlinear gravitational potential of CDM (such as N-body simulations).

It is important to point out that in our calculations we have remained agnostic about the nonlinear gravitational evolution, imposing only kinematic relations. We would need to consider explicit models for the nonlinear dynamics of CDM, such as N-body simulations or perturbative methods, in order to evaluate the higher order terms in Eq. (5.51) since they would involve a derivative of the velocity which requires the equations of motion. Additionally, in our approach we do not have a handle on the neutrino density contrast at scales that are comparable to the free-streaming scale as that would require evaluating the time integrals in Eq. (5.37), which also requires a model for the nonlinear dynamics of CDM.

In this work we adopt two strategies to go around those issues. The first is to assume that $k_{\rm fs}/k_{\rm NL} \ll 1$. As depicted in Fig. 5.3, this a very good approximation for an individual neutrino mass of $m_\nu = 0.05$eV. However, in the case of a large individual neutrino mass of $m_\nu = 0.15$eV (corresponding to $\sum m_{\nu i} \approx 0.45$eV), this small parameter can be as large as $k_{\rm fs}/k_{\rm NL} \sim 2/3$ producing $(2/3)^2 \sim 40\%$ level corrections to our calculations so that we can still confidently predict the overall magnitude of distortion effects, even for such high values for the individual neutrino mass. Our second strategy is to introduce a simple interpolation between Eqs. (5.42) and (5.51) to ensure that our final formula for the neutrino density contrast can be applied at all scales. For example, under the approximation that the neutrino density contrast is proportional to the CDM density contrast with a scale dependent factor [which is equivalent to dropping the second $p(\vec{k})$ term in the right-hand side of Eq. (5.51)], the following formula is known to give a good approximation at all scales [132]:

$$\delta_\nu(\vec{k}) \approx \frac{1}{(1+\frac{k}{k_{\rm fs}})^2}\delta(\vec{k}) = \frac{\left(\frac{k_{\rm fs}}{k}\right)^2}{(1+\frac{k_{\rm fs}}{k})^2}\delta(\vec{k})\,. \tag{5.52}$$

Note that this reduces to both Eqs. (5.42) and (5.51) in the appropriate limits, as it must. We now slightly generalize this prescription to account for the additional second $p(\vec{k})$ term in the right-hand side of Eq. (5.51). Since this contribution must not be present at scales $k \ll k_{\rm fs}$, we simply add an additional power in the denominator,

$$\begin{aligned}\delta_\nu(\vec{k}) &\approx \frac{\left(\frac{k_{\rm fs}}{k}\right)^2}{(1+\frac{k_{\rm fs}}{k})^2}\delta(\vec{k}) + \gamma\frac{\left(\frac{k_{\rm fs}}{k}\right)^2}{(1+\frac{k_{\rm fs}}{k})^3}\frac{p(\vec{k})}{\sigma_\nu} \\ &= \frac{\left(\frac{k_{\rm fs}}{k}\right)^2}{(1+\frac{k_{\rm fs}}{k})^2}\left[\delta(\vec{k}) + \gamma\frac{p_{\rm rel}(\vec{k})}{\sigma_\nu}\right]\,,\end{aligned} \tag{5.53}$$

where we introduce the relative (nonlinear) momentum,

$$p_{\rm rel}(\vec{k}) = \frac{p(\vec{k})}{1+\frac{k_{\rm fs}}{k}}\,, \tag{5.54}$$

which is just a high-pass filtered nonlinear contribution to the CDM momentum field with a cutoff at the free-streaming scale. In that sense the distortion in the neutrino density field is sensitive to the relative bulk flow between CDM and neutrinos and not just the CDM

bulk flow, or in other words, we have to remember to impose the cutoff at the free-streaming scale as we did here (also see the discussion in Sec. 5.4.3 for a physical interpretation). In the limit $k_{\rm fs}/k_{\rm NL} \ll 1$ any nonlinear physics happens on scales where Eq. (5.51) safely applies and hence our interpolation is not really needed. However, as discussed above we will consider applications with m_ν somewhat larger than the current cosmological bounds where $k_{\rm fs}/k_{\rm NL} \lesssim 2/3$ for which the interpolation can play an important role, i.e., to cutoff the distortion effect at the free-streaming scale. We then expect some inaccuracies in our modeling that nevertheless allow for an estimation of the size of the distortion effect in potential observables.

We can gain some intuition on Eq. (5.53) by considering what it looks like in case one simply introduces a constant (position independent) relative velocity between the neutrino and CDM fluids $\vec{v}_{\rm rel} = \sigma_{\rm rel}\hat{z}$ (say along the z direction); this is a good approximation at scales below the velocity coherence length (which is also at sub-free streaming scales) and is indeed the approach taken in previous works [168, 169, 171, 172]. In that case $p_{\rm rel}(\vec{k}) = i(\hat{k}\cdot\hat{z})\sigma_{\rm rel}\delta(\vec{k})$ and Eq. (5.51) reads

$$\delta_\nu(\vec{k}) = \left(\frac{k_{\rm fs}}{k}\right)^2 \left[1 + i\gamma(\hat{k}\cdot\hat{z})\frac{\sigma_{\rm rel}}{\sigma_\nu}\right]\delta(\vec{k})\,. \tag{5.55}$$

An imaginary term appears, which corresponds to a dipole distortion of the neutrino density field along the direction of the relative velocity. Furthermore, we are working in the regime $\sigma_{\rm rel}/\sigma_\nu \ll 1$, where

$$e^{i\gamma(\hat{k}\cdot\hat{z})\frac{\sigma_{\rm rel}}{\sigma_\nu}} \approx 1 + i\gamma(\hat{k}\cdot\hat{z})\frac{\sigma_{\rm rel}}{\sigma_\nu}\,, \tag{5.56}$$

and hence we can absorb the distortion in a phase that corresponds to a shift in coordinates in position space, i.e., going to the neutrino rest frame. To briefly elaborate on this point, if the CDM velocity field is completely coherent, then one can study the problem in the rest frame of CDM, where the neutrino momentum distribution appears to have a dipole component. In this situation the Fourier-space momentum field $p(\vec{k})$ is simply proportional to the density $\delta(\vec{k})$, as opposed to related via a convolution over the velocity field. Our Equation 5.53 is more general in that it accounts for the fact that the relative bulk flow is

not coherent over arbitrarily large scales, so one cannot generally boost to a frame where all the CDM is at rest. This approach therefore allows for a more complete treatment of potential observational signatures which henceforth will be the focus of our attention.

5.5.2 Dynamical friction revisited

We have investigated the solution to the Boltzmann equation for the neutrino distribution function in the background of a general nonlinear CDM distribution, and we showed how one can extract the distortion in the neutrino density field produced by the CDM bulk flow, building up to our central result in Eq. (5.53). Our next step is to estimate the impact of this on the CDM dynamics. The motivation for doing so is two fold: First, we want to connect the formalism described in the previous section with our 1-halo results, or more specifically Eq. (5.27). This will enable us to derive the 2-halo term, and to confirm that the second term in the right-hand side of Eq. (5.53), the nonlinear momentum term, is indeed producing the dynamical friction effect. Second, if the impact of dynamical friction on the CDM is found to be large, i.e., halos are slowing down significantly due to the anisotropic neutrino wakes then we may use that as an avenue for detection of the dynamical friction effect, through a halo velocity bias or direct measurements of galaxy peculiar velocities, e.g., via the kinetic Sunyaev Zeldovich (kSZ) effect. Otherwise, we should focus on potential signatures of Eq. (5.53) on the large-scale structure as observables.

The additional clustering of neutrinos as given by the second term in the right-hand side of Eq. (5.53) translates into a contribution to the large-scale gravitational field, using Poisson's equation,

$$\vec{g}(\vec{k}) = 4\pi G a \frac{i\hat{k}}{k}\delta\rho_\nu^{(1)}(\vec{k}) = \frac{4\pi m_\nu^4 G^2 \bar{\rho}_0}{(1+\frac{k_{\rm fs}}{k})^3} p(\vec{k})\frac{i\hat{k}}{k^3}\,, \tag{5.57}$$

where $\delta\rho_\nu^{(1)}(\vec{k}) = \bar{\rho}_\nu\delta_\nu^{(1)}(\vec{k})$ is the distortion in the neutrino density perturbation produced by the CDM bulk flow and we used Eqs. (5.30), (5.31), (5.38), (5.50) and (5.53). Note that time dependences are implicit in our equations. We wish to compute the large-scale structure analog of the characteristic time scale in Eq. (5.27). In this case the force per unit

mass acting on the CDM structure is given by $\delta\vec{g}$, since the integral of the density times the gravitational field produces the overall force on the CDM within a given region of space. Additionally, due to the stochastic nature of cosmological fields we need to take the average such that we arrive at:

$$\begin{aligned}\tau^{-1} &= -\frac{\langle\,\vec{v}\cdot\delta\;\vec{g}\,\rangle}{\langle v^2\rangle} = -\frac{1}{\sigma^2}\langle\,\vec{p}\cdot\vec{g}\,\rangle \\ &= \frac{2}{\pi}\frac{G^2 m_\nu^4}{\sigma^2}\bar{\rho}_0\int d\log k\,\frac{P_{pp}(k)}{(1+\frac{k_{\rm fs}}{k})^3}\,,\end{aligned} \tag{5.58}$$

where in the first line we used Eq. (5.28) in combination with $\vec{p}=\delta\vec{v}$, and Eq. (5.57) to arrive at the second line.[8] In order to make progress in Eq. (5.58) we need an expression for the power spectrum of the (nonlinear) momentum field $P_{pp}(k)$. This is what we discuss next, following the work of [199]. We are interested in the divergence $p(\vec{k})=i\hat{k}\cdot\vec{p}(\vec{k})$. In position space $\vec{p}=\delta\vec{v}$ so we need to evaluate the convolution to obtain,

$$p(\vec{k}) = \sigma\frac{k_{\rm NL}}{k}\int\frac{d^3\vec{k}_1}{(2\pi)^3}\int\frac{d^3\vec{k}_2}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}_{12})\alpha(\vec{k}_1,\vec{k}_2)\delta(\vec{k}_1)\frac{\theta(\vec{k}_2)}{aH}\,, \tag{5.59}$$

where $\vec{k}_{12}=\vec{k}_1+\vec{k}_2$, $\alpha(\vec{k}_1,\vec{k}_2)=(\vec{k}_{12}\cdot\vec{k}_2)/k_2^2$, $\delta^{(3)}(\vec{k})$ stands for the Dirac delta function and for future convenience we have inserted a factor of $\sigma k_{\rm NL}/aH=1$. It is now straightforward to compute the two-point function. We find:

$$\langle p(\vec{k})p(\vec{k}')\rangle = (2\pi)^3\delta^{(3)}(\vec{k}+\vec{k}')P_{pp}(k)\,, \tag{5.60}$$

where,

$$P_{pp}(k) = \sigma^2\int\frac{d^3\vec{k}'}{(2\pi)^3}\left(\frac{k_{\rm NL}}{k'}\right)^2(\hat{k}\cdot\hat{k}')\left\{(\hat{k}\cdot\hat{k}')P_{\delta\delta}(|\vec{k}-\vec{k}'|)\frac{P_{\theta\theta}(k')}{(aH)^2}+\frac{kk'\left[1-\frac{k'}{k}(\hat{k}\cdot\hat{k}')\right]}{|\vec{k}-\vec{k}'|^2}\frac{P_{\delta\theta}(|\vec{k}-\vec{k}'|)}{aH}\frac{P_{\delta\theta}(k')}{aH}\right\}. \tag{5.61}$$

For wavelengths below the velocity coherence scale, $k\gg k_{\rm coh}$, we can expand the integrand to leading order in k'/k to obtain [200]

$$P_{pp}(k)\underset{k\gg k_{\rm coh}}{=}\sigma^2 P_{\delta\delta}(k)\int\frac{d^3\vec{k}'}{(2\pi)^3}\left(\frac{k_{\rm NL}}{k'}\right)^2(\hat{k}\cdot\hat{k}')^2\frac{P_{\theta\theta}(k')}{(aH)^2} = \frac{\sigma^2}{3}P_{\delta\delta}(k)\,, \tag{5.62}$$

[8] We assume that the CDM nonlinear momentum can be written as the gradient of a potential. We do not expect vorticity to contribute at the large quasilinear scales of our interest $\sim k_{\rm NL}\sim 0.1{\rm Mpc}^{-1}$, e.g., see Fig.4 in [198].

where we have used Eqs. (5.28) and (5.29). In words, the CDM bulk flow is coherent on scales $k \gg k_{\rm coh}$, with velocity $v = \sigma/\sqrt{3}$. On the other hand, at larger scales this is no longer the case causing a suppression in the momentum power spectrum when compared to Eq. (5.62). This can be parameterized as follows: On large perturbative scales we can set $\theta(\vec{k}) = faH\delta(\vec{k})$ into Eq. (5.61), with $f = d\log D_{\rm L}/d\log a$ the linear growth rate, and evaluate the ratio

$$\begin{aligned}\beta(k) &= \left.\frac{P_{pp}(k)}{\frac{\sigma^2}{3}P_{\delta\delta}(k)}\right|_{\text{linear theory}} \\ &= 3f^2\int\frac{d^3\vec{k'}}{(2\pi)^3}\left(\frac{k_{\rm NL}}{k'}\right)^2(\hat{k}\cdot\hat{k'})\left\{(\hat{k}\cdot\hat{k'})+\frac{kk'\left[1-\frac{k'}{k}(\hat{k}\cdot\hat{k'})\right]}{|\vec{k}-\vec{k'}|^2}\right\}\frac{P^{\rm L}_{\delta\delta}(|\vec{k}-\vec{k'}|)P^{\rm L}_{\delta\delta}(k')}{P^{\rm L}_{\delta\delta}(k)}\,,\end{aligned} \tag{5.63}$$

with $P^{\rm L}_{\delta\delta}(k)$ the linear power spectrum. Note that $\beta(k) = 1$ for $k \gg k_{\rm coh}$ as we have demonstrated above, and this can be used as a definition of the velocity coherence scale. From $P^{\rm L}_{\delta\delta} \sim D^2_{\rm L}(a)$, Eqs. (5.28) and (5.29) it follows that $\beta(k)$ is independent of redshift, and hence so is the velocity coherence scale. This function is plotted in Fig. 5.4 for our reference ΛCDM cosmology from where we can read off $k_{\rm coh} \sim 0.1{\rm Mpc}^{-1}$. At scales above the velocity coherence scale, $k < k_{\rm coh}$, the momentum field, and hence the distortion effect, is suppressed. Our model for the momentum power spectrum at all scales now reads,

$$P_{pp}(k) = \frac{\sigma^2}{3}\beta(k)P_{\delta\delta}(k)\,, \tag{5.64}$$

with $P_{\delta\delta}(k)$ the nonlinear power spectrum. Substituting Eq. (5.64) into Eq. (5.58) yields

$$\tau^{-1}(a) = \frac{2}{3\pi}G^2m_\nu^4\bar{\rho}_0\int d\log k\,\frac{\beta(k)}{\left[1+\frac{k_{\rm fs}(a)}{k}\right]^3}P_{\delta\delta}(a,k)\,. \tag{5.65}$$

Note that for clarity we insert the time dependences in this final formula, which are the ones implicit through $P_{\delta\delta}(a,k)$ and $k_{\rm fs}(a)$. This can be readily evaluated with a nonlinear power spectrum for the neutrino masses and redshifts of interest. We use the Boltzmann code CLASS to produce the nonlinear power spectrum with the HMcode prescription [201].

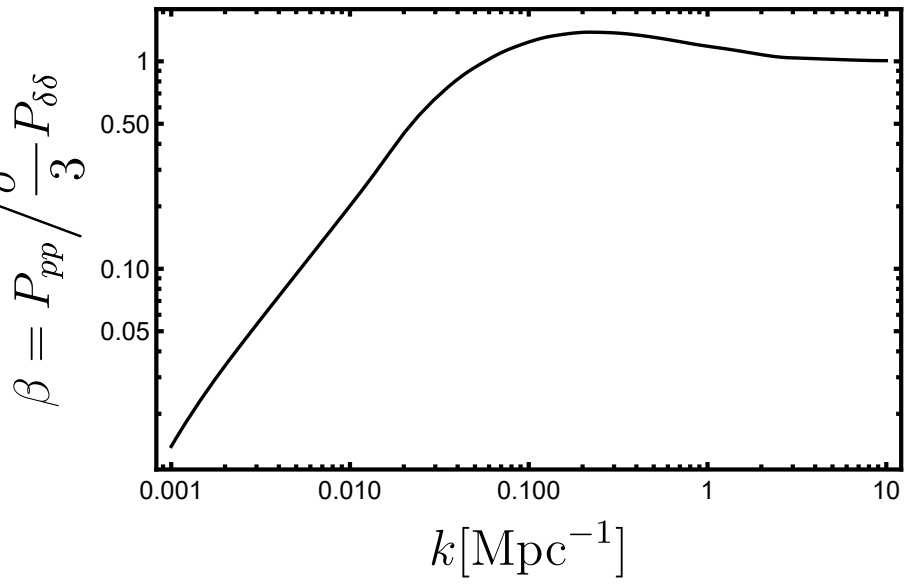

Figure 5.4: The function $\beta(k)$ as given by Eq. (5.63), which captures the coherence of the CDM velocity field as a function of k. The CDM bulk flow is coherent on scales $k \gg k_{\rm coh}$, with velocity $v = \sigma/\sqrt{3}$. On the other hand, at larger scales this is no longer the case causing a suppression in the momentum power spectrum relative to $\sigma^2 P_{\delta\delta}/3$. From the plot above we read off $k_{\rm coh} \sim 0.1 \mathrm{Mpc}^{-1}$ for our reference ΛCDM cosmology.

It turns out that the integral in Eq. (5.65) is dominated by its contribution from large quasilinear scales where the nonlinear corrections are small, and in fact we explicitly verified that simply using the linear power spectrum produces the same numerical results.[9] We then substitute $P_{\delta\delta}(a,k) = [D_{\rm L}(a)/D_{\rm L}(a=1)]^2 P^{\rm L}_{\delta\delta}(a=1,k)$ into Eq. (5.65) to obtain (using the approximate formula in [197] for the linear growth factor),

$$(\tau H)^{-1} \approx \sum_{m_\nu} 2.2 \times 10^{-4} \left(\frac{m_\nu}{0.1\mathrm{eV}}\right)^{2.8} (1+z)^{-1.83}\,, \tag{5.66}$$

where we assume our reference ΛCDM cosmology and add a sum over neutrino mass eigenstates. The reader is reminded that dynamical friction only occurs for neutrino masses small

[9] In the next section we will use the halo model to interpret this result as the statement that the 2-halo term dominates over the 1-halo term on average.

enough that the free-streaming scale is larger than the non-linear scale (roughly $m_\nu \lesssim 0.2$eV). The power law approximation in Eq. (5.66) reproduces the numerical results of Eq. (5.65) with errors that can be as large as tens of percents, for neutrino masses in the range 0.05eV $\leq m_\nu \leq$ 0.15eV, and redshifts $z \leq 2$. Note that $\Delta v/v \sim 1/\tau H$ is the net average fractional decrease in the velocity of halos due to the dynamical friction over the age of the universe $\sim 1/H$.

Increasing the neutrino mass causes a dramatic increase in the effect, but there is a significant deviation from the naive scaling from $\sim m_\nu^4$ to $\sim m_\nu^{2.8}$. Recall that our framework assumes $k_{\rm fs}/k_{\rm NL} \ll 1$, while $k_{\rm fs}/k_{\rm NL} \sim 2/3$ already for $m_\nu = 0.15$eV. At that point the suppression at the free-streaming scale is playing an important role, with the effect vanishing in the limit $k_{\rm fs}/k_{\rm NL} \gg 1$. This is relevant since a scenario where $k_{\rm fs}/k_{\rm NL} > 1$ has been ruled out by cosmological observations as we discussed previously, but laboratory experiments based on measurements of beta decay only set an upper bound to a weighted sum of the masses $m_{\nu,\beta} < 0.8$eV [152]. Furthermore, nonstandard scenarios such as dynamical dark energy and unstable neutrino species allow for less stringent bounds on the neutrino mass scale from cosmological probes [153–155, 202, 203]. Finally, at higher redshift the effect is smaller due to both linear growth of structure and the smaller age of the universe $\sim 1/H$.

The overall amplitude in Eq. (5.66) is quite small, indicating that the decrease in the velocity of halos due to the dynamical friction effect produced by anisotropic neutrino wakes are unlikely to be detected in the foreseeable future. That includes measurements of galaxy peculiar velocities via redshift space distortions, the kSZ effect, or a large-scale velocity bias that deviates from unity. In the absence of massive neutrinos, a large scale velocity bias of $b_v \equiv 1$ follows from the equivalence principle [204] (and any initial nonzero $|b_v - 1|$ decays away [205]). In [115] a $\gtrsim 0.5\%$ deviation from unity in the large scale velocity bias was not found in simulations with massive neutrinos, which is consistent with our results. Previous work has found a more pronounced dynamical friction effect [168], but there the cutoffs at the velocity coherence and free-streaming scales were not consistently accounted for.

We are still left with the task of connecting Eq. (5.65), our result for the velocity de-

cay time, to the results of Sec. 5.4, specifically our 1−halo calculation of that quantity in Eq. (5.27). We will postpone this discussion to Sec. 5.6 where we carry out numerical calculations using the halo model. There we will see that the 1-halo term will naturally reemerge, but now together with a 2-halo term accounting for the clustering of nearby halos.

5.5.3 Large-scale structure observables

We calculated the average fractional decrease in halo velocities due to dynamical friction and found that it is too small to allow for a detection of the anisotropic neutrino wakes. We then take the alternative route of searching for potential observational signatures in the large scale structure. For instance, we can use Eq. (5.53) to evaluate the CDM-ν cross power spectrum,

$$P_{\delta_\nu\delta}(k) = \frac{\left(\frac{k_{\rm fs}}{k}\right)^2}{(1+\frac{k_{\rm fs}}{k})^2}\left[P_{\delta\delta}(k) + \frac{\gamma}{1+\frac{k_{\rm fs}}{k}}\frac{P_{p\delta}(k)}{\sigma_\nu}\right] . \tag{5.67}$$

The additional second term in the right-hand side of Eq. (5.67) is the contribution from the distortion effect that we expect to be small, due to its scaling with $\sim k_{\rm fs}/k_{\rm NL}$, and negative since the relative bulk flow between neutrinos and CDM should reduce their cross correlation. We postpone some numerical calculations, using standard perturbartion theory (SPT) as our model for the nonlinear dynamics of CDM, to Sec. 5.6. For now we simply point out that the signature of the distortion effect in the two-point function is small and very degenerate with standard nonlinear structure formation.

Instead, we will argue that three-point cross-correlations allow for a separation of the distortion effect from standard gravitational nonlinearities, and that the resulting signal can be potentially measured in upcoming large scale structure surveys. We can use Eq. (5.53) to compute the neutrino-CDM-CDM bispectrum:

$$\langle\delta_\nu(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)B(k_1,k_2,k_3)\,, \tag{5.68}$$

where[10]

$$B(k_1,k_2,k_3)=\frac{\left(\frac{k_{\rm fs}}{k_1}\right)^2}{(1+\frac{k_{\rm fs}}{k_1})^2}\left[\langle\delta(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle'+\frac{\gamma}{1+\frac{k_{\rm fs}}{k_1}}\langle\frac{p(\vec{k}_1)}{\sigma_\nu}\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle'\right], \tag{5.69}$$

is written in terms of CDM-only bispectra.[11] Note that the first term inside the brackets in the right-hand side of Eq. (5.69) is symmetric under the exchange $\vec{k}_i \leftrightarrow \vec{k}_j$ for any pair of wavevectors, while the second term is only symmetric under $\vec{k}_2 \leftrightarrow \vec{k}_3$. Because of this it is straightforward to obtain a smoking gun of the distortion effect: Take the antisymmetric combination of Eq. (5.69) under $\vec{k}_1 \leftrightarrow \vec{k}_3$ and consider the squeezed limit where $k_1 \approx k_3 \gg k_2$. We postpone explicit calculations using SPT as our model for the nonlinear dynamics of CDM to Section 5.6, and for now we provide some intuition on this observable. In the presence of a long-wavelength velocity potential $\phi(\vec{K})$, the (local) ν-CDM cross correlation picks up a dipole,

$$\langle\delta_\nu(\vec{k}_1)\delta(\vec{k}_2)\rangle_{\phi(\vec{K})}\sim\frac{\vec{k}_1\cdot\vec{K}}{k_1}P^{\rm MON}_{\delta\delta_\nu}(k_1)\phi(\vec{K})\,, \tag{5.70}$$

due to neutrinos clustering behind the moving CDM, where $\vec{K}=\vec{k}_1+\vec{k}_2$.[12] This is qualitatively similar to the picture in Fig. 5.1 for the clustering of neutrinos around a moving point mass halo. Due to isotropy of the background, the CDM bulk velocity points in different directions in disparate regions of space so that the dipole averages out to zero, and the net effect in the two-point function is a small suppression in cross power as we pointed out below Eq. (5.67). On the other hand, the presence of a long-wavelength mode $\vec{K}$ provides a preferred direction which allows the signal to be extracted from three-point correlations.

Additionally, note from Eq. (5.70) that the dipole is proportional to the ν-CDM monopole which is related to the CDM power spectrum by a known scale-dependent transfer function.

[10]Note that these are only functions of the magnitude of the wave vectors since the angles between them are constrained by the requirement that they fit into a triangle configuration, i.e., $\vec{k}_1+\vec{k}_2+\vec{k}_3=0$. For example, $\hat{k}_1\cdot\hat{k}_2=(k_3^2-k_1^2-k_2^2)/2k_1k_2$.

[11]A prime in correlation functions indicates that the ever present factor of $(2\pi)^3\delta^{(3)}(\sum\vec{k}_i)$ is to be omitted.

[12]Indeed, this is the signature obtained in previous studies that investigated this effect using simplified models at scales below the velocity coherence scale, in two different but related languages: Dynamical friction [168] and large-scale relative flow [171].

This implies that a measurement of the distortion effect is not limited by cosmic variance since one can reconstruct the potential $\phi(\vec{K})$ from measurements of both the monopole and dipole, regardless of the specific realization of the small scale CDM power. In the next section we will see that a multi-tracer approach involving measurements of both the matter field and galaxies can be applied to extract this signal since galaxies only trace the CDM, while the matter field traces neutrinos as well. This benefits from cosmic variance cancellation, which makes it possible for this faint signal of neutrino masses to be potentially detectable in upcoming surveys.

5.6 Numerical calculations of large-scale structure observables

In the previous section we extracted the distortion in the neutrino density field, due to the peculiar motion of halos, from the Boltzmann Eq. (5.7). Our approach assumes a rather general CDM distribution on large scales while remaining agnostic about the details of the nonlinear gravitational evolution. We were able to determine the size of the dynamical friction effect and potential observational signatures on the large scale structure. However, it will prove useful to assume specific models for the nonlinear structure formation in order to carry out numerical calculations of interest. Specifically, in Sec. 5.6.1 we use the halo model to derive how halos of different masses contribute to the average in Eqs. (5.58) and (5.65). This will enable us to reproduce the results of Sec. 5.4 [specifically Eq. (5.27)] but also to derive the 2-halo term. In Sec. 5.6.2 and Sec. 5.6.3, we will apply the standard perturbation theory (SPT) to compute the power spectra and bispectra in Eq. (5.67) and (5.69) to study the observational signatures of the distortion effect in the large scale structure, and use it to forecast its observability in future surveys.

5.6.1 The 2-halo contribution to dynamical friction

We start with the halo model [39], which asserts that cold dark matter halos are the basic building blocks of CDM structure. We assume the NFW profile for the halos as suggested by numerical simulations [206, 207], and we use the expression in [208] for the halo concentration

[see Eqs. (8), (10) and (11) in this reference] which is calibrated from simulations. We adopt the convention that halos are defined as spherical regions with an average density of $\Delta = 200$ times the critical density. In order to write down the halo model expression for the power spectrum, we need two additional ingredients: the halo mass function dn/dM that gives the (comoving) number density of halos in a given halo mass window dM, and the linear halo bias $b(M)$ that fixes the correlation between nearby halos in terms of matter two-point correlations. As we will see, we will not need their explicit expressions as it will suffice to enforce the usual requirement that the mean bias of halos is unity, that is,

$$\frac{1}{\bar{\rho}_0}\int dM \frac{dn}{dM} M b(M) = 1\,, \tag{5.71}$$

where $\bar{\rho}_0$ is the comoving matter density (or matter density evaluated at $z = 0$) as in previous sections. According to the halo model:

$$\begin{aligned} P_{\delta\delta}(k) = &\int dM \frac{dn}{dM}\left(\frac{M}{\bar{\rho}_0}\right)^2 |u(k|M)|^2 + \\ &+ \int dM \frac{dn}{dM}\int dM' \frac{dn}{dM'}\frac{MM'}{\bar{\rho}_0^2} u(k|M)u(k|M')b(M)b(M')P^L_{\delta\delta}(k)\,, \end{aligned} \tag{5.72}$$

where $u(k|M)$ is the Fourier transform of the normalized NFW profile, $P^L_{\delta\delta}(k)$ is the linear theory power spectrum and time dependences are implicit. Now substituting Eq. (5.72) into Eq. (5.65) gives

$$\begin{aligned} \tau^{-1} = \frac{1}{\bar{\rho}_0}\int dM \frac{dn}{dM} M \Bigg\{ \frac{2}{3\pi} G^2 m_\nu^4 M \int d\log k \, \frac{g(k)}{\left(1+\frac{k_{\rm fs}}{k}\right)^3} \Bigg[|u(k|M)|^2 + \\ + \int dM' \frac{dn}{dM'}\frac{M'}{M} u(k|M)u(k|M')b(M)b(M')P^L_{\delta\delta}(k) \Bigg] \Bigg\}\,. \end{aligned} \tag{5.73}$$

This can be interpreted as an average over halos, from which one can read off the expression for the inverse characteristic velocity decay time for a halo of mass M:

$$\tau^{-1}(M) = \frac{2}{3\pi} G^2 m_\nu^4 M \int d\log k \, \frac{g(k)}{\left(1+\frac{k_{\rm fs}}{k}\right)^3} \left[|u(k|M)|^2 + \int dM' \frac{dn}{dM'}\frac{M'}{M} u(k|M)u(k|M')b(M)b(M')P^L_{\delta\delta}(k) \right]. \tag{5.74}$$

A comparison of Eq. (5.74) with Eq. (5.27), when ignoring the second term in the right-hand side of Eq. (5.74) which is the 2-halo term, suggests the identification

$$\log \Lambda(M) = \int d\log k \, \frac{g(k)}{\left(1+\frac{k_{\rm fs}}{k}\right)^3} |u(k|M)|^2 \,. \tag{5.75}$$

Let us stop for a moment to compare this expression with the estimate made in Sec. 5.4, i.e., $\Lambda \sim \lambda_{\rm coh}/R$ with $\lambda_{\rm coh}$ the CDM velocity coherence scale and R the halo radius. The normalized NFW profile is such that $u(k|M) \approx 1$ for $k \lesssim 1/R$ and drops to zero at smaller scales (see Fig.9 in [39]). Based on Fig. 5.4, it is also true that $g(k)/(1+k_{\rm fs}/k)^3 \approx 1$ at scales $k \gtrsim k_{\rm coh}$ and drops to zero at larger scales (assuming $k_{\rm coh} \gg k_{\rm fs}$, which is approximately true for the neutrino masses we consider). We then conclude that the integrand in Eq. (5.75) is approximately a step function that assumes the value unity between the scales $k_{\rm coh}$ and $1/R$ and drops to zero outside this range, when assuming the hierarchy of scales $k_{\rm fs} \ll k_{\rm coh} \ll 1/R$. In that case,

$$\log \Lambda(M) = \int_{k_{\rm coh}}^{1/R} \frac{dk}{k} \sim \log(1/k_{\rm coh}R) \implies \Lambda \sim \lambda_{\rm coh}/R \,, \tag{5.76}$$

which is in agreement with our previous estimate. More generally Eq. (5.75) can be well fit by a power law:

$$\Lambda(M) \approx K_\Lambda \left(\frac{10^{13} M_\odot}{M}\right)^\chi \,, \tag{5.77}$$

with $K_\Lambda = \Lambda(10^{13} M_\odot)$ a numerical coefficient that depends on the neutrino mass and redshift, and we find $\chi \approx 0.35$ to be robust against changes in the neutrino mass and redshift for the values we consider. We present numerical values for the coefficient K_Λ in Tab. 5.1. The free-streaming length $2\pi/k_{\rm fs}$, which is smaller for higher values of the neutrino mass and lower redshifts as can be seen in Fig. 5.3, acts as a cutoff to the integral in Eq. (5.75). This (mostly) sets the dependence of K_Λ on the neutrino mass and redshift.

After having carefully investigated the 1-halo term, we now proceed to study the contribution to Eq. (5.74) from the clustering of nearby halos which can be parameterized by

$$\begin{aligned} \Sigma(M) &= \frac{1}{\log\Lambda(M)} \int d\log k \, \frac{g(k)}{\left(1+\frac{k_{\rm fs}}{k}\right)^3} \int dM' \frac{dn}{dM'} \frac{M'}{M} u(k|M) u(k|M') b(M') P^L_{\delta\delta}(k) \\ &= \frac{1}{\log\Lambda(M)} \int d\log k \, I(k|M) \Gamma(k) \,, \end{aligned} \tag{5.78}$$

$K_\Lambda = \Lambda(10^{13} M_\odot)$	$m_\nu = 0.05$eV	$m_\nu = 0.1$eV	$m_\nu = 0.15$eV
$z = 0$	60	28	18
$z = 1$	79	38	24

Table 5.1: Numerical values for $K_\Lambda = \Lambda(10^{13} M_\odot)$ involved in the power-law fit of Eq. (5.77) to the 1-halo cutoff Λ.

where:

$$
\begin{aligned}
I(k|M) &= \frac{\bar{\rho}_0}{M} \frac{g(k)}{\left(1 + \frac{k_{\rm fs}}{k}\right)^3} u(k|M) P^L_{\delta\delta}(k)\,, \\
\Gamma(k) &= \frac{1}{\bar{\rho}_0} \int dM \frac{dn}{dM} M b(M) u(k|M)\,,
\end{aligned}
\tag{5.79}
$$

are dimensionless functions. Since $u(k|M) \approx 1$ at scales $k \lesssim 1/R$ while dropping to zero as $k \gg 1/R$, we conclude that $\Gamma(k)$ is unity at sufficiently large scales and drops to zero at small scales, using Eq. (5.71). The transition between the two regimes happen at the scale $k_{\rm tr} \sim 2\pi/R_{\rm peak}$, where $R_{\rm peak}$ is the radius of the halo whose mass $M_{\rm peak}$ gives the largest contribution to the integral that defines the function $\Gamma(k)$, which is in general dominated by larger halos so we choose $M_{\rm peak} \sim 10^{15} M_\odot$ with the associated $k_{\rm tr} \sim 2{\rm Mpc}^{-1}$. On the other hand, the function $I(k|M)$ peaks at scales much larger than this transition scale and hence it suffices to set $\Gamma(k) \approx 1$ in Eq. (5.78).[13] This greatly simplifies the calculation (note that it is unnecessary to introduce the halo mass function or halo bias), and it shows that dynamical friction is mostly insensitive to the specifics of the halo model parameters. We arrive at

$$
\begin{aligned}
\Sigma(M) &= \frac{1}{\log \Lambda(M)} \frac{\bar{\rho}_0}{M} \int d\log k \, \frac{g(k)}{\left(1 + \frac{k_{\rm fs}}{k}\right)^3} u(k|M) P^L_{\delta\delta}(k)\,, \\
&\approx \frac{1}{\log \Lambda(M)} \frac{\bar{\rho}_0}{M} \int d\log k \, \frac{g(k)}{\left(1 + \frac{k_{\rm fs}}{k}\right)^3} P^L_{\delta\delta}(k) \implies \Sigma(M) \approx \frac{K_\Sigma}{\log \Lambda(M)} \frac{10^{13} M_\odot}{M}\,,
\end{aligned}
\tag{5.80}
$$

[13] We verified this explicitly with the halo mass function from [209], and halo bias from [210]. Note that this approximation is analogous to setting $P^{\rm 2h}_{\delta\delta}(k) = P^L_{\delta\delta}(k)$ in the halo model and is made in, e.g., the halo reaction approach for the nonlinear power spectrum [211] in beyond ΛCDM cosmologies.

where,

$$K_\Sigma = \frac{\bar{\rho}_0}{10^{13} M_\odot} \int d\log k \, \frac{g(k)}{\left(1 + \frac{k_{\rm fs}}{k}\right)^3} P^L_{\delta\delta}(k) \,, \tag{5.81}$$

is a dimensionless number, and in order to go from the first to second line in Eq. (5.80) we used the fact that $u(k|M) \approx 1$ at scales that contribute to the integral for the range of halo masses we are considering, i.e., $(10^9 - 10^{16}) M_\odot$. In Tab. 5.2 we present numerical values for the coefficient K_Σ. It decreases with neutrino mass due to the cutoff at the free-streaming scale, and it decreases with redshift due to the linear growth of matter power.

K_Σ	$m_\nu = 0.05$eV	$m_\nu = 0.1$eV	$m_\nu = 0.15$eV
$z = 0$	60	28	17
$z = 1$	30	15	10

Table 5.2: Numerical values for the coefficient K_Σ defined in Eq. (5.81), involved in the contribution to the dynamical friction effect from the clustering of halos.

We can finally substitute Eqs. (5.75)-(5.81) into Eq. (5.74) to obtain

$$\tau^{-1}(M) = \frac{2}{3\pi} \log\Lambda(M) G^2 M m_\nu^4 [1 + b(M)\Sigma(M)] \,. \tag{5.82}$$

The function $\Sigma(M)$ gives the ratio of two to one halo contributions to the inverse characteristic time, up to the halo bias $b(M)$. From Eqs. (5.77), (5.80) and the numerical values in Tables 5.1 and 5.2, we conclude that the 2-halo term dominates for halo masses $M \lesssim 10^{14} M_\odot$ while the 1-halo term dominates for halo masses $M \gtrsim 10^{14} M_\odot$ [when assuming $b(M) \approx 1$].

The inverse characteristic time in Eq. (5.82) is plotted in Fig. 5.5 under the assumption that $b(M) = 1$ for simplicity. It asymptotically approaches a constant for small halo masses, as determined by the numerical values in Eqs. (5.65) and (5.66). This follows from a comparison between Eqs. (5.80-(5.82) and (5.65), in the regime where the 2-halo term dominates. This result can be interpreted as follows: Halos of sufficiently low mass can be thought of as surfing on the neutrino wakes produced by larger halos. On the other extreme of Fig. 5.5,

the dynamical friction experienced by larger halos is due to their own (strong) wakes leading to a more pronounced effect. This, in combination with the fact that the halo bias $b(M)$ increases at the high mass end (and can reach values as high as ~ 10) indicates that neutrino wakes can lead to a percent-level decrease in halo velocities.[14]

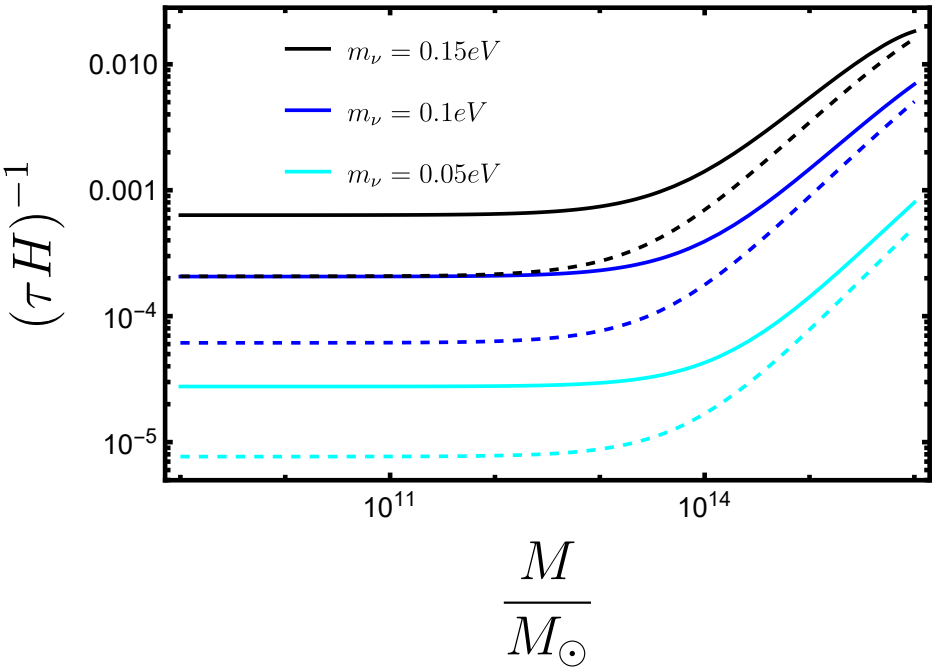

Figure 5.5: The inverse characteristic velocity decay time, due to neutrino dynamical friction, in units of the Hubble expansion rate, as given by Eq. (5.82). This also corresponds to the fractional decrease in the velocity of halos due to the dynamical friction over the age of the universe. The individual neutrino mass varies from $m_\nu = 0.15$eV to $m_\nu = 0.05$eV from top to bottom. Solid curves correspond to $z = 0$, and dashed curves to $z = 1$.

We have accomplished our goal of comparing our simplified framework in Sec. 5.4, culminating on Eq. (5.27), with the more general approach of Sec. 5.5 leading to Eq. (5.82), that solves both issues with the Chandrasekhar dynamical friction formula as previously pointed

[14]However, such high mass halos are rare and hence have limited statistics.

out: arbitrariness in the choice of a cutoff Λ, and lack of halo clustering effects.

5.6.2 Distortions to the neutrino-CDM cross power spectrum

In the previous section we have explicitly checked that the distortion in the neutrino density field due to the peculiar motion of halos, as given by the second term in the right-hand side of Eq. (5.53), is connected to the dynamical friction effect. We are now finally read to apply SPT to explore its potential signatures on the large scale structure, offering new opportunities for a cosmological measurement of the neutrino masses.

We start with the CDM-ν two-point cross-correlation, or the cross-power spectrum in Fourier space. We repeat Eq. (5.67) for convenience,

$$P_{\delta_\nu\delta}(k) = \frac{\left(\frac{k_{\rm fs}}{k}\right)^2}{(1+\frac{k_{\rm fs}}{k})^2}\left[P_{\delta\delta}(k) + \frac{\gamma}{1+\frac{k_{\rm fs}}{k}}\frac{P_{p\delta}(k)}{\sigma_\nu}\right], \tag{5.83}$$

and Eq. (5.59) for the nonlinear contribution to the CDM momentum field,

$$p(\vec{k}) = \sigma\frac{k_{\rm NL}}{k}\int\frac{d^3\vec{k}_1}{(2\pi)^3}\int\frac{d^3\vec{k}_2}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}_{12})\alpha(\vec{k}_1,\vec{k}_2)\delta(\vec{k}_1)\frac{\theta(\vec{k}_2)}{aH}\,. \tag{5.84}$$

To leading order in perturbation theory [41],

$$\begin{aligned}
\delta(\vec{k}) &= \delta_L(\vec{k}) + \int\frac{d^3\vec{k}_1}{(2\pi)^3}\int\frac{d^3\vec{k}_2}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}_{12})F_2(\vec{k}_1,\vec{k}_2)\delta_L(\vec{k}_1)\delta_L(\vec{k}_2)\,,\\
\theta(\vec{k}) &= -faH\left[\delta_L(\vec{k}) + \int\frac{d^3\vec{k}_1}{(2\pi)^3}\int\frac{d^3\vec{k}_2}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}_{12})G_2(\vec{k}_1,\vec{k}_2)\delta_L(\vec{k}_1)\delta_L(\vec{k}_2)\right],
\end{aligned} \tag{5.85}$$

with $\delta_L(\vec{k})$ the CDM linear fluctuation field, and perturbation theory kernels

$$\begin{aligned}
F_2(\vec{k}_1,\vec{k}_2) &= \frac{5}{7} + \frac{1}{2}\frac{\vec{k}_1\cdot\vec{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}\frac{(\vec{k}_1\cdot\vec{k}_2)^2}{k_1^2k_2^2}\,,\\
G_2(\vec{k}_1,\vec{k}_2) &= \frac{3}{7} + \frac{1}{2}\frac{\vec{k}_1\cdot\vec{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{4}{7}\frac{(\vec{k}_1\cdot\vec{k}_2)^2}{k_1^2k_2^2}\,.
\end{aligned} \tag{5.86}$$

Combining Eqs. (5.83)-(5.86), we arrive at the one-loop expression to the CDM-ν cross-power

spectrum[15]

$$P^{\text{1-loop}}_{\delta\delta_\nu}(k) = P^L_{\delta\delta_\nu}(k)\left[1 + \left.\frac{\Delta P^{\text{1-loop}}_{\delta\delta_\nu}(k)}{P^L_{\delta\delta_\nu}(k)}\right|_{\text{SPT}} + \left.\frac{\Delta P^{\text{1-loop}}_{\delta\delta_\nu}(k)}{P^L_{\delta\delta_\nu}(k)}\right|_{\text{Dist}}\right], \tag{5.87}$$

with,

$$P^L_{\delta\delta_\nu}(k) = \frac{\left(\frac{k_{\text{fs}}}{k}\right)^2}{(1+\frac{k_{\text{fs}}}{k})^2} P^L_{\delta\delta}(k)\,, \tag{5.88}$$

the linear theory cross power spectrum and

$$\begin{aligned}
\left.\frac{\Delta P^{\text{1-loop}}_{\delta\delta_\nu}(k)}{P^L_{\delta\delta_\nu}(k)}\right|_{\text{SPT}} &= 2\int \frac{d^3\vec{k}'}{(2\pi)^3}\left\{\left[F_2(\vec{k}', \vec{k}-\vec{k}')\right]^2 \frac{P^L_{\delta\delta}(|\vec{k}-\vec{k}'|)}{P^L_{\delta\delta}(k)} + 3F_3(\vec{k}, \vec{k}', -\vec{k}')\right\} P^L_{\delta\delta}(k')\,, \\
\left.\frac{\Delta P^{\text{1-loop}}_{\delta\delta_\nu}(k)}{P^L_{\delta\delta_\nu}(k)}\right|_{\text{Distortion}} &= -f\frac{\sigma}{\sigma_\nu}\frac{\gamma}{1+\frac{k_{\text{fs}}}{k}}\int \frac{d^3\vec{k}'}{(2\pi)^3}\frac{k_{\text{NL}}}{k'}(\hat{k}\cdot\hat{k}')\times \\
&\times\left[F_2(\vec{k}', -\vec{k})P^L_{\delta\delta}(k') + G_2(\vec{k}'-\vec{k}, \vec{k})P^L_{\delta\delta}(|\vec{k}-\vec{k}'|) + F_2(\vec{k}-\vec{k}', \vec{k}')\frac{P^L_{\delta\delta}(|\vec{k}-\vec{k}'|)P^L_{\delta\delta}(k')}{P^L_{\delta\delta}(k)}\right].
\end{aligned} \tag{5.89}$$

In Eq. (5.87), the second term in the right hand side gives the SPT one-loop correction to the power spectrum, while the third term represents the contribution from the distortion in neutrino density field due to the peculiar motion of halos, whose effect is to slightly suppress the cross power. In Fig. 5.6 we plot the one-loop expression to the CDM-ν cross power spectrum, as given by Eq. (5.87), for the reference values $m_\nu = 0.1$eV and $z = 0$. We can see that in principle one needs to properly account for the distortion effect in order to produce accurate predictions for the nonlinear corrections to the cross power spectrum at scales around $k_{\text{NL}} \approx 0.1\text{Mpc}^{-1}$, but the effect is quite small.

Note that, in principle, two-point cross correlations of CDM and ν are sensitive to the distortion effect above. However, this small signal is degenerate with the much larger contribution from standard structure formation, and hence analysis involving two-point correlations cannot isolate the distortion effect. However, we expect three-point cross correlations

[15]The one-loop power spectrum requires a next to leading order calculation in perturbation theory, i.e., we will also need the kernel $F_3(\vec{k_1}, \vec{k_2}, \vec{k_3})$. An expression for it can be found in [41].

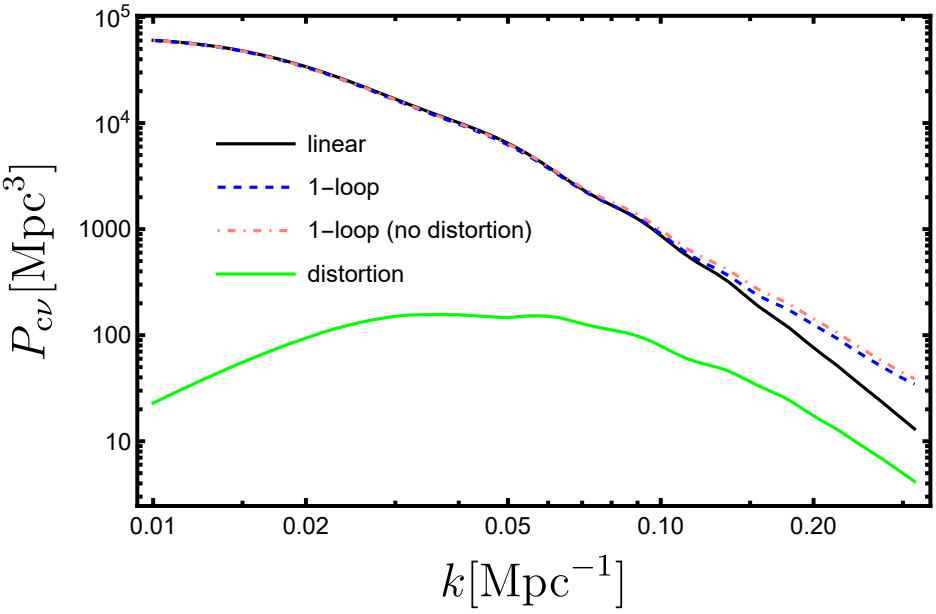

Figure 5.6: Contributions to the CDM-ν cross power spectrum, for the reference values $m_\nu = 0.1\text{eV}$ and $z = 0$. The black solid curve is the linear spectrum produced by CLASS, while the blue dashed curve is the one-loop expression from Eq. (5.87). In the pink dashed curve we also show the SPT one-loop result when not including the contribution from the distortion effect, i.e., dropping the third term in the right hand side of Eq. (5.87). Finally, the solid green curve isolates the distortion contribution to the one-loop result in blue (in absolute value, since the contribution is negative).

of CDM and ν to reveal the distortion effect. For instance, as we argued in Sec. 5.5, in the presence of a long wavelength CDM velocity field, a dipole appears in the local CDM-ν cross power. In a different language, the presence of a specific nonzero squeezed limit bispectrum involving CDM and ν correlations would be a smoking gun of the distortion effect.

5.6.3 Neutrino distortion to the bispectrum

We are interested in three-point correlations of the form $\langle\delta_\nu(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle$. Our starting point is Eq. (5.69) for the bispectrum, which we repeat here for convenience:

$$B(k_1,k_2,k_3)=\frac{\left(\frac{k_{\rm fs}}{k_1}\right)^2}{(1+\frac{k_{\rm fs}}{k_1})^2}\left[\langle\delta(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle'+\frac{\gamma}{1+\frac{k_{\rm fs}}{k_1}}\langle\frac{p(\vec{k}_1)}{\sigma_\nu}\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle'\right]. \tag{5.90}$$

This is written in terms of CDM-only bispectra that can be readily evaluated to tree level in SPT using Eqs. (5.84)-(5.86)

$$\langle\delta(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle'=2\left[F_2(\vec{k}_2,\vec{k}_3)P^L_{\delta\delta}(k_2)P^L_{\delta\delta}(k_3)+F_2(\vec{k}_1,\vec{k}_3)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_3)+F_2(\vec{k}_1,\vec{k}_2)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_2)\right], \tag{5.91}$$

and,

$$\langle\frac{q(\vec{k}_1)}{\sigma_\nu}\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle'=-f\frac{\sigma}{\sigma_\nu}\frac{k_{\rm NL}}{k_1}\left[\alpha(\vec{k}_2,\vec{k}_3)+\alpha(\vec{k}_3,\vec{k}_2)\right]P^L_{\delta\delta}(k_2)P^L_{\delta\delta}(k_3)\,. \tag{5.92}$$

We can then write Eq. (5.90) as a sum of SPT and distortion (dist) contributions,

$$B(k_1,k_2,k_3)=B_{\rm SPT}(k_1,k_2,k_3)+B_{\rm dist}(k_1,k_2,k_3)\,, \tag{5.93}$$

with,

$$\begin{aligned}B_{\rm SPT}(k_1,k_2,k_3)&=2\frac{\left(\frac{k_{\rm fs}}{k_1}\right)^2}{(1+\frac{k_{\rm fs}}{k_1})^2}\left[F_2(\vec{k}_2,\vec{k}_3)P^L_{\delta\delta}(k_2)P^L_{\delta\delta}(k_3)+F_2(\vec{k}_1,\vec{k}_3)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_3)+F_2(\vec{k}_1,\vec{k}_2)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_2\right.\\ B_{\rm dist}(k_1,k_2,k_3)&=\gamma f\frac{\sigma}{\sigma_\nu}\frac{\left(\frac{k_{\rm fs}}{k_1}\right)^2}{(1+\frac{k_{\rm fs}}{k_1})^3}\left(\frac{k_{\rm NL}}{k_3}\mu_{13}+\frac{k_{\rm NL}}{k_2}\mu_{12}\right)P^L_{\delta\delta}(k_2)P^L_{\delta\delta}(k_3)\,,\end{aligned} \tag{5.94}$$

where we used the relations $\alpha(\vec{k}_1,\vec{k}_2)=(\vec{k}_{12}\cdot\vec{k}_2)/k_2^2$ and $\vec{k}_1+\vec{k}_2+\vec{k}_3=0$, and introduce the notation $\mu_{ij}=\hat{k}_i\cdot\hat{k}_j$.

We are interested in the contribution from the distortion effect, captured by $B_{\rm dist}$, which we will see can be isolated from B_{SPT}. The distortion term, $B_{\rm dist}(k_1,k_2,k_3)$, is symmetric

under the exchange $k_2 \leftrightarrow k_3$, so we can assume without loss of generality that $k_3 \leq k_2$ and parameterize different triangle configurations by the values of k_2, $x_3 = k_3/k_2 \leq 1$ and $x_1 = k_1/k_2$. From the requirement that three wavevectors must fit into a triangle configuration, it follows that $1 - x_3 \leq x_1 \leq 1 + x_3$. In Fig. 5.7 we plot the triangle shape dependence of the bispectrum for a choice of $k_2 = 0.05\text{Mpc}^{-1}$ and our fiducial values $m_\nu = 0.1\text{eV}$ and $z = 0$. The bispectrum peaks at squeezed configurations for which $k_1 \approx k_2 \gg k_3$. The difference in sign between the $x_1 = 1 + x_3$ and $x_1 = 1 - x_3$ cases is due to the fact that the bispectrum is proportional to the angle between wavevectors and hence depends on their relative orientations. In Fig. 5.8 we plot the same bispectrum along elongated triangle configurations for which $x_1 = 1 + x_3$.

We can provide a physical interpretation for this bispectrum. For this consider the squeezed limit with $k_1 \approx k_2 \gg k_3$ in the second line of Eq. (5.94), and use Eq. (5.88). We arrive at

$$B_{\rm dist}(k_1, k_2, k_3) \approx \gamma f \frac{\sigma}{\sigma_\nu} \frac{k_{\rm NL}}{k_3} \mu_{13} \frac{P^L_{\delta\delta_\nu}(k_1)}{1 + \frac{k_{\rm fs}}{k_1}} P^L_{\delta\delta}(k_3)\,. \tag{5.95}$$

In order to interpret Eq. (5.95), note that:

$$\frac{\vec{v}(\vec{k}_3)}{\sigma_\nu} = -\frac{i\hat{k}_3}{k_3}\frac{\theta(\vec{k}_3)}{\sigma_\nu} = i\hat{k}_3 f \frac{\sigma}{\sigma_\nu}\frac{k_{\rm NL}}{k_3}\delta(\vec{k}_3)\,, \tag{5.96}$$

where we used the relation $\theta(\vec{k}_3) = -(faH)\delta(\vec{k}_3)$ and insert the identity $\sigma k_{\rm NL}/aH = 1$. We may then write,

$$B_{\rm dist}(k_1, k_2, k_3) \approx \left\langle \langle \delta_\nu(\vec{k}_1)\delta(\vec{k}_2)\rangle_{\phi(\vec{K})}\delta(\vec{k}_3)\right\rangle\,, \tag{5.97}$$

with $\vec{K} = -\vec{k}_3 = \vec{k}_1 + \vec{k}_2$ and,

$$\langle \delta_\nu(\vec{k}_1)\delta(\vec{k}_2)\rangle_{\phi(\vec{K})} = \frac{\vec{k}_1\cdot\vec{K}}{k_1}\frac{P^L_{\delta\delta_\nu}(k_1)}{1+\frac{k_{\rm fs}}{k_1}}\phi(\vec{K})\,, \tag{5.98}$$

is a dipole contribution to the local CDM-ν cross correlation in the presence of a long wavelength velocity potential $\phi(\vec{K})$ defined as, in position space,

$$\gamma\frac{\vec{v}}{\sigma_\nu} = \vec{\nabla}\phi\,. \tag{5.99}$$

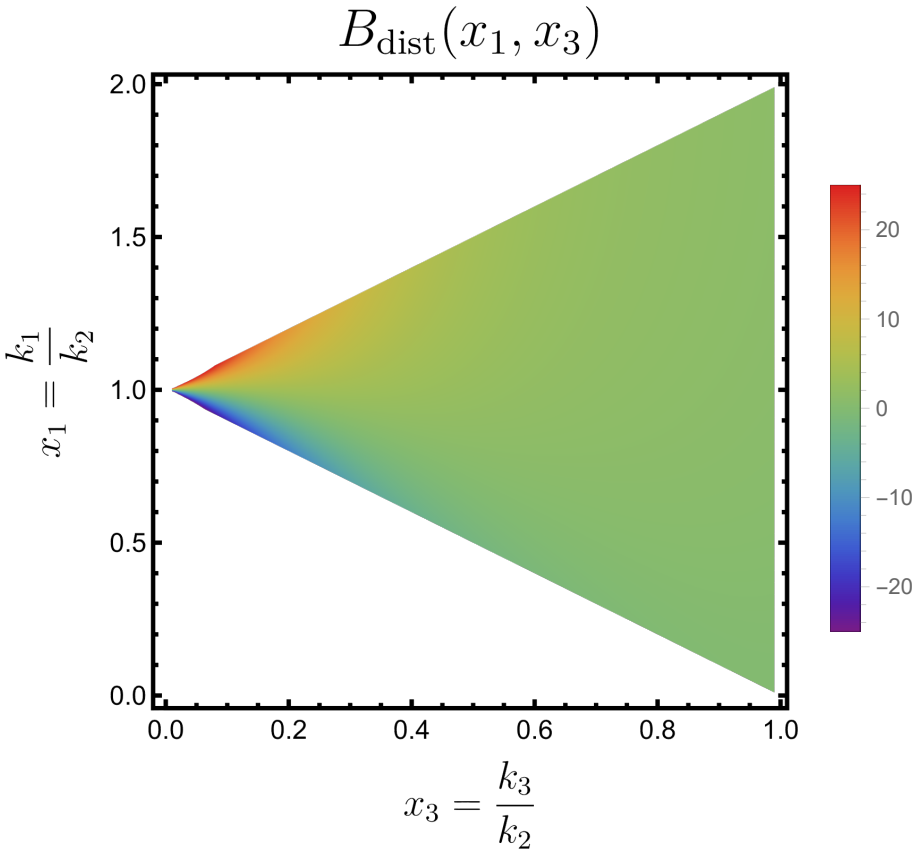

Figure 5.7: Contribution to the bispectrum from the neutrino distortion effect as given by the second line in Eq. (5.94), normalized by its value at the equilateral triangle configuration $B_{\rm dist}(x_1 = 1, x_3 = 1)$, as a function of the triangle shape parameterized by $x_1 = k_1/k_2$ and $x_3 = k_3/k_2$. We fix $k_2 = 0.05{\rm Mpc}^{-1}$, and assume our fiducial values of a single neutrino mass state with $m_\nu = 0.1{\rm eV}$ and $z = 0$. The left corner of the triangle corresponds to the squeezed limit $k_3 \ll k_1, k_2$. The upper boundary of the triangle corresponds to the "flattened" triangle configuration with $\vec{k}_3$ aligned with $\vec{k}_2$, and $\vec{k}_1$ pointing in the exact opposite direction of $\vec{k}_2$, $\vec{k}_3$. The lower boundary of the triangle corresponds to the opposite limit of $\vec{k}_1$ parallel with $\vec{k}_3$ and $\vec{k}_2$ antiparallel to $\vec{k}_1$, $\vec{k}_3$. The change in sign of $B_{\rm dist}$ along these flattened configurations in the limit $k_3 \ll k_1, k_2$ illustrates the anisotropy in the bispectrum under the exchange of $\vec{k}_1$ and $\vec{k}_2$ or equivalently, under exchanging CDM and ν density fields in $\langle \delta_\nu(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle$.

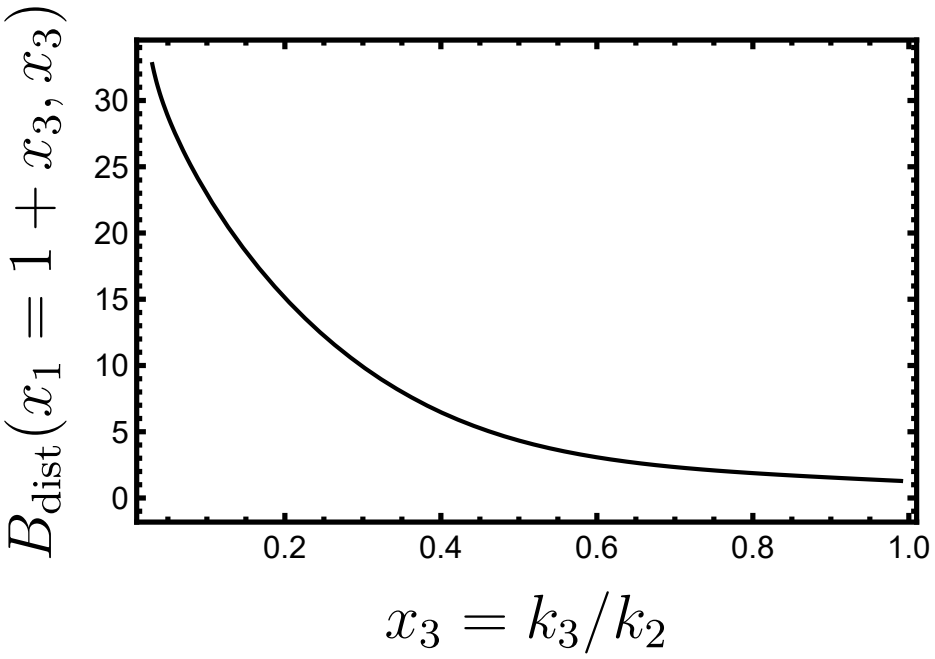

Figure 5.8: Contribution to the bispectrum from the neutrino distortion effect as given by the second line in Eq. (5.94), normalized by its value at the equilateral triangle configuration $B_{\rm dist}(x_1 = 1, x_3 = 1)$, as a function of $x_3 = k_3/k_2$ and for elongated configurations for which $x_1 = 1 + x_3$. We fix $k_2 = 0.05{\rm Mpc}^{-1}$, and assume our fiducial values of a single neutrino mass state $m_\nu = 0.1{\rm eV}$ and $z = 0$.

We expect Eq. (5.98) to be robust against higher order perturbation theory corrections, provided one applies the replacement $P^L_{\delta\delta_\nu}(k_1) \to P^{\rm MON}_{\delta\delta_\nu}(k_1) = ((\frac{k_{\rm fs}}{k})^2/(1 + \frac{k_{\rm fs}}{k})^2)P_{\delta\delta}(k_1)$ to the monopole contribution to the local CDM-ν cross power spectrum, where $P_{\delta\delta}(k_1)$ is the nonlinear CDM power spectrum.

Note from Eq. (5.98) that the dipole is proportional to both the large-scale displacement field $\phi(\vec{K})$ and the short scale monopole contribution to the cross power spectrum $P^{\rm MON}_{\delta\delta_\nu}(k_1)$. A measurement of the squeezed limit bispectrum can then be phrased as a reconstruction of the displacement field $\phi(\vec{K})$, independent of the specific realization of the small scale CDM power. As a consequence, such a measurement is not limited by cosmic variance but rather

by the finite number of galaxies in the survey as we will see in the next subsection.

Equation (5.98) is in qualitative agreement with the results obtained in [172] in a simplified model where the authors introduce an effective relative displacement between CDM and neutrino fluid elements to show that shift nonlinearities arise in the neutrino density field, which is exactly what we find in a explicit investigation of the solution to the Boltzmann equation for the distribution function of relic neutrinos around nonlinear CDM structure, i.e., from first principles. This novel approach provides some additional insight into the distortion effect and its potential observational signatures; it is now clear how the neutrino thermal velocity enters the picture, and that the dipole distortion is indeed directly connected to the dynamical friction effect. Additionally, and from a more practical point of view, we now have an explicit expression for the effective displacement potential as given by Eq. (5.99). We also have an expression for the bispectrum that goes beyond the squeezed limit as given by the second line in Eq. (5.94), and that is important for two reasons. First, it opens up the opportunity to check if the squeezed limit indeed dominates the contribution to the signal-to-noise and to account for additional information in other triangle configurations. Second, it allows us to quantitatively assess the degeneracy between the distortion contribution and standard nonlinear structure formation.

To elaborate further in the latter point, we now argue that the distortion contribution can in principle be separated from the SPT contribution. For this, the key observation is that the expression in Eq. (5.91) is symmetric under the exchange $\vec{k}_1 \leftrightarrow \vec{k}_2$, while the expression in Eq. (5.92), i.e., the contribution from the distortion effect, does not share the same symmetry. In order to exploit this, we consider (where the superscript A denotes an antisymmetrization),

$$\begin{aligned} B^A(k_1,k_2,k_3) &= \frac{1}{2}\left[B(k_1,k_2,k_3) - B(k_2,k_1,k_3)\right] , \\ &= B^A_{\rm SPT}(k_1,k_2,k_3) + B^A_{\rm dist}(k_1,k_2,k_3) , \end{aligned} \tag{5.100}$$

which once again can be written as a sum of SPT and distortion (dist) contributions. The

first reads,

$$B^A_{\rm SPT}(k_1,k_2,k_3) = \left[\frac{\left(\frac{k_{\rm fs}}{k_1}\right)^2}{(1+\frac{k_{\rm fs}}{k_1})^2} - \frac{\left(\frac{k_{\rm fs}}{k_2}\right)^2}{(1+\frac{k_{\rm fs}}{k_2})^2}\right]\Big[F_2(\vec{k}_2,\vec{k}_3)P^L_{\delta\delta}(k_2)P^L_{\delta\delta}(k_3) + F_2(\vec{k}_1,\vec{k}_3)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_3) + F_2(\vec{k}_1,\vec{k}_2)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_2)\Big]\,, \tag{5.101}$$

and the second,

$$B^A_{\rm dist}(k_1,k_2,k_3) = \frac{\gamma}{2} f \frac{\sigma}{\sigma_\nu}\Bigg\{\frac{\left(\frac{k_{\rm fs}}{k_1}\right)^2}{(1+\frac{k_{\rm fs}}{k_1})^3}\left(\frac{k_{\rm NL}}{k_3}\mu_{13}+\frac{k_{\rm NL}}{k_2}\mu_{12}\right)P^L_{\delta\delta}(k_2)P^L_{\delta\delta}(k_3) - \frac{\left(\frac{k_{\rm fs}}{k_2}\right)^2}{(1+\frac{k_{\rm fs}}{k_2})^3}\left(\frac{k_{\rm NL}}{k_3}\mu_{23}+\frac{k_{\rm NL}}{k_1}\mu_{21}\right)P^L_{\delta\delta}(k_1)P^L_{\delta\delta}(k_3)\Bigg\}\,, \tag{5.102}$$

where we combined Eqs. (5.94) and (5.100). In the squeezed limit where $k_1 \approx k_2 \gg k_3$ we then obtain $B^A_{\rm SPT}(k_1,k_2,k_3) \approx 0$ and $B^A_{\rm dist}(k_1,k_2,k_3) \approx B_{\rm dist}(k_1,k_2,k_3)$ such that the antisymmetrization procedure cleans out the SPT contribution and hence provides a smoking gun for the distortion effect. This argument assumes that all the information from the distortion effect in the bispectrum comes from the squeezed limit, which at this point is a reasonable assumption based on Figs. 5.7 and 5.8.

In the next section we extend our analysis with a calculation of the signal-to-noise ratio based on a multi-tracer approach. In order to build intuition on the size of the distortion effect, let us first compare our squeezed limit bispectrum in Eq. (5.95), which we repeat here in simpler notation for convenience,

$$B_{\rm dist} = f_\nu \gamma f \frac{\sigma}{\sigma_\nu}\frac{k_{\rm NL}}{k_L}\mu \frac{\left(\frac{k_{\rm fs}}{k_S}\right)^2}{\left(1+\frac{k_{\rm fs}}{k_S}\right)^3} P_{\delta\delta}(k_S)P_{\delta\delta}(k_L)\,, \tag{5.103}$$

with the familiar squeezed limit bispectrum (of the late-time matter density contrast) gen-

erated by the presence of local primordial non-Gaussianity [212],

$$B_{\rm local} = \frac{6\Omega_{\rm m}(a)}{g(a)}(aH)^2 f_{\rm NL}\frac{1}{k_L^2 T(k_L)}P_{\delta\delta}(k_S)P_{\delta\delta}(k_L)\,. \tag{5.104}$$

In Eq. (5.103), $k_S = k_1 \approx k_2$, $k_L = k_3$, μ is the cosine of the angle between short and long wave vectors and we used Eq. (5.52). We also inserted a factor of the fractional contribution of neutrinos to the matter density f_ν which always multiplies the neutrino density contrast δ_ν extracted from a measurement of the total mass fluctuation $\delta_m = (1-f_\nu)\delta + f_\nu\delta_\nu$. In Eq. (5.104), $\Omega_{\rm m}(a)$ is the fractional contribution of matter to the critical density, $g(a) = D_{\rm L}(a)/a$ with $D_{\rm L}(a)$ the linear growth factor and $T(k)$ is the matter transfer function. The shape and scale dependence of the two bispectra are different, but both peak at squeezed triangle configurations and we can formally compare Eqs. (5.103) and (5.104) to define an effective non-Gaussian parameter to the distortion bispectrum:

$$f_{\rm NL}^{\rm eff} = \frac{f_\nu}{\sigma_\nu^2}\gamma\frac{fg}{6\Omega_{\rm m}}\sqrt{\frac{3}{2}\Omega_{\rm m}}\left(\frac{k_L}{k_S}\right)\mu T(k_L)\frac{\left(\frac{k_{\rm fs}}{k_S}\right)^2}{\left(1+\frac{k_{\rm fs}}{k_S}\right)^3}\,, \tag{5.105}$$

where we used Eqs. (5.29) and (5.31). For a numerical estimate we evaluate this quantity at redshift $z = 0$, with $\Omega_{\rm m} = 0.32$, $f = \Omega_{\rm m}^{4/7}$ and $g = 1$. We also set $k_S = k_{\rm fs}$, $k_L/k_S = 0.1$, $\mu = 1$ and $T(k_L) \approx 1$. We the obtain $f_{\rm NL}^{\rm eff} \approx 0.5$ for an individual neutrino mass $m_\nu = 0.05$eV, $f_{\rm NL}^{\rm eff} \approx 4$ for $m_\nu = 0.1$eV and $f_{\rm NL}^{\rm eff} \approx 14$ for $m_\nu = 0.15$eV. In the case of three neutrino mass eigenstates, the net effective non-Gaussian parameter is given by the sum of the individual contributions to $f_{\rm NL}^{\rm eff}$. For comparison, the current best constraints on local non-Gaussianity from galaxy surveys are $f_{\rm NL} = -33 \pm 28$ at 95% confidence, though this comparison is not straightforward since those measurements come from joint analysis of the power spectrum and bispectrum, including scale-dependent bias and loop-corrections. Yet, near-term surveys target $\sigma(f_{\rm NL}) = 0.2 - 0.5$, with significant constraining power coming from bispectrum measurements analogous to those proposed here [213, 214]. Of course, the signal shapes and scalings of neutrinos and $f_{\rm NL}$ are different but this comparison nevertheless provides a rough estimate of the survey needs to use the bispectrum to study neutrino wakes.

5.6.4 *Multi-tracer bispectrum forecast*

In the previous section we studied the contribution to the bispectrum from the distortion effect, and used the fact that the signal peaks at squeezed triangle configurations to argue that it can be separated from the contributions to the bispectrum from standard nonlinear structure formation [the SPT term in Eq. (5.94)]. There are two more ingredients we need to account for: The intrinsic noise from cosmic variance, and the number of modes for a given triangle configuration. This will enable us to forecast the observability of the distortion effect in future surveys and to quantitatively assess how degenerate the distortion signal is with standard nonlinear structure formation.

In order to isolate the distortion contribution to the bispectrum in Eq. (5.94) we adopt the antissymetrization strategy of Eq. (5.100), as argued in the previous section. For this we need a probe of the total matter field $\delta_{\rm m} = (1-f_\nu)\delta + f_\nu\delta_\nu$ and also a galaxy data, $\delta_g = b\delta$, with b the linear bias, $f_\nu = \bar{\rho}_\nu/\bar{\rho}$ the fractional contribution of neutrinos to the matter budget. Note that we follow the standard practice of assuming the galaxy field to trace CDM and baryons, excluding neutrinos [106, 114, 115, 215–219].

It follows that, schematically

$$B^A_{\rm obs} = \frac{1}{2}\langle(\delta_{\rm m}\delta_{\rm g} - \delta_{\rm g}\delta_{\rm m})\delta_{\rm m}\rangle \;\approx \frac{1}{2}f_\nu b\langle(\delta_\nu\delta - \delta\delta_\nu)\delta\rangle\,, \tag{5.106}$$

is the bispectrum we wish to extract from the surveys, when combined with Eqs. (5.100)-(5.102), where we work to leading order in the neutrino fraction $f_\nu \ll 1$ [16]. This includes our signal of the distortion in the neutrino density field due to the peculiar motion of halos but also a background contribution from standard nonlinear structure formation. We are then left with the problem of forecasting the observability of our effect in future surveys, when also accounting for potential degeneracies with the physics of standard small scale structure formation.

[16] One could also choose $B^A_{\rm obs} = \frac{1}{2}\langle(\delta_{\rm m}\delta_{\rm g} - \delta_{\rm g}\delta_{\rm m})\delta_{\rm g}\rangle$, with a galaxy density contrast on the long-wavelength mode as opposed to a matter density contrast. To leading order in f_ν this produces the same observable (up to a factor of the linear bias b), with some additional (shot) noise due to the finite sampling of the galaxy density field.

In Appendix 5.8 we derive an optimal estimator for the bispectrum in Eq. (5.106). The following formula for the cumulative signal-to-noise ratio per galaxy (or the fisher matrix F) is obtained for our signal, which corresponds to the contribution to the bispectrum from the distortion effect as given by Eq. (5.102):

$$\frac{\mathrm{SNR}^2}{N_g} = F_{B_{\rm dist}B_{\rm dist}} = 2f_\nu^2 b^2 \int \frac{d^3\vec{k}_1}{(2\pi)^3}\int \frac{d^3\vec{k}_2}{(2\pi)^3}\int \frac{d^3\vec{k}_3}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}_{123})\frac{[B^A_{\rm dist}(k_1,k_2,k_3)]^2}{P_{\delta\delta}(k_3)[P_{\delta\delta}(k_1)+P_{\delta\delta}(k_2)+P_{\delta\delta}(k_3)]}\,, \tag{5.107}$$

where SNR^2 is the cumulative signal-to-noise ratio squared, $N_{\rm g}$ is the total number of galaxies in the galaxy survey, and $P_{\delta\delta}(k)$ stands for the (linear) power spectrum. We can now use Eq. (5.100) to rewrite this expression as follows:

$$\begin{aligned} F_{B_{\rm dist}B_{\rm dist}} = \frac{f_\nu^2 b^2}{2}\int \frac{d^3\vec{k}_1}{(2\pi)^3}\int \frac{d^3\vec{k}_2}{(2\pi)^3}\int \frac{d^3\vec{k}_3}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}_{123})\frac{B_{\rm dist}(k_1,k_2,k_3)}{P_{\delta\delta}(k_2)P_{\delta\delta}(k_3)[P_{\delta\delta}(k_1)+P_{\delta\delta}(k_2)+P_{\delta\delta}(k_3)]}\times \\ \times\left\{[P_{\delta\delta}(k_2)+P_{\delta\delta}(k_3)]\,B_{\rm dist}(k_1,k_2,k_3) - P_{\delta\delta}(k_2)B_{\rm dist}(k_2,k_1,k_3) - P_{\delta\delta}(k_3)B_{\rm dist}(k_3,k_1,k_2)\right\}\,. \end{aligned} \tag{5.108}$$

Due to the Dirac delta enforcing the requirement that the wavevectors fit into a triangle configuration, we can immediately integrate over $\vec{k}_1$ such that we are left with integrals over k_2, k_3 and $\mu_{23} = \hat{k}_2\cdot\hat{k}_3$. First, we replace the variable k_3 by $x_3 = k_3/k_2$. Since the integrand in Eq. (5.108) is symmetric under the exchange $k_2 \leftrightarrow k_3$, we can assume $x_3 \leq 1$ without loss of generality. Second, we replace μ_{23} with $x_1 = k_1/k_2$ according to:

$$\mu_{23} = \hat{k}_2\cdot\hat{k}_3 = \frac{x_1^2-1-x_3^2}{2x_3} \implies d\mu_{23} = \frac{x_1}{x_3}dx_1\,, \tag{5.109}$$

and the range of integration is $1-x_3 \leq x_1 \leq 1+x_3$. Implementing these changes in the integration variables leads to:

$$\int \frac{d^3\vec{k}_1}{(2\pi)^3}\int \frac{d^3\vec{k}_2}{(2\pi)^3}\int \frac{d^3\vec{k}_3}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}_{123}) = \frac{1}{4\pi^4}\int_0^\infty \frac{dk_2}{k_2}\int_0^1 dx_3\int_{1-x_3}^{1+x_3} dx_1\, k_2^6 x_1 x_3\,, \tag{5.110}$$

and hence to our final formula for the Fisher matrix,

$$\begin{aligned} F_{B_{\rm dist}B_{\rm dist}} = \frac{f_\nu^2 b^2}{8\pi^4}\int_0^\infty \frac{dk_2}{k_2}\int_0^1 dx_3\int_{1-x_3}^{1+x_3} dx_1\, k_2^6 x_1 x_3\frac{B_{\rm dist}(k_1,k_2,k_3)}{P_{\delta\delta}(k_2)P_{\delta\delta}(k_3)[P_{\delta\delta}(k_1)+P_{\delta\delta}(k_2)+P_{\delta\delta}(k_3)]}\times \\ \times\left\{[P_{\delta\delta}(k_2)+P_{\delta\delta}(k_3)]\,B_{\rm dist}(k_1,k_2,k_3) - P_{\delta\delta}(k_2)B_{\rm dist}(k_2,k_1,k_3) - P_{\delta\delta}(k_3)B_{\rm dist}(k_3,k_1,k_2)\right\}\,. \end{aligned} \tag{5.111}$$

We can integrate over x_1 and x_3 (the triangle shape), for a given fixed k_2, to investigate the scale dependence of the signal-to-noise. This is plotted in Fig. 5.9 for our reference values $m_\nu = 0.1\text{eV}$ and $z = 0$. It peaks at large quasi-linear scales, as set by the neutrino free-streaming scale, and hence one does not enhance the signal-to-noise by adding more modes. This means, in particular, that we are not limited by the precise modeling of complicated nonlinear and galaxy formation physics (and a linear halo bias should be sufficient). The measurement is also not degenerate with the optical depth to reionization since it is independent from the total matter power spectrum, and it is not cosmic variance limited either since the signal-to-noise can be made arbitrarily large by taking the limit $N_g \to \infty$. Indeed, cosmic variance cancellation is a common theme of multi-tracer approaches [186, 189].

To elaborate on this, let us briefly make a comparison to the more familiar case of how cosmic variance cancellation applies to measuring a linear galaxy bias factor b in $\delta_g = b\delta$ (e.g. [220]). The linear bias factor can be measured without cosmic variance because it represents the linear response of the fluctuation in the galaxy density field δ_g to a fluctuation in the matter density field δ and one only needs a single realization of δ and δ_g to determine this response b. Moreover, the particular amplitude or shape of $P_{\delta\delta}$ is, in principle, unimportant for this so long as $P_{g\delta} = bP_{\delta\delta}$ and one measures both $P_{g\delta}$ and $P_{\delta\delta}$. In our case, the distortion to the neutrino density field is a linear response of the neutrino field δ_ν to the local realization of the small-scale CDM momentum field $p = \delta v$. One similarly only needs a single realization of δv to determine this response [also see the discussion below Eq. (5.70)].

Returning to Eq. (5.111), we can also integrate over k_2 and x_1, for a given fixed x_3. This is plotted in Fig. 5.10, again for $m_\nu = 0.1\text{eV}$ and $z = 0$. It peaks at a small but nonzero value of $x_3 \sim 0.1$ due to a combination of two effects: The bispectrum prefers the $x_3 \to 0$ limit as can be seen in Figs. 5.7 and 5.8, but the number of modes is suppressed in this limit. As a consequence, real surveys need to have a sufficiently large volume in order to probe the distortion bispectrum close to the squeezed limit. For our reference value $m_\nu = 0.1\text{eV}$, we need $k_{\text{min}} \lesssim k_{\text{L}} \sim x_3 k_{\text{S}} \approx 0.005\text{Mpc}^{-1}$, from Figs. 5.9 and 5.10.

We finally proceed to evaluate Eq. (5.111) numerically, for the individual neutrino masses

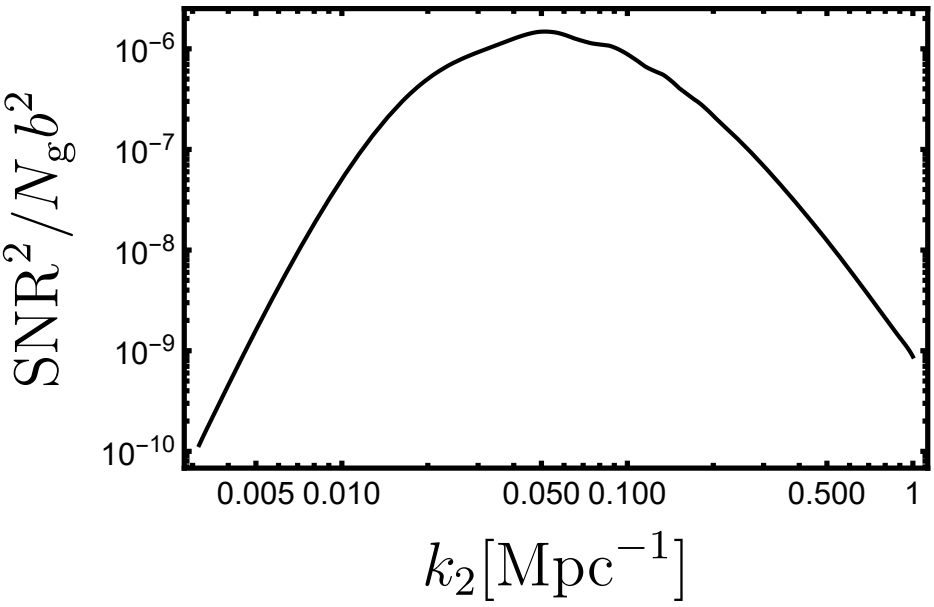

Figure 5.9: Cumulative distortion bispectrum signal-to-noise ratio squared per galaxy, divided by the galaxy bias squared, as a function of scale k_2 after integrating Eq. (5.111) over the triangle shape. We choose the reference values of a single neutrino mass state with $m_\nu = 0.1$eV and measure the bispectrum at $z = 0$. With this choice, consistency with oscillation data requires three neutrino mass states each with $m_{\nu i} \approx 0.1$eV and the net signal-to-noise squared would increase by roughly a factor of three.

and redshifts of interest, in order to forecast the observability of the effect in future surveys. The results are presented in Tab. 5.3. At $z = 0$ we find that $\mathrm{SNR}^2/b^2 N_g \approx (2.5\times10^{-9} - 1.6\times 10^{-6})$ for individual neutrino masses in the range $m_\nu = (0.05 - 0.15)$eV. [17] This corresponds to (setting $b = 1$ for the galaxy bias) a range of $N_g \gtrsim (4\times10^9 - 6\times10^6)$ for the minimal number of galaxies in the survey in order to reach a 3σ detection of the distortion effect. The signal-to-noise increases with the neutrino mass and decreases with redshift, as expected from our intuition from the dynamical friction effect [see Eq(5.66)]. However, we should keep

[17] The total signal-to-noise for all three neutrino mass eingenstates can then be obtained by summing the individual contributions in quadrature, i.e., $SNR^2_{\rm tot} = \sum_{m_\nu} SNR^2_{\rm tot}|_{m_\nu}$.

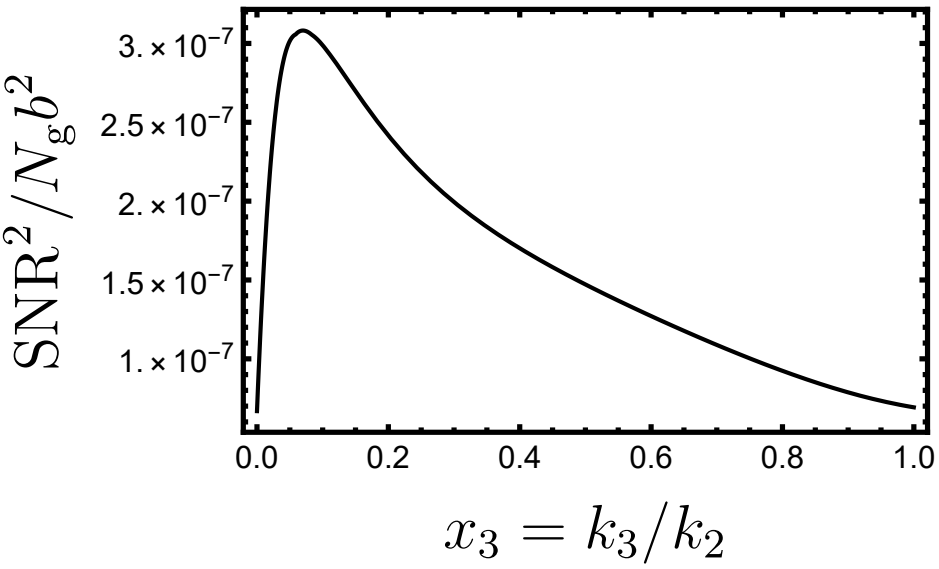

Figure 5.10: Cumulative distortion bispectrum signal-to-noise ratio squared per galaxy, divided by the galaxy bias squared, as a function of the long-wavelength mode $x_3 = k_3/k_2$ after integrating Eq. (5.111) over k_2 and $x_1 = k_1/k_3$. We choose the reference values of a single neutrino mass state with $m_\nu = 0.1$eV and measure the bispectrum at $z = 0$. With this choice, consistency with oscillation data requires three neutrino mass states each with $m_{\nu i} \approx 0.1$eV and the net signal-to-noise squared would be increased by roughly a factor of three.

in mind that we work under the assumption that $k_{\rm fs}/k_{\rm NL} \ll 1$ (see Fig. 5.3). For a sufficiently large neutrino mass $k_{\rm fs} \sim k_{\rm NL}$, and hence we expect the signal-to-noise to eventually decrease due to the cutoff at the free-streaming scale.

When accounting for the cosmological upper bound $\sum_\nu m_\nu \leq 0.12$eV, together with oscillation experiments, we are led to the optimistic scenario of two neutrino eigenstates with mass $m_\nu \approx 0.05$eV under the inverted hierarchy. This yields a total number of galaxies $N_g \approx 8 \times 10^8$ in order to reach a 2-σ detection, which may be possible with future surveys. This specific choice of neutrino mass splittings produces the highest possible signal-to-noise ratio while saturating the current cosmological upper bound in the sum of mass eigenstates.

SNR^2/b^2N_g	$m_\nu = 0.05\mathrm{eV}$	$m_\nu = 0.1\mathrm{eV}$	$m_\nu = 0.15\mathrm{eV}$
$z = 0$	2.5×10^{-9}	1.6×10^{-7}	1.6×10^{-6}
$z = 1$	8.6×10^{-11}	6.3×10^{-9}	6.8×10^{-8}

Table 5.3: Numerical values for the cumulative distortion bispectrum signal-to-noise ratio squared per galaxy, divided by the galaxy bias squared, for the (individual) neutrino masses and redshifts of interest. For three massive neutrino states, the net signal-to-noise squared is roughly the sum of the signal-to-noise squared for each state.

Our estimates for the signal-to-noise ratio in Tab. 5.3 are more pessimistic than previous estimates found in the literature for the observability of the CDM-ν relative flow effect in the large-scale structure [168, 171, 172]. We attribute this to the fact that previous approaches to study the effect of CDM-ν dynamical friction/relative flow are based on the introduction of an ad-hoc effective displacement between CDM and ν fluids, and hence do not fully account for the cutoff at the velocity coherence and free-streaming scales. Our first-principles approach, based on an explicit extraction of the distortion effect from the solution to the Boltzmann equation for the neutrino distribution function in the background of the nonlinear CDM structure, does not suffer from this problem. The cutoffs at both the velocity coherence and free-streaming scale are naturally accounted for. Yet, the presence of these cutoffs reduces the amplitude of the signal.

We previously made the observation that standard nonlinear structure formation also gives a contribution to the bispectrum of interest, as given by Eq. (5.101). We also argued that the antisymmetrization effectively separates it from the distortion contribution, provided that all the information comes from the squeezed limit $x_3 \to 0$. However, the signal-to-noise ratio peaks at some small but nonzero value of x_3, as indicated in Fig. 5.10. One may then worry that this can lead to a significant degeneracy between the distortion and SPT

contributions to the bispectrum, effectively erasing our signal. To quantitatively investigate this possibility, we first proceed to define a Fisher matrix $F_{B_{\rm SPT}B_{\rm SPT}}$ exactly as before in Eq. (5.107), by simply replacing $B_{\rm dist} \to B_{\rm SPT}$. The next step is to introduce:

$$F_{B_{\rm dist}B_{\rm SPT}} = 2f_\nu^2 b^2 \int \frac{d^3\vec{k}_1}{(2\pi)^3} \int \frac{d^3\vec{k}_2}{(2\pi)^3} \int \frac{d^3\vec{k}_3}{(2\pi)^3} (2\pi)^3 \delta^{(3)}(\vec{k}_{123}) \frac{B_{\rm dist}(k_1,k_2,k_3) B_{\rm SPT}(k_1,k_2,k_3)}{P_{\delta\delta}(k_3)[P_{\delta\delta}(k_1)+P_{\delta\delta}(k_2)+P_{\delta\delta}(k_3)]}\,, \tag{5.112}$$

in terms of which we can define the cosine of an angle between the two bispectra as follows,

$$\mu = \frac{F_{B_{\rm dist}B_{\rm SPT}}}{\sqrt{F_{B_{\rm dist}B_{\rm dist}} F_{B_{\rm SPT}B_{\rm SPT}}}} \tag{5.113}$$

with $|\mu| \leq 1$, $|\mu| = 1$ corresponding to the case of perfect degeneracy and $|\mu| = 0$ to no degeneracy between the two bispectra. For all individual neutrino masses and redshifts of interest, we find $|\mu| \leq 0.23$ when choosing $k_{\rm max} = 0.3{\rm Mpc}^{-1}$, further reducing to $|\mu| \leq 0.13$ for $k_{\rm max} = 1{\rm Mpc}^{-1}$. Of course, such a choice of maximum wavenumber actually requires a better handle into nonlinear structure formation and galaxy formation physics in order to accurately model the nonlinear dynamics. However, a naive application of SPT combined with a simple linear galaxy bias already illustrates the point that one can best disentangle the signal from the background by pushing to smaller scales. This is to be expected since these two effects mostly operate at different scales: The distortion contribution peaks at large quasi-linear scales (see Fig. 5.9), while we expect the contributions from standard nonlinear structure formation to show up at smaller nonlinear scales. However, even for a choice of $k_{\rm max} = 0.3{\rm Mpc}^{-1}$ the degeneracy is already small. This is because the distortion signal is dominated by triangle configurations that are close to the squeezed limit (see Fig. 5.10), while we expect the contribution from standard nonlinear structure formation to be suppressed in the squeezed limit [221]. We expect similar arguments to also hold for the contributions to the bispectrum from local galaxy formation physics, or, for instance higher-order terms in a bias expansion (e.g. [44]), which would justify our use of a simple linear bias. On the other hand, these terms would add additional sources of stochasticity to the noise, analogous to higher-order terms we have already dropped in the calculation of the covariance matrix in Appendix 5.8. We leave detailed investigations of these considerations to future work.

5.7 Conclusion

At late times relic neutrinos become nonrelativistic and cluster anisotropically behind moving cold dark matter structures, which generates a distortion in the neutrino density field. This effect gets imprinted into cosmological observables and can potentially be detected with upcoming surveys via its signature on three-point cross correlations of matter and galaxies.

In Sec. 5.4 we first considered a CDM distribution consisting of a single moving point mass halo, which leads to the anisotropic clustering of neutrinos as shown in Fig. 5.1. This in turn produces a dynamical friction effect that slows the halo down according to Chandrasekhar's formula, Eq. (5.24).

In Sec. 5.5 we moved on to a more general (and hence realistic) nonlinear cold dark matter (CDM) distribution, relying solely on the validity of the continuity and Poisson's equation, and hence remaining agnostic about the nonlinear gravitational evolution. We showed how to extract the distortion in the neutrino density field produced by the peculiar motion of halos from the solution to the Boltzmann equation for the neutrino distribution function, as given by Eqs. (5.51) and (5.53). This enabled us to greatly improve on, and solve a few problems with, our previous framework based on a single moving point mass halo. The consistency of our approach required a hierarchy of scales between the neutrino free-streaming scale $k_{\rm fs}$, and the scale of nonlinearities $k_{\rm NL}$, i.e., $k_{\rm fs} \ll k_{\rm NL}$, which is consistent with current cosmological upper bounds on the neutrino mass scale as illustrated in Fig. 5.3.

We determined that the average decrease of halo peculiar velocities due to the dynamical friction effect is small and hence future experiments based on measurements of galaxy velocities will likely remain insensitive to the distortion effect in the foreseeable future [see Eq. (5.66)]. However, the distortion effect also gets imprinted in the large scale structure, and one has to account for it in order to accurately model the nonlinear corrections to the CDM-ν cross power spectrum, see Eq. (5.67). We also argued that a clean probe of the distortion effect is to be found in three-point cross correlations of CDM and neutrinos, or its Fourier transform the bispectrum, as given by Eq. (5.69). We emphasize that Eqs. (5.67)

and (5.69) are agnostic about the nonlinear dynamics of CDM and can be evaluated with any given model for the CDM power spectrum and bispectrum. This fact can also be used to make inferences about neutrino wakes from data that are also agnostic about the nonlinear dynamics of CDM, that is, they will depend only on the snapshot of CDM field in a given region.

In Sec. 5.6 we considered simplified models for the nonlinear gravitational evolution in order to make some numerical calculations, building upon the framework developed in Sec. 5.5. The halo model is applied to extend the results of Sec. 5.4 based on a single moving halo, now accounting for the clustering of nearby halos and its contribution to the dynamical friction effect. We calculate the decrease in halo peculiar velocities as a function of halo mass, see Fig. 5.5. Next we introduce 1-loop standard perturbation theory (SPT) to extract the distortion contribution to the CDM-ν cross power spectrum, as shown in Fig. 5.6. We find it to be a small decrease in power, which is also degenerate with standard nonlinear structure formation. Finally, we compute the bispectrum to tree-level in SPT. The contribution from the distortion effect peaks at squeezed triangle configurations, see Figs. 5.7 and 5.8. This has a simple physical interpretation: The local CDM-ν cross power acquires a dipole in the presence of a long-wavelength displacement potential, as given by Eq. (5.98). We show that this signal can be extracted using three-point cross correlations of matter and galaxies, i.e., in a multi-tracer approach.

In Eq. (5.98) the dipole is proportional to the monopole and hence one can extract the displacement potential from the ratio between the bispectrum and the power spectrum [172], independently of the specific realization of the small scale CDM power. This implies that the signal is not limited by cosmic variance or potential inaccuracies in the modeling of complicated nonlinear structure and galaxy formation physics (this can also be seen from the fact that the signal-to-noise ratio peaks at large quasi-linear scales as illustrated in Fig. 5.9). This signature of neutrino masses on the large-scale structure has some other desirable features; it is not degenerate with the optical depth to reionization, and is also model independent in the sense that a dynamical dark energy component cannot reproduce

the effect.[18] Finally, it is not directly proportional to the neutrino mass. The dynamical friction effect and distortion contributions to the bispectrum have an approximate scaling of $\sim m_\nu^4$ and $\sim m_\nu^3$ for sufficiently small neutrino masses,[19] respectively. These effects are hence dominated by the most massive neutrino eigenstate (see Eq. (5.66) and Tab. 5.3). This could, in principle, provide a cosmological window into the neutrino mass splittings [98, 222, 223].

In practice, a detection of the distortion in the neutrino density field due the peculiar motion of halos will be challenging, and ultimately depends on the neutrino mass scale that is actually realized in nature. We estimate $N_g \gtrsim (4\times10^9 - 6\times10^6)$ [for individual neutrino masses in the range $m_\nu = (0.05-0.15)$eV] for the minimum number of galaxies in the survey in order to reach a 3σ detection (in the optimistic scenario where all the galaxies sit at low redshifts $z \approx 0$, see Tab. 5.3). In the optimistic scenario of two neutrino eigenstates with mass $m_\nu \approx 0.05$eV under the inverted hierarchy, which saturates the cosmological upper bound $\sum_\nu m_\nu \leq 0.12$eV and is consistent with oscillation experiments, a total number of galaxies $N_g \approx 8\times10^8$ is required in order to reach a 2-σ detection. For reference, the Baryon Oscillation spectroscopic survey (BOSS) completed spectroscopy on $N_g \sim 10^6$ galaxies [224], and the ongoing Dark Energy Spectroscopic Instrument (DESI) will reach $N_g \sim 4\times10^7$ [225]. A futuristic spectroscopy survey may reach $N_g \sim 10^9$ [226]. Also, a more realistic forecast will have to take into account that the matter field is only probed indirectly via the gravitational deflection it causes on the trajectory of light rays, i.e., via weak lensing. Hence, what we actually probe is the projected matter field along the line-of-sight. We can then consider photometric surveys since the exact knowledge about the redshift of galaxies is no longer essential. In one hand we expect this leads to a degradation on the signal due to the mixing of scales in projection, while on the other hand it allows for a much larger number of galaxies. For example, the Dark Energy Survey catalogs include hundreds of millions

[18] Here we assume a standard dynamical dark energy component that changes the expansion history, and hence the linear growth of structure, but does not cluster.

[19] These approximate scalings assume the hierarchy $k_{\rm fs} \ll k_{\rm NL}$. For a neutrino mass sufficiently large such that $k_{\rm fs} \sim k_{\rm NL}$, the free-streaming scales starts acting as a cutoff which reduces the amplitude of distortion effects, see Eq. (5.66) and comments below.

of galaxies [227, 228] while the upcoming Vera C. Rubin Observatory (LSST) will extract photometric redshifts of $N_g \sim 4 \times 10^9$ galaxies at $z \lesssim 1$ [229]. It would be interesting to do a detailed exploration of prospects for constraining this signal with photometric surveys, we leave this for future work. In summary, it appears that future cosmological surveys have the statistical power to detect the distortion effect and further investigation is warranted.

It would be interesting to directly extract the distortion signal in the bispectrum from N-body simulations with massive neutrinos, and compare to our calculations [184]. In the theory side, we plan on improving our modeling to relax the necessity of a hierarchy of scales, $k_{\rm fs} \ll k_{\rm NL}$. This requires one to assume a specific model for the nonlinear dynamics of CDM, and we speculate that an approach along the lines of the Renormalized Perturbation Theory (RPT) [230] can provide the necessary ingredients to account for the effects of CDM bulk flows in the neutrino field to all orders in perturbation theory. Also, it seems likely that our analysis can be extended to a more general warm dark matter component, as all we had to assume is that the background distribution functions peaks at some finite characteristic momentum (the temperature). For instance, it has been previously suggested that the CDM-baryon relative velocity leads to a unique signature in the galaxy bispectrum [231–233]. We leave an investigation of the connection between these two effects to later work.

5.8 Appendix: Signal-to-noise derivation

We carry out the derivation of the optimal estimator for the bispectrum in Eq. (5.106),

$$\frac{1}{2}\langle[\delta_{\rm m}(\vec{k}_1)\delta_{\rm g}(\vec{k}_2) - \delta_{\rm g}(\vec{k}_1)\delta_{\rm m}(\vec{k}_2)]\delta_{\rm m}(\vec{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)B^A_{\rm obs}(k_1,k_2,k_3)\,, \quad (5.114)$$

following the approach in [234]. Here $\delta^{(3)}(\vec{k})$ stands for the Dirac delta. Let us start with the most general three-point estimator for the bispectrum amplitude:

$$\begin{aligned}\hat{\mathcal{E}} &= \frac{1}{2VF}\int\frac{d^3\vec{k}_1}{(2\pi)^3}\int\frac{d^3\vec{k}_2}{(2\pi)^3}\int\frac{d^3\vec{k}_3}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}_{123})W(k_1,k_2,k_3)\left[\delta_{\rm m}(\vec{k_1})\delta_{\rm g}(\vec{k_2}) - \delta_{\rm g}(\vec{k_1})\delta_{\rm m}(\vec{k_2})\right]\delta_{\rm m}(\vec{k}_3)\\ &= \frac{1}{VF}\int_{\{\vec{k}\}}\delta(\vec{k}_{123})W(k_1,k_2,k_3)\delta_{\rm m}(\vec{k_1})\delta_{\rm g}(\vec{k_2})\delta_{\rm m}(\vec{k}_3)\,,\end{aligned} \quad (5.115)$$

with V the survey volume, and $W(k_1, k_2, k_3)$ a weight function to be found in such a way as to minimize the noise. We take it to satisfy the condition $W(k_1, k_2, k_3) = -W(k_2, k_1, k_3)$ without loss of generality, since a symmetric part would give a vanishing contribution to the estimator. We also introduced, for simplicity of notation:

$$\int_{\{\vec{k}\}} \delta(\vec{k}_{123}) \equiv \int \frac{d^3\vec{k}_1}{(2\pi)^3} \int \frac{d^3\vec{k}_2}{(2\pi)^3} \int \frac{d^3\vec{k}_3}{(2\pi)^3} (2\pi)^3 \delta^{(3)}(\vec{k}_{123}) \,. \tag{5.116}$$

Finally, the Fisher matrix F is a normalization coefficient that we fix by requiring that $\langle \hat{\mathcal{E}} \rangle = 1$, which implies

$$F = \int_{\{\vec{k}\}} \delta(\vec{k}_{123}) W(k_1, k_2, k_3) B^A_{\rm obs}(k_1, k_2, k_3) \,, \tag{5.117}$$

after taking the expectation value of Eq. (5.115), using Eq. (5.114), and letting $(2\pi)^3\delta^{(3)}(0) \to V$ as usual. We now proceed to compute the leading order Gaussian contribution to the noise $\langle \hat{\mathcal{E}}^2 \rangle$, which effectively means only keeping terms in the resulting six-point function that can be written as a product of three two-point functions, i.e., the PPP contributions. We obtain after a straightforward calculation:

$$\begin{aligned} \langle \hat{\mathcal{E}}^2 \rangle = \frac{1}{VF^2} \int_{\{\vec{k}\}} \delta(\vec{k}_{123}) W(k_1, k_2, k_3) \Big\{ & W(k_1, k_2, k_3) \left[P_{\rm mm}(k_1) P_{\rm gg}(k_2) - P_{\rm mg}(k_1) P_{\rm mg}(k_2) \right] P_{\rm mm}(k_3) + \\ & + W(k_1, k_3, k_2) \left[P_{\rm mm}(k_1) P_{\rm mg}(k_3) - P_{\rm mg}(k_1) P_{\rm mm}(k_3) \right] P_{\rm mg}(k_2) + \\ & + W(k_3, k_2, k_1) \left[P_{\rm gg}(k_2) P_{\rm mm}(k_3) - P_{\rm mg}(k_2) P_{\rm mg}(k_3) \right] P_{\rm mm}(k_1) \Big\} \,. \end{aligned} \tag{5.118}$$

We now make use of,

$$\begin{aligned} P_{\rm mm}(k) &\approx P_{\delta\delta}(k) \,, \\ P_{\rm gg}(k) &\approx b^2 P_{\delta\delta}(k) + \frac{1}{n_{\rm g}} \,, \\ P_{\rm mg}(k) &\approx b P_{\delta\delta}(k) \,, \end{aligned} \tag{5.119}$$

where we work to leading order in the fractional contribution of neutrinos to the energy density $f_\nu \ll 1$, $P_{\delta\delta}(k)$ the power spectrum including a shot-noise contribution to the power

spectrum of galaxies, $1/n_{\rm g}$, with $n_{\rm g}$ the number density of galaxies in the survey, and for simplicity we assume a constant bias factor b [20]. We arrive at

$$\langle\hat{\mathcal{E}}^2\rangle = \frac{1}{2N_g F^2}\int_{\{\vec{k}\}}\delta(\vec{k}_{123})W(k_1,k_2,k_3)\Big\{\left[W(k_1,k_2,k_3)+W(k_3,k_2,k_1)\right]P_{\delta\delta}(k_1)P_{\delta\delta}(k_3)- \\ -\left[W(k_2,k_1,k_3)+W(k_3,k_1,k_2)\right]P_{\delta\delta}(k_2)P_{\delta\delta}(k_3)\Big\}\,, \tag{5.120}$$

where $N_{\rm g}=n_{\rm g}V$ the total number of galaxies in the survey volume V. The Eq. (5.120) can be conveniently written as follows

$$\langle\hat{\mathcal{E}}^2\rangle = \frac{1}{N_{\rm g}VF^2}\int_{\{\vec{k}\}}\delta(\vec{k}_{123})\int_{\{\vec{k}'\}}\delta(\vec{k}'_{123})W(k_1,k_2,k_3)\mathrm{Cov}(\vec{k}_1,\vec{k}_2,\vec{k}_3;\vec{k}'_1,\vec{k}'_2,\vec{k}'_3)W(k'_1,k'_2,k'_3)\,, \tag{5.121}$$

in terms of a covariance matrix:

$$\mathrm{Cov}(\vec{k}_1,\vec{k}_2,\vec{k}_3;\vec{k}'_1,\vec{k}'_2,\vec{k}'_3) = \frac{1}{2}\Big[\delta(\vec{k}'_1-\vec{k}_1)\delta(\vec{k}'_2-\vec{k}_2)\delta(\vec{k}'_3-\vec{k}_3)P_{\delta\delta}(k_1)P_{\delta\delta}(k_3)+ \\ +\delta(\vec{k}'_1-\vec{k}_3)\delta(\vec{k}'_2-\vec{k}_2)\delta(\vec{k}'_3-\vec{k}_1)P_{\delta\delta}(k_1)P_{\delta\delta}(k_3)-\delta(\vec{k}'_1-\vec{k}_2)\delta(\vec{k}'_2-\vec{k}_1)\delta(\vec{k}'_3-\vec{k}_3)P_{\delta\delta}(k_2)P_{\delta\delta}(k_3)- \\ -\delta(\vec{k}'_1-\vec{k}_3)\delta(\vec{k}'_2-\vec{k}_1)\delta(\vec{k}'_3-\vec{k}_2)P_{\delta\delta}(k_2)P_{\delta\delta}(k_3)\Big]\,. \tag{5.122}$$

We now have an optimization problem, where we would like to minimize the noise $\langle\hat{\mathcal{E}}^2\rangle$, as given by Eq. (5.121), when subject to the constraint Eq. (5.117). By taking the variation of the noise with respect to the weight function and setting it to zero, we obtain the following equation:

$$\int_{\{\vec{k}'\}}\delta(\vec{k}'_{123})\mathrm{Cov}(\vec{k}_1,\vec{k}_2,\vec{k}_3;\vec{k}'_1,\vec{k}'_2,\vec{k}'_3)W(k'_1,k'_2,k'_3) = \delta(\vec{k}_{123})B_{\rm obs}(k_1,k_2,k_3)\,, \tag{5.123}$$

up to an arbitrary normalization constant. The solution to Eq. (5.123) can be parameterized as follows,

$$W(k_1,k_2,k_3) = \int_{\{\vec{k}'\}}\delta(\vec{k}'_{123})\Sigma(\vec{k}_1,\vec{k}_2,\vec{k}_3;\vec{k}'_1,\vec{k}'_2,\vec{k}'_3)B_{\rm obs}(k'_1,k'_2,k'_3)\,, \tag{5.124}$$

[20] This assumption can be relaxed to accommodate a scale dependent bias which is more realistic at the scales we are considering, and it does not lead to qualitatively different results. We stick to a constant bias for simplicity.

in terms of a precision matrix $\Sigma(\vec{k}_1, \vec{k}_2, \vec{k}_3; \vec{k}'_1, \vec{k}'_2, \vec{k}'_3)$. The substitution of Eq. (5.124) into Eq. (5.123) now yields

$$\int_{\{\vec{k}'\}} \delta(\vec{k}'_{123}) \mathrm{Cov}(\vec{k}_1, \vec{k}_2, \vec{k}_3; \vec{k}'_1, \vec{k}'_2, \vec{k}'_3) \Sigma(\vec{k}'_1, \vec{k}'_2, \vec{k}'_3; \vec{k}''_1, \vec{k}''_2, \vec{k}''_3) = \delta(\vec{k}''_1 - \vec{k}_1)\delta(\vec{k}''_2 - \vec{k}_2)\delta(\vec{k}''_3 - \vec{k}_3)\,, \tag{5.125}$$

such that the precision matrix is the inverse of the covariance matrix of Eq. (5.122). As we will see, in practice it is more convenient to directly solve Eq. (5.123) for the weight function instead of solving for the precision matrix. The substitution of Eq. (5.122) into Eq. (5.123) leads to:

$$\begin{aligned}[P_{\delta\delta}(k_1) + P_{\delta\delta}(k_2)]P_{\delta\delta}(k_3)W(k_1, k_2, k_3) + P_{\delta\delta}(k_1)P_{\delta\delta}(k_3)W(k_3, k_2, k_1) - \\ - P_{\delta\delta}(k_2)P_{\delta\delta}(k_3)W(k_3, k_1, k_2) = 2B_{\rm obs}(k_1, k_2, k_3)\,.\end{aligned} \tag{5.126}$$

It is now straightforward to check that

$$W(k_1, k_2, k_3) = \frac{2B_{\rm obs}(k_1, k_2, k_3)}{P_{\delta\delta}(k_3)\left[P_{\delta\delta}(k_1) + P_{\delta\delta}(k_2) + P_{\delta\delta}(k_3)\right]}\,, \tag{5.127}$$

is the solution we are looking for, after using the identity

$$B_{\rm obs}(k_1, k_2, k_3) + B_{\rm obs}(k_3, k_1, k_2) + B_{\rm obs}(k_2, k_3, k_1) = 0 \tag{5.128}$$

which follows from Eq. (5.114). Alternatively, the Eq. (5.127) can be derived after considering the Eq. (5.126) in combination with the same equation when written in terms of a different permutation of the wavenumbers. We now have all the ingredients necessary to determine the cumulative signal-to-noise ratio. First note from Eqs. (5.117),(5.121) and (5.123) that

$$\mathrm{SNR}^2 = \frac{1}{\langle \hat{\mathcal{E}}^2 \rangle} = N_{\rm g} F \implies F = \frac{\mathrm{SNR}^2}{N_{\rm g}}\,, \tag{5.129}$$

is given by, when substituting the Eq. (5.127) into Eq. (5.117):

$$F = 2 \int \frac{d^3\vec{k}_1}{(2\pi)^3} \int \frac{d^3\vec{k}_2}{(2\pi)^3} \int \frac{d^3\vec{k}_3}{(2\pi)^3} (2\pi)^3 \delta^{(3)}(\vec{k}_{123}) \frac{\left[B_{\rm obs}(k_1, k_2, k_3)\right]^2}{P_{\delta\delta}(k_3)\left[P_{\delta\delta}(k_1) + P_{\delta\delta}(k_2) + P_{\delta\delta}(k_3)\right]}\,, \tag{5.130}$$

when also writing the integration variables in full, according to Eq. (5.116). This, in combination with Eq. (5.129), is the formula we need for the cumulative signal-to-noise ratio.

Chapter 6

COSMOLOGICAL PERTURBATION THEORY FOR LARGE SCALE STRUCTURE IN PHASE SPACE

6.1 Abstract

We develop a framework for Large Scale Structure (LSS) perturbation theory, that solves the Vlasov-Poisson system of equations for the distribution function in full phase space. This approach relaxes the usual *apriori* assumption of negligible velocity dispersion underlying the Standard Perturbation Theory (SPT). We apply the new method to rederive the usual SPT kernels up to third order in the perturbative expansion. We also show that a counterterm, identical to the one introduced by standard Effective Field Theory (EFT) methods, naturally arises within our framework. We finish by making a precise connection to EFT techniques, which reveals the necessity of the EFTofLSS to self-consistently model the long-wavelength fluid, and illustrates the importance of having theoretical control over short distance fluctuations.

6.2 Introduction

Perturbation theory methods for the evolution of large-scale structure in the universe play a central role in our understanding of gravitational instability in the nonlinear regime [41, 133, 235–238]. In one hand, they provide flexibility to go beyond standard scenarios [239–243] and valuable intuition on the nonlinear gravitational evolution. They also do not require expensive computational resources, contrary to simulations [37]. On the other hand, perturbation theory methods have a limited range of applicability as N-body simulations are required to accurately model the dynamics on sufficiently small scales [244].

The collective dynamics of N-body particles in an expanding background, interacting

solely via gravity, is encoded in the collisionless Boltzmann (or Vlasov) equation in phase space coupled to the Poisson equation, the Vlasov-Poisson system. The standard perturbation theory (SPT) approach is then based on a truncation of the Boltzmann hierarchy in its first two moments, which corresponds to the assumption of an ideal pressureless fluid [41]. An alternative but equivalent approach to model the dynamics of late time cosmological fluctuations perturbatively is Lagrangian Perturbation Theory (LPT), where one solves for the displacement field connecting the initial (Lagrangian) and final (Eulerian) particle positions [134, 245, 246] [1].

Over the years many improvements to SPT and LPT have been proposed, and some of them are now a central piece of the theory modeling involved in the analysis of real data [52–54, 248–261]. Such improvements can be broadly divided into two classes. The first corresponds to the set of tools that do not attempt to describe an imperfect fluid, but are rather based on a reorganization of the perturbative expansion and resummations of certain classes of diagrams to all orders in perturbation theory. Examples of such methods are Renormalized Perturbartion Theory (RPT) [230, 262–265] and Infrared Resummation (IR) schemes [48, 51, 266–272].

In this work our goal will be to shed light on the second class of improvements to SPT, i.e. those that accommodate the inevitable deviations from a pressureless fluid that arise during nonlinear structure formation. In this class the leading framework is the Effective Field Theory of Large Scale Structure (EFTofLSS) [42, 43, 273–278], which also addresses the issue of sensitivity to uncontrolled short distance physics [2], and has been very successful in pushing the regime of validity of perturbation theory methods towards smaller scales [279–281]. It does so, however, at the cost of introducing new nuisance parameters to the theoretical model.

We will be focusing on the simplest case of perturbation theory to one-loop, describing the

[1] It can be shown that LPT matches SPT order by order in the perturbative expansion (e.g. [247]). See [50] for a thorough comparison between these two approaches.

[2] As can be seen from the fact that modes with arbitrarily high frequencies are running on the loops in SPT.

two-point clustering of matter in real space (as opposed to redshift space). In that case only a single new parameter is needed, the effective sound speed, which can be determined either by observations or through matching to N-body simulations. Additional free parameters become necessary when considering higher order terms in the perturbative expansion [282], when modeling biased tracers [44] and for higher point correlation functions [283–285] as well.

On the other hand, gravity-only N-body simulations have no free parameters, and we are entering a new era where efficient emulators are available to interpolate the predictions from simulations in broad regions of parameter (and even theory) space [46, 286–295]. This motivates the search for a theoretical framework which does not introduce new nuisance parameters. The starting point of SPT is the assumption of vanishing velocity dispersion (which corresponds to an ideal fluid), and this leaves open the possibility that one can account for the dissipative terms in a fully perturbative framework that truncates the Boltzmann hierarchy at a higher than second moment [296–306].

In what follows we will pursue the question of how far one can go with old-fashioned cosmological perturbation theory methods, without using Effective Field Theory (EFT) ingredients, in terms of accurately predicting observables of interest. To accomplish this we will develop a framework to perturbatively solve the Vlasov-Poisson system of equations directly in phase space, expanding upon ideas first introduced in [307]. This framework circumvents the need to artificially truncate the Boltzmann hierarchy and hence relaxes the usual *apriori* assumptions of negligible vorticity and velocity dispersion (see [308, 309] for other previous approaches to nonlinear structure formation in phase space).

We will first use this framework to rederive the familiar SPT kernels. This exercise underscores the fact that a vanishing vorticity and velocity dispersion should be seen as a consequence of the perturbative expansion, rather than an assumption (in accordance with the results obtained in [307]). This is consistent with the expectation that these effects are of intrinsically nonperturbative nature. An important outcome of our formalism will be the fact that nonlinearities backreact into the background distribution function. That is, even

if our starting point is a background distribution function consistent with the assumption of cold dark matter (proportional to a Dirac delta function at zero momentum), gravitational nonlinearities will introduce some additional time-dependent contributions. In fact, we will see that the Boltzmann equation can be rephrased as a coupled set of equations, one for the background distribution function and another for its fluctuations.

This suggests a natural procedure to improve on SPT: To use the (*apriori* unknown) fully nonlinear background distribution function as a source term to solve for the fluctuation in the distribution function perturbatively. This effectively enables one to account for a nonzero average velocity dispersion, and introduces an additional contribution to the one-loop power spectrum with the exact same form as the effective sound speed counterterm in the EFTofLSS. While this reveals the possibility to arrive at the right ingredients from a purely old-fashioned perturbation theory approach, we will show that the framework is necessarily incomplete for reasons that are related to the lack of theoretical control over short distance fluctuations. This problem can be cured by standard EFT methods, which emerge as a necessary ingredient to self-consistently model a nonzero velocity dispersion.

We will be focusing on the minimal scenario of a ΛCDM universe, particularly its late time dynamics (redshifts $z \lesssim 100$) on subhorizon scales $k \gg aH$ (where $a(t)$ is the cosmological scale factor, $H = d\log a/dt$ is the Hubble expansion rate and t is cosmic time) [3]. All numerical calculations assume a fiducial ΛCDM cosmology with $\Omega_{\mathrm{m},0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $h = 0.7$, where $H_0 = 100h\mathrm{km/s/Mpc}$ is the Hubble expansion rate today.

We structure this paper as follows: In Sec.6.3 we review the framework underlying Standard Perturbation Theory (SPT), to establish notation and for later comparison with our new perturbation theory scheme based directly on the Vlasov-Poisson system of equations in phase space, developed in Sec.6.4. In Sec.6.5 we first show how an EFT-like counterterm naturally emerges within our framework, and then proceed to make the connection to EFT methods more precise, which underscores the necessity of the EFTofLSS to self-consistently

[3] We do however wish to consider time scales that go beyond a Hubble time $t \gtrsim t_H \sim 1/H$.

model a nonzero velocity dispersion. Our main results are summarized in Sec.6.6. Additionally, Appendices 6.7, 6.8 and 6.9 derive important equations which are used in the main text. Appendices 6.10 and 6.11 contain explicit formulas for the one-loop power spectrum and tree-level bispectrum in SPT, to aid the reader with numerical calculations involving the full time dependencies of perturbation theory kernels in ΛCDM.

6.3 Standard Perturbation Theory

The collective behavior of particles interacting only gravitationally in an expanding universe is governed by the collisionless Boltzmann, or Vlasov, equation

$$\frac{\partial f}{\partial \eta} + \frac{d\vec{x}}{d\eta} \cdot \frac{\partial f}{\partial \vec{x}} + \frac{d\vec{q}}{d\eta} \cdot \frac{\partial f}{\partial \vec{q}} = 0 \,, \tag{6.1}$$

for the phase space distribution function $f(\eta, \vec{x}, \vec{q})$. We work with the superconformal time defined by $d\eta = dt/a^2(t)$. Additionally, $\vec{x}$ are comoving coordinates and $\vec{q}$ is the comoving momentum such that $(d\vec{x}/d\eta) = a^2(d\vec{x}/dt) = (\vec{q}/m)$, with m the particle mass. We then have $(d\vec{q}/d\eta) = -ma^2\vec{\nabla}\phi$, with $\phi(\eta, \vec{x})$ the gravitational potential. Eq. (6.1) now reads,

$$\frac{\partial f}{\partial \eta} + \frac{\vec{q}}{m} \cdot \frac{\partial f}{\partial \vec{x}} - ma^2(\eta) \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} = 0 \,. \tag{6.2}$$

Next we take moments of Eq. (6.2) to arrive at fluid equations, following standard procedure [41]. First define the energy density

$$\rho(\eta, \vec{x}) = ma^{-3}(\eta) \int \frac{d^3\vec{q}}{(2\pi)^3} f(\eta, \vec{x}, \vec{q}) \,. \tag{6.3}$$

The derivative of Eq. (6.3) with respect to η, which we here denote by a prime, yields using Eq. (6.2)

$$\rho' + 3\mathcal{H}\rho + a\vec{\nabla} \cdot \vec{\Pi} = 0 \,, \tag{6.4}$$

where $\mathcal{H} = d \log a/d\eta = a^2 H$ and

$$\vec{\Pi}(\eta, \vec{x}) = ma^{-4}(\eta) \int \frac{d^3\vec{q}}{(2\pi)^3} \, \vec{q} \, f(\eta, \vec{x}, \vec{q}) \,, \tag{6.5}$$

is the fluid momentum. We once again take a derivative of Eq. (6.5) with respect to super-conformal time, and use Eq. (6.2) to arrive at

$$\Pi_i' + 4\mathcal{H}\Pi_i + 2a\partial^j K_{ij} + a\rho\partial_i\phi = 0\,, \tag{6.6}$$

when written in terms of a kinetic energy density tensor

$$K_{ij}(\eta,\vec{x}) = \frac{1}{2m}a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}\,q_i q_j\,f(\eta,\vec{x},\vec{q})\,. \tag{6.7}$$

Eqs. (6.4) and (6.6) are called the first and second moments of the Boltzmann equation, corresponding to the continuity and Euler equations respectively. Note that these moments do not depend explicitly on the particle mass m, which in fact can be absorbed into suitable redefinitions of the comoving momentum $\vec{q}$ and distribution function f as follows: $\vec{q}\to m\vec{q}$ and $f\to f/m^4$. For this reason we simply set $m=1$ moving forward.

Both Eqs. (6.4) and (6.6) pick-up corrections that are sizable on horizon scales and for large thermal velocities that can appear, for example, in warm dark matter models. For instance, note that the energy density in Eq. (6.3) is in reality a mass density since it only includes the rest mass contribution to the particle proper energy. Such corrections are negligible for cold dark matter and at sub-horizon scales.

In principle we can continue by taking the derivative of Eq. (6.7) with respect to super-conformal time, to arrive at an equation of motion for the second moment, involving the third moment of the Boltzmann equation. Repeating this procedure indefinitely generates coupled equations of motion for even higher moments, the so-called Boltzmann hierarchy. Instead, the starting point of SPT is to truncate the resulting Boltzmann hierarchy at its second moment. To see how this works let us first introduce the field velocity $\vec{v}(\eta,\vec{x})$ as follows,

$$\vec{\Pi}(\eta,\vec{x}) = \rho(\eta,\vec{x})\vec{v}(\eta,\vec{x})\,, \tag{6.8}$$

in terms of which the Euler Eq. (6.6) becomes

$$v_i' + \mathcal{H}v_i + av_j\partial^j v_i + a\partial_i\phi + \frac{a}{\rho}\partial^j\tau_{ij} = 0\,, \tag{6.9}$$

after substituting Eq. (6.8) into Eq. (6.6) and dividing by the energy density. This involves the stress tensor

$$\begin{aligned}\tau_{ij}(\eta,\vec{x}) &= 2K_{ij}(\eta,\vec{x}) - \rho(\eta,\vec{x})v_i(\eta,\vec{x})v_j(\eta,\vec{x}) \\ &= a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}\left[q_i - a(\eta)v_i(\eta,\vec{x})\right]\left[q_j - a(\eta)v_j(\eta,\vec{x})\right]f(\eta,\vec{x},\vec{q})\,.\end{aligned} \tag{6.10}$$

Note from its definition in the second line of Eq. (6.10) that the stress tensor is sourced by velocity dispersion at a fixed comoving position (dispersion with respect to averaging over momentum at a given point in configuration space), and hence can only be nonvanishing at the onset of shell-crossing when particle trajectories intersect.

In order to turn Eqs. (6.4) and (6.9) into a closed system, we first need to introduce the Poisson equation satisfied by the gravitational potential:

$$\nabla^2\phi = 4\pi G a^2(\rho - \bar{\rho})\,, \tag{6.11}$$

where only the fluctuations around the average density $\bar{\rho}(\eta) = \langle\rho(\eta,\vec{x})\rangle \propto a^{-3}(\eta)$ contribute to the gravitational potential. Eq. (6.11) also picks-up corrections on large scales and for large thermal velocities, but they are negligible for cold dark matter on sub-horizon scales.

One next proceeds with the standard assumption of a negligible stress tensor: $\tau_{ij} \approx 0$. Under this assumption the velocity field is fully specified by its divergence, i.e., the vorticity degrees of freedom are negligible. To see why that is, define $w(\eta,\vec{x}) = \vec{\nabla}\times\vec{v}(\eta,\vec{x})$, and take the curl of Eq. (6.9) to obtain

$$w_i' + \mathcal{H}w_i - a[\vec{\nabla}\times(\vec{v}\times\vec{w})]_i = -a\epsilon_i^{jk}\partial_j\left(\frac{1}{\rho}\partial^l\tau_{kl}\right)\,. \tag{6.12}$$

The source term in the right-hand side of Eq. (6.12) vanishes when $\tau_{ij} \approx 0$, such that vorticity can be neglected if it is not present in the initial conditions. This latter assumption is justified since the vorticity decays with the expansion of the universe in linear perturbation theory [41]. We can even relax the assumption of $\tau_{ij} \approx 0$, and consider a diagonal stress tensor of the form $\tau_{ij} = p\delta_{i,j}$, where $p = p(\eta,\vec{x})$ is the pressure. In this case

$$-\epsilon_i^{jk}\partial_j\left(\frac{1}{\rho}\partial^l\tau_{kl}\right) = \left(\frac{\vec{\nabla}\rho}{\rho^2}\times\vec{\nabla}p\right)_i\,, \tag{6.13}$$

which vanishes whenever the pressure is an arbitrary function of the density, $p = p(\rho)$, as in an adiabatic fluid. We then once again reach the conclusion that vorticity can be neglected [310].

Once both the stress tensor and vorticity are neglected, the continuity and Euler Eqs. (6.4) and (6.9) can be solved perturbatively in fluctuations around the homogeneous background, see [41]. SPT accurately describes the clustering of matter in single-stream regions (often denoted by voids) where particle trajectories do not intersect. In such regions the stress tensor vanishes. However, at sufficiently small scales the gravitational evolution becomes strongly coupled, leading to shell-crossing and the subsequent formation of bound structures via gravitational collapse where particle trajectories do intersect [133, 311]. This is signaled by the emergence of a nonzero stress tensor.

While on large scales isotropy is a good approximate symmetry of the perturbations, the local process of gravitational collapse is anisotropic and happens at different rates along different axes, as determined by the engenvectors of the tidal tensor $\partial_i \partial_j \phi$ [4]. The gravitational collapse then proceeds in a hierarchical triaxial way: First into cosmic sheets, followed by cosmic filaments until the remaining axis finally collapses and dark matter halos form. The outcome is an intricate cosmic web where dark matter halos can be found within filaments which themselves can be found within sheets [313–317].

Cosmic sheets, filaments and halos all correspond to multi-stream regions where the stress tensor does not vanish [318], and hence SPT breaks down. Note that although $\tau_{ij} = 0$ is an *apriori* assumption of the standard perturbative framework, we expect the emergence of a nonzero stress to be an intrinsically nonperturbative phenomena in nonlinear gravitational evolution [5]. Indeed, in the next section we will see that the Vlasov-Poisson system of Eqs. (6.2) and (6.11) can be solved perturbatively in full phase space, allowing us to relax the *apriori* assumption of a vanishing stress tensor.

---

[4] The recent paper [312] argues for an energy shear tensor criteria.

[5] A nonzero vorticity is also generated by nonperturbative effects [319, 320], and we can think of it as contributing to the effective stress tensor.

6.4 Cosmological perturbations in phase space

In this section we will develop a framework to directly solve the Vlasov-Poisson system of equations perturbatively in full phase space. It will then become clear that a negligible stress is a consequence of the perturbative expansion, rather than an assumption. We will also show that small scale nonlinearities backreact into the background distribution function, which enables one to naturally account for a nonzero average velocity dispersion by introducing the unknown fully nonlinear background distribution function into the formalism. This adds a new term to the nonlinear power spectrum which has the exact same form as the effective sound speed counterterm in the EFTofLSS.

6.4.1 An iterative solution to Vlasov-Poisson

We first present a derivation of SPT based on a perturbative solution to the Vlasov-Poisson system of equations in full phase space, extending upon the work of [307]. Let us repeat here for convenience the collisionless Boltzmann (or Vlasov) Eq. (6.2)

$$\frac{\partial f}{\partial \eta} + \vec{q}\cdot\frac{\partial f}{\partial \vec{x}} = a^2(\eta)\frac{\partial \phi}{\partial \vec{x}}\cdot\frac{\partial f}{\partial \vec{q}}\,, \tag{6.14}$$

where we set $m = 1$[6], and moved the nonlinear term to the right-hand side for future convenience. At first we will remain agnostic about what is sourcing the gravitational potential $\phi(\eta, \vec{x})$, so we will delay writing down the Poisson Eq. (6.11).

Next we split Eq. (6.14) into a coupled set of equations, one for the background distribution function $\bar{f}(\eta, q) = \langle f(\eta, \vec{x}, \vec{q})\rangle$ defined by an ensemble average [7], and another for its fluctuations $\delta f = f - \bar{f}$. This procedure is not strictly necessary as Eq. (6.14) can be solved perturbatively as is, but it will prove useful for later developments. The ensemble average of Eq. (6.14) reads

$$\frac{\partial \bar{f}}{\partial \eta} = a^2(\eta)\Big\langle\frac{\partial \phi}{\partial \vec{x}}\cdot\frac{\partial f}{\partial \vec{q}}\Big\rangle\,, \tag{6.15}$$

[6] A justification for this choice can be found in Sec. 6.3, in the discussion below Eq. (6.7).

[7] This can also be thought of as a volume average. A precise operational definition will be given later in this section.

where we used the fact that $\bar{f}(\eta, q)$ is position independent to drop the term proportional to its spatial gradient [8]. Note that the right-hand side of Eq. (6.15) includes the ensemble average of a term quadratic in fluctuations and is hence nonvanishing. For this reason small scale nonlinearities will backreact into the background distribution function. Subtracting Eq. (6.15) from Eq. (6.14) produces the equation for the fluctuation to the distribution function $\delta f(\eta, \vec{x}, \vec{q})$:

$$\frac{\partial \delta f}{\partial \eta} + \vec{q} \cdot \frac{\partial \delta f}{\partial \vec{x}} = a^2(\eta) : \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} : , \tag{6.16}$$

where we introduced the normal ordering symbol as subtracting ensemble averages, that is,

$$: \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} : = \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} - \left\langle \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} \right\rangle . \tag{6.17}$$

Since $f = \bar{f} + \delta f$ appears on the right-hand side of Eqs. (6.15) and (6.16) these are a coupled set of equations, which we will now write in their integral forms. For Eq. (6.15) this is straightforward and follows from an integration over superconformal time

$$\bar{f}(\eta, q) = \bar{f}^{(0)}(q) + \int_0^\eta d\eta' a^2(\eta') \left\langle \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} \right\rangle \Big|_{\eta'} , \tag{6.18}$$

where $\bar{f}^{(0)}(q)$ is the background distribution function before picking up nonlinear corrections. For a cold dark matter species, $\bar{f}^{(0)}(q) \propto \delta^{(3)}(\vec{q})$.

In Fourier space [9] Eq. (6.16) becomes a first order ODE and it is then a straightforward exercise to rephrase it as an integral equation:

$$\delta f(\eta, \vec{k}, \vec{q}) = \int_0^\eta d\eta' a^2(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta - \eta')} \left[: \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} : \right] \Big|_{\eta', \vec{k}} , \tag{6.19}$$

where [10]

$$\left[: \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} : \right] \Big|_{\eta', \vec{k}} = \int \frac{d^3\vec{k}_1}{(2\pi)^3} \frac{d^3\vec{k}_2}{(2\pi)^3} (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}_1 - \vec{k}_2) : \phi(\eta', \vec{k}_1) i\vec{k}_1 \cdot \frac{\partial f}{\partial \vec{q}} \Big|_{\eta', \vec{k}_2, \vec{q}} : , \tag{6.20}$$

---

[8] Or thinking in terms of a volume average, that term becomes a total derivative and hence leads to a surface contribution, which we assume vanishes at spatial infinity with suitable boundary conditions.

[9] Let $\vec{\nabla} \to i\vec{k}$.

[10] Note that we also set $\delta f(\eta = 0, \vec{k}, \vec{q}) = 0$. This is justified here since the initial conditions only play a role at horizon scales while at the subhorizon scales of interest the source term completely dominates.

denotes a convolution. These are the ingredients we need to start solving this coupled set of Boltzmann equations in phase space.

We are now ready to consider an iterative solution to Eqs. (6.18) and (6.19) in the form of

$$\begin{aligned} \bar{f}(\eta, q) &= \bar{f}^{(0\text{th})}(\eta, q) + \bar{f}^{(1\text{st})}(\eta, q) + \bar{f}^{(2\text{nd})}(\eta, q) + \cdots \\ \delta f(\eta, \vec{k}, \vec{q}) &= \delta f^{(0\text{th})}(\eta, \vec{k}, \vec{q}) + \delta f^{(1\text{st})}(\eta, \vec{k}, \vec{q}) + \delta f^{(2\text{nd})}(\eta, \vec{k}, \vec{q}) + \cdots , \end{aligned} \tag{6.21}$$

for a given external gravitational potential $\phi(\eta, \vec{k})$. This can be represented diagrammatically using circles connected by a horizontal line, with the number of circles denoting the order in the iterative expansion. Additionally, we use a solid horizontal line to denote a term in the iterative expansion for the fluctuation in the distribution function, and a dashed horizontal line is used for the background distribution function.

For example, the zeroth order term for the background distribution function, $\bar{f}^{(0\text{th})}(\eta, q)$, corresponds to the diagram shown in Fig.6.1. On the other hand, the third order term in the iterative expansion for the distribution function fluctuation, $\delta f^{(3\text{rd})}(\eta, \vec{k}, \vec{q})$, is represented by the diagram drawn in Fig.6.2.

Figure 6.1: Diagram for the zeroth order term in the iterative expansion for the background distribution function.

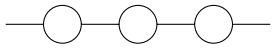

Figure 6.2: Diagram for the third order term in the iterative expansion for the fluctuation in the distribution function.

An iterative solution of this form was considered previously in [183] (without the background-fluctuation split) in the context of massive neutrinos, and can be interpreted as a reconstruction of particle trajectories in powers of the gradient of the gravitational potential in units

of the Hubble scale, i.e., $\sim \nabla^2\phi/H^2$ is the effective expansion parameter (we will further comment on this at the end of this subsection). To zeroth order we simply have,

$$\begin{aligned} \bar{f}^{(0\text{th})}(\eta, q) &= \bar{f}^{(0)}(q) \\ \delta f^{(0\text{th})}(\eta, \vec{k}, \vec{q}) &= 0\,, \end{aligned} \tag{6.22}$$

corresponding to the unperturbed trajectory. We now substitute this zeroth order solution into the right-hand side of Eqs. (6.18) and (6.19) to obtain the first order solutions. To evaluate the ensemble averages all one needs to know are the statistical properties of the external gravitational potential, which we take as a given. We arrive at,

$$\begin{aligned} &\bar{f}^{(1\text{st})}(\eta, q) = 0\,, \\ &\delta f^{(1\text{st})}(\eta, \vec{k}, \vec{q}) = i\vec{k}\cdot\frac{\partial \bar{f}^{(0)}}{\partial \vec{q}}\int_0^\eta d\eta' a^2(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}\phi(\eta', \vec{k})\,, \end{aligned} \tag{6.23}$$

where we assume $\langle\phi(\eta, \vec{k})\rangle = 0$. We can now repeat this procedure and substitute Eq. (6.23) into the right-hand side of Eqs. (6.18) and (6.19). Let us first stop for a moment to introduce notation which will be used throughout the manuscript,

$$\int \frac{d^3\vec{k}_1}{(2\pi)^3}\cdots\frac{d^3\vec{k}_N}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\vec{k}-\vec{k}_1-\cdots-\vec{k}_N) \equiv \int^{\vec{k}}_{\{\vec{k}_1,\cdots,\vec{k}_N\}}\,, \tag{6.24}$$

and in terms of which the second order solutions are

$$\begin{aligned} \bar{f}^{(2\text{nd})}(\eta, q) &= \int_0^\eta d\eta' a^2(\eta')\int^{\vec{k}}_{\{\vec{k}_1,\vec{k}_2\}}\langle\phi(\eta', \vec{k}_1) i\vec{k}_1\cdot\frac{\partial\delta f^{(1\text{st})}}{\partial\vec{q}}\Big|_{\eta',\vec{k}_2,\vec{q}}\rangle \\ \delta f^{(2\text{nd})}(\eta, \vec{k}, \vec{q}) &= \int_0^\eta d\eta' a^2(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}\int^{\vec{k}}_{\{\vec{k}_1,\vec{k}_2\}}:\phi(\eta', \vec{k}_1) i\vec{k}_1\cdot\frac{\partial\delta f^{(1\text{st})}}{\partial\vec{q}}\Big|_{\eta',\vec{k}_2,\vec{q}}:\,, \end{aligned} \tag{6.25}$$

after plugging in the second line of Eq. (6.23) into Eq. (6.25). Here we do want to go up to third order since this is required for a one-loop calculation of the power spectrum. For that we substitute Eq. (6.25) into the right-hand side of Eqs. (6.18) and (6.19), which involves some additional terms when compared to the second order solution because $f^{(2\text{nd})} = \bar{f}^{(2\text{nd})} + \delta f^{(2\text{nd})}$ is the quantity that appears as a source, and now $\bar{f}^{(2\text{nd})} \neq 0$ (as opposed to $\bar{f}^{(1\text{st})} = 0$ so that

the background term does not contribute at second order). We arrive at,

$$\begin{aligned}\bar{f}^{(3\mathrm{rd})}(\eta,q) &= \int_0^\eta d\eta' a^2(\eta') \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} \langle \phi(\eta',\vec{k}_1) i\vec{k}_1 \cdot \frac{\partial \bar{f}^{(2\mathrm{nd})}}{\partial \vec{q}}\Big|_{\eta',\vec{k}_2,\vec{q}} \rangle \\ &+ \int_0^\eta d\eta' a^2(\eta') \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} \langle \phi(\eta',\vec{k}_1) i\vec{k}_1 \cdot \frac{\partial \delta f^{(2\mathrm{nd})}}{\partial \vec{q}}\Big|_{\eta',\vec{k}_2,\vec{q}} \rangle \,,\end{aligned} \tag{6.26}$$

and

$$\begin{aligned}\delta f^{(3\mathrm{rd})}(\eta,\vec{k},\vec{q}) &= \int_0^\eta d\eta' a^2(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')} \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} :\phi(\eta',\vec{k}_1) i\vec{k}_1 \cdot \frac{\partial \bar{f}^{(2\mathrm{nd})}}{\partial \vec{q}}\Big|_{\eta',\vec{k}_2,\vec{q}}: \\ &+ \int_0^\eta d\eta' a^2(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')} \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} :\phi(\eta',\vec{k}_1) i\vec{k}_1 \cdot \frac{\partial \delta f^{(2\mathrm{nd})}}{\partial \vec{q}}\Big|_{\eta',\vec{k}_2,\vec{q}}: \,.\end{aligned} \tag{6.27}$$

Note that it is only at third order in this expansion that the background distribution function backreacts into the fluctuations (this will be important later). The recursion relations found in Eqs. (6.26) and (6.27) can be straightforwardly generalized to higher orders in the iterative expansion, and for a known external gravitational potential this is the full story (and this is the extent to which the iterative solution was considered in [183], since one can safely ignore the backreaction of neutrino fluctuations to the total gravitational potential which is dominated by cold dark matter).

However, in practice we know that the Poisson equation relates the gravitational potential to the density field, which is itself obtained from the distribution function via a momentum integration as in Eq. (6.3). What this means is that the iterative solution we wrote down is, in reality, an integral equation that we will solve perturbatively in what follows. As mentioned previously, the effective expansion parameter is $\sim \nabla^2\phi/H^2$. From the Poisson Eq. (6.11) and the Friedmann equation

$$H^2 = \frac{8\pi G}{3}\rho_{\mathrm{cri}} = \frac{8\pi G}{3}\frac{\bar{\rho}}{\Omega_{\mathrm{m}}(a)} \,, \tag{6.28}$$

with $\Omega_{\mathrm{m}}(a) = \bar{\rho}(a)/\rho_{\mathrm{crit}}(a)$ the fractional contribution of matter to the energy budget, the effective expansion parameter is of order $\nabla^2\phi/H^2 \sim \delta = (\rho-\bar{\rho})/\bar{\rho}$ the matter density contrast. Our framework is then an old-fashioned cosmological perturbation theory scheme,

in the sense that we can expect it to share the same limitations as traditional methods associated to the fact that the density contrast becomes large at the onset of nonlinearities, signaling the breakdown of the perturbative expansion [244, 321]. In Sec.6.5 we will have more to say about the implications of this observation.

6.4.2 The perturbative expansion

When coupled to the Poisson Eq. (6.11), which we here rewrite (in Fourier space) in terms of the Friedmann Eq. (6.28) evaluated at the present time [11]

$$k^2\phi = -\frac{3}{2}\Omega_{\rm m,0}H_0^2\frac{\delta}{a}\,, \tag{6.29}$$

the iterative (formal) solution we studied before becomes an integral equation that we will here solve in a perturbative expansion. Since the matter density contrast acts as the effective expansion parameter, we look for a solution in the form:

$$\delta(\eta,\vec{k}) = \delta^{(1)}(\eta,\vec{k}) + \delta^{(2)}(\eta,\vec{k}) + \cdots\,, \tag{6.30}$$

which in view of Eq. (6.29) translates to a similar expansion for the gravitational potential,

$$\phi(\eta,\vec{k}) = \phi^{(1)}(\eta,\vec{k}) + \phi^{(2)}(\eta,\vec{k}) + \cdots\,. \tag{6.31}$$

Let us now investigate how the perturbative expansion works explicitly. To leading order the only contribution comes from a single insertion of $\phi^{(1)}$ into the expression for the first order iterative solution in Eq. (6.23). We represent this by the diagram in Fig.6.3, where the number N of wiggly lines connecting to a given circle (in this case $N=1$) determines the order $\phi^{(N)}$ of that insertion [12]. This diagram evaluates to,

$$\delta f^{(1)}(\eta,\vec{k},\vec{q}) = i\vec{k}\cdot\frac{\partial\bar{f}^{(0)}}{\partial\vec{q}}\int_0^\eta d\eta' a^2(\eta')e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}\phi^{(1)}(\eta',\vec{k})\,. \tag{6.32}$$

[11] Quantities evaluated at the present time carry a subscript 0. For example, H_0 is the present day value of the Hubble expansion rate.

[12] In general there needs to be at least one wiggly line connecting to any given circle, and the order in perturbation theory can be read from the total number of wiggly lines in a diagram.

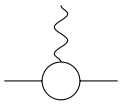

Figure 6.3: The only diagram contributing to leading order in the perturbative expansion.

Also recall that $\bar{f}^{(1\mathrm{st})}(\eta, q) = 0$ from Eq. (6.23), so the same diagram with a dashed horizontal line vanishes. Now we use the Poisson Eq. (6.29) to write this as

$$\delta f^{(1)}(\eta, \vec{k}, \vec{q}) = -\frac{i\vec{k}}{k^2} \cdot \frac{\partial \bar{f}^{(0)}}{\partial \vec{q}} \frac{3}{2} \Omega_{\mathrm{m},0} H_0^2 \int_0^\eta d\eta' a(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')} \delta^{(1)}(\eta', \vec{k}) \,. \tag{6.33}$$

Next integrate Eq. (6.33) with respect to momentum, using Eq. (6.3) and the fact that $\bar{f}^{(0)}(q) \propto \delta^{(3)}(\vec{q})$ [13], to arrive at an integral equation for $\delta^{(1)}(\eta, \vec{k})$. One obtains,

$$\delta^{(1)}(\eta, \vec{k}) = \frac{3}{2} \Omega_{\mathrm{m},0} H_0^2 \int_0^\eta d\eta' a(\eta') \delta^{(1)}(\eta', \vec{k})(\eta - \eta') \,. \tag{6.34}$$

As expected the linear theory evolution does not couple different wavenumbers, and in fact a separable solution of the form

$$\delta^{(1)}(\eta, \vec{k}) = D_{\mathrm{L}}(\eta) \delta_{\mathrm{L}}(\vec{k}) \,, \tag{6.35}$$

can be found, where $D_{\mathrm{L}}(\eta)$ is the linear growth factor (normalized to unity when evaluated today) and $\delta_{\mathrm{L}}(\vec{k})$ is the present day linear density field. The latter quantity represents the initial conditions for the matter density contrast, when rescaled to the present time under the assumption of linear evolution. It is a Gaussian stochastic random field whose power spectrum can be extracted from linear Boltzmann solvers (we use the Cosmic Linear Anisotropy Solving System, CLASS [47]). This provides a precise operational definition for the ensemble averages in our formalism, as they can always be decomposed in terms of the two-point function of the linear density field,

$$\langle \delta_{\mathrm{L}}(\vec{k}) \delta_{\mathrm{L}}(\vec{k'}) \rangle = (2\pi)^3 \delta^{(3)}(\vec{k} + \vec{k'}) P_{\mathrm{L}}(k) \,, \tag{6.36}$$

[13] This allows for a straightforward integration over momentum after an integration by parts.

with $P_{\rm L}(k)$ the linear theory power spectrum at redshift $z = 0$. As we will see, a generic term in the perturbative expansion for the density contrast, in Eq. (6.30), scales as $\delta^{(n)} \sim (\delta^{(1)})^n = (D_{\rm L}\delta_{\rm L})^n$ and so organizes itself in powers of the initial condition for the density contrast.

The evolution equation satisfied by the linear growth factor, Eq. (6.34), can be recast as a second order ODE by taking two derivatives of this equation with respect to superconformal time

$$\frac{d^2 D_{\rm L}}{d\eta^2} - \frac{3}{2}\Omega_{{\rm m},0} H_0^2 a(\eta) D_{\rm L}(\eta) = 0\,, \tag{6.37}$$

and this looks more familiar when written in terms of the scale factor,

$$\frac{d^2 D_{\rm L}}{da^2} + \frac{1}{a}\left(3 + \frac{d\log H}{d\log a}\right)\frac{dD_{\rm L}}{da} - \frac{3}{2}\Omega_{{\rm m},0} H_0^2 \frac{D_{\rm L}(a)}{a^5 H(a)^2} = 0\,. \tag{6.38}$$

In Appendix 6.7 we derive the well-known analytic solution to this equation, as a special case of the more general scenario involving the presence of a source term on the right-hand side of Eq. (6.38). There are two independent solutions to this second order ODE

$$\begin{aligned} D_{\rm L}^{+}(a) &= H(a)\int_0^a \frac{da'}{(a')^3 H^3(a')} \\ D_{\rm L}^{-}(a) &= H(a)\,. \end{aligned} \tag{6.39}$$

The mode $D_{\rm L}^{-}(a)$ decays with the expansion of the universe, and it quickly becomes negligible in comparison to the growing mode $D_{\rm L}^{+}(a)$. We then drop the decaying mode, and arrive at

$$D_{\rm L}(a) = \frac{H(a)}{H_0}\left[\int_0^1 \frac{da'}{(a')^3 H^3(a')}\right]^{-1}\int_0^a \frac{da'}{(a')^3 H^3(a')}\,, \tag{6.40}$$

where we fix the normalization by imposing $D_{\rm L}(a = 1) = 1$. We also introduce the linear growth rate, $f(a) = d\log D_{\rm L}/d\log a$. Accurate fitting functions for the numerical evaluation of these quantities can be found in [322].

To second order in perturbation theory two diagrams contribute to the distribution function fluctuation as shown in Fig. 6.4. The diagram on the left represents a single insertion of $\phi^{(2)}$ into the first order iterative solution Eq. (6.23), and the diagram on the right represents

two insertions of $\phi^{(1)}$ into the second order iterative solution Eq. (6.25). For the background distribution function at second order in perturbation theory only a single diagram, as depicted in Fig.6.5, gives a nonzero contribution.

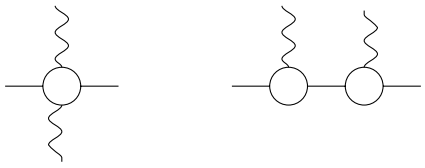

Figure 6.4: Two diagrams contribute to the distribution function fluctuation at second order in perturbation theory. The total number of wiggly lines reveals the order in perturbation theory.

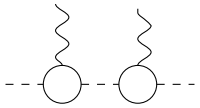

Figure 6.5: The only diagram contributing to the background distribution function at second order in perturbation theory.

In Appendix 6.8 we include detailed calculations of diagrams to second order in perturbation theory, while in the main text we focus on summarizing the main results. The diagram in Fig.6.5 evaluates to

$$\bar{f}^{(2)}(\eta, q) = -\left(\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\right)^2 \int_0^\eta d\eta' a(\eta') D_{\mathrm{L}}(\eta') \int_0^{\eta'} d\eta'' a(\eta'') D_{\mathrm{L}}(\eta'') \times \\ \times \int \frac{d^3\vec{k}'}{(2\pi)^3} \frac{P_{\mathrm{L}}(k')}{(k')^4} i\vec{k}' \cdot \frac{\partial}{\partial \vec{q}} \left[i\vec{k}' \cdot \frac{\partial \bar{f}^{(0)}}{\partial \vec{q}} e^{i\vec{k}' \cdot \vec{q}(\eta' - \eta'')} \right] . \tag{6.41}$$

Note that this is a total derivative with respect to momentum. We then see from Eq. (6.3) that this does not lead to a renormalization of the background mass density, which is the statement of particle number conservation. It is also straightforward to check, from Eqs. (6.5)

and (6.41), that the fluid momentum (and hence the fluid velocity) does not pick up a background renormalization either since $\bar{f}^{(2)}(\eta, q)$ is only a function of the absolute value of comoving momentum, due to statistical isotropy. As a consequence, any perturbative framework based solely on the density and velocity fields necessarily misses this renormalization of the background distribution function. We expect this to remain true at higher orders in the perturbative expansion since particle number conservation and statistical isotropy should hold to all orders.

That being said, it is straightforward to argue that the background distribution function needs to be renormalized even within SPT. Taking the ensemble average of the stress tensor using Eq. (6.7) and the first line of Eq. (6.10) yields

$$\tau(\eta) \equiv \langle \tau_i^i \rangle\big|_\eta = a^{-5}(\eta) \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}(\eta, q) - \langle \rho v^2 \rangle\big|_\eta \,, \tag{6.42}$$

where only the trace can be nonvanishing (upon averaging) due to statistical isotropy. Since the stress tensor vanishes in SPT by construction, the background distribution functions needs to pick up backreactions beyond $\bar{f}^{(0)}(q) \propto \delta^{(3)}(\vec{q})$ in order to cancel the contribution from $\langle \rho v^2 \rangle \neq 0$. Indeed, we show in Appendix 6.8 that Eq. (6.41) implies (where we suppress the time dependence for simplicity of notation when it is convenient to do so),

$$a^{-5} \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}(q) = \bar{\rho}\, a^2 H^2 f^2 D_{\rm L}^2 \int_0^\infty \frac{dk'}{2\pi^2} P_{\rm L}(k') \,. \tag{6.43}$$

This corresponds to the bulk flow (or to be more precise, the linear theory mean square velocity), and exactly cancels the leading contribution to $\langle \rho v^2 \rangle$ in a perturbative expansion, to produce a vanishing averaged stress tensor [14]. This is the first indication that our framework based on a perturbative solution to the Vlasov-Poisson system of equations in phase space is reproducing nothing other than SPT.

Moving on to the distribution function fluctuations, a calculation of the two diagrams in Fig.6.4 (carried out explicitly in Appendix 6.8) leads to a second order density contrast

[14] Within SPT our expectation is that higher loop corrections to the background distribution function (of which $\bar{f}^{(4)}(\eta, q)$ is the next to leading order term as we will see shortly) will exactly cancel the higher order contributions to $\langle \rho v^2 \rangle$ such that the consistency relation $\tau(\eta) = 0$ is satisfied.

which admits a decomposition into a sum of separable terms as follows

$$\delta^{(2)}(a,\vec{k}) = c_1^{(2)}(a)h_1^{(2)}(\vec{k}) + c_2^{(2)}(a)h_2^{(2)}(\vec{k})\,, \tag{6.44}$$

where the scale dependent functions $h_i^{(2)}(\vec{k})$ are given by

$$\begin{aligned} h_1^{(2)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} \alpha^{(s)}(\vec{k}_1,\vec{k}_2)\,{:}\delta_{\rm L}(\vec{k}_1)\delta_{\rm L}(\vec{k}_2){:} \\ h_2^{(2)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} \beta(\vec{k}_1,\vec{k}_2)\,{:}\delta_{\rm L}(\vec{k}_1)\delta_{\rm L}(\vec{k}_2){:}\,, \end{aligned} \tag{6.45}$$

with $\alpha(\vec{k}_1,\vec{k}_2) = (\vec{k}_1\cdot\vec{k}_{12})/k_1^2$ and $\beta(\vec{k}_1,\vec{k}_2) = k_{12}^2(\vec{k}_1\cdot\vec{k}_2)/2k_1^2k_2^2$. Here $\vec{k}_{12} = \vec{k}_1+\vec{k}_2$ and $\alpha^{(s)}(\vec{k}_1,\vec{k}_2) = [\alpha(\vec{k}_1,\vec{k}_2)+\alpha(\vec{k}_2,\vec{k}_1)]/2$ stands for the symmetric combination. We derive explicit analytic expressions for the time-dependent coefficients, $c_i^{(2)}(a)$, in Appendix 6.7.

Let us now move on to the third order in the perturbative expansion. The background distribution function picks up no third order contributions, or to any odd order in perturbation theory more generally, due to the assumption of a Gaussian linear random field. In pictorial language, all diagrams with a horizontal dashed line and an odd number of wiggly lines vanish. There are, however, four distinct diagrams contributing to the distribution function fluctuation at third order in perturbation theory, as illustrated in Fig.6.6. After evaluating these diagrams, we find that the third order density contrast can also be decomposed as a sum of separable terms [15]

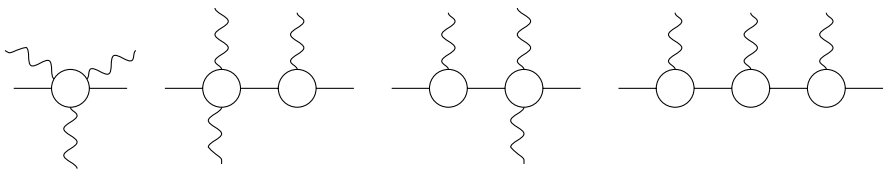

Figure 6.6: Four diagrams contribute to the distribution function fluctuation at third order in perturbation theory.

[15] One can organize such a decomposition into a proper basis of scale dependent functions [323], but we are not going to worry about this here.

$$\delta^{(3)}(a,\vec{k}) = \sum_{i=1}^{6} c_i^{(3)}(a) h_i^{(3)}(\vec{k})\,. \tag{6.46}$$

The scale dependent functions are,

$$\begin{aligned}
h_1^{(3)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2,\vec{k}_3\}}^{\vec{k}} \alpha(\vec{k}_1,\vec{k}_{23})\alpha^{(s)}(\vec{k}_2,\vec{k}_3)\,\delta_{\rm L}(\vec{k}_1)\,{:}\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3){:}\\
h_2^{(3)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2,\vec{k}_3\}}^{\vec{k}} \alpha(\vec{k}_1,\vec{k}_{23})\beta(\vec{k}_2,\vec{k}_3)\,\delta_{\rm L}(\vec{k}_1)\,{:}\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3){:}\\
h_3^{(3)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2,\vec{k}_3\}}^{\vec{k}} \alpha(\vec{k}_{23},\vec{k}_1)\alpha^{(s)}(\vec{k}_2,\vec{k}_3)\,\delta_{\rm L}(\vec{k}_1)\,{:}\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3){:}\\
h_4^{(3)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2,\vec{k}_3\}}^{\vec{k}} \alpha(\vec{k}_{23},\vec{k}_1)\beta(\vec{k}_2,\vec{k}_3)\,\delta_{\rm L}(\vec{k}_1)\,{:}\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3){:}\\
h_5^{(3)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2,\vec{k}_3\}}^{\vec{k}} \beta(\vec{k}_1,\vec{k}_{23})\alpha^{(s)}(\vec{k}_2,\vec{k}_3)\,\delta_{\rm L}(\vec{k}_1)\,{:}\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3){:}\\
h_6^{(3)}(\vec{k}) &= \int_{\{\vec{k}_1,\vec{k}_2,\vec{k}_3\}}^{\vec{k}} \beta(\vec{k}_1,\vec{k}_{23})\beta(\vec{k}_2,\vec{k}_3)\,\delta_{\rm L}(\vec{k}_1)\,{:}\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3){:}\,,
\end{aligned} \tag{6.47}$$

and analytic formulas for the time-dependent coefficients, $c_i^{(3)}(a)$, are provided in Appendix 6.8.

Our results in this section are in exact agreement with the SPT predictions for the density contrast, obtained following the traditional approach outlined in Sec.6.3 (see, e.g., [323–334])[16]. However, let us stress that at no point in our framework did we have to assume a vanishing stress tensor or vorticity. Instead, an ideal pressureless fluid follows as a consequence of the perturbative solution to the Vlasov-Poisson system of equations in phase space. This is consistent with the intuition that dissipative effects are due to intrisically nonperturbative effects in nonlinear gravitational evolution. We attribute the following physical interpretation to this result: Our perturbative expansion is based on

[16]Note that normal ordering symbols appear in Eqs. (6.45) and (6.47) for the perturbation theory kernels. However, in SPT one obtains the same kernels without any normal ordering symbols. This is inconsequential as the forbidden contractions would lead to a vanishing contribution. The normal ordering symbols then simply organize for us which of the contractions lead to a nonvanishing contribution and hence need to be evaluated in the first place.

the iterative approach developed in Sec.6.4.1, which reconstructs particle trajectories in an expansion in powers of the gradient of the gravitational field, as first showed in [183]. Now, the shell-crossing singularity occurs when particles start to accumulate in small regions of space, which require particle trajectories to turn around. This happens when the gradient of the gravitational field becomes sufficiently large, likely beyond the convergence radius of the iterative solution. This is why the perturbative expansion does not account for the shell-crossing singularity.

This result is consistent with the findings of [307], where the author offers the following explanation as to why the perturbative expansion based on the Vlasov-Poisson system reproduces SPT: At early times, before the onset of shell-crossing, the fluid description of the system holds exactly and hence must match what is obtained from the Boltzmann equation. Since the structure of the perturbative expansion at all times is set in terms of powers of the initial conditions via $\delta_{\rm L}(k)$, there is no room to accommodate for shell-crossing at late times once it first occurs. Beyond this simple argument, there are no definitive proofs that the perturbative expansion based on the Vlasov-Poisson system exactly reproduces SPT to all orders in perturbation theory [17].

The formalism we develop in this paper is an improvement over the one introduced in [307], mainly for the added clarity on how the nonlinearities in gravitational evolution renormalize the background distribution function. This feature enables us to make a connection between the cosmological perturbation theory in phase space and the EFTofLSS, which will be the subject of our attention in the next section.

We intend Sec.6.4 to be a useful pedagogical reference for numerical calculations in SPT with the full ΛCDM time-dependent kernels, up to third order in the perturbative expansion. For this reason, we supplement the main text with explicit analytic formulas for the time-dependent coefficients in Appendix 6.7, one-loop power spectrum in Appendix 6.10 and

[17] In particular, we were not able to find a symmetry argument which manifestly forbids the generation of a nonzero average stress tensor, i.e., which explains the exact cancellation of diagrams identified in Eqs. (6.42) and (6.43).

tree-level bispectrum in Appendix 6.11.

6.5 Connection to EFT methods

In Sec. 6.4 we showed that the Boltzmann equation can be recast as a coupled set of equations, one for the background distribution function, Eq. (6.15), and another for its fluctuations, Eq. (6.16). We then solved these two coupled set of equations perturbatively in full phase space, while also imposing the Poisson Eq. (6.29). We saw that the outcome of this procedure is identical to SPT, in the sense that both frameworks lead to the same density contrast at second and third order in perturbation theory.

The key difference is that our perturbative framework based on the Vlasov-Poisson system of equations in phase space is based solely on the most fundamental object, the distribution function. In fact, one important outcome of our study was the observation that the background distribution function picks up backreactions from gravitational nonlinearities, a feature that is not transparent in the traditional approach based on the density and velocity fields alone since these quantities do not get renormalized. We expect such backreactions into the background distribution function to arise not only from perturbative scales, but also from nonperturbative ones. For example, within halos the average distribution function was found to have a Gaussian core with exponential wings due to virial velocities [335, 336]. This observation suggests a natural path towards improving on SPT: To insert the *apriori* unknown fully nonlinear background distribution function into Eq. (6.16) for its fluctuations, which should then be solved perturbatively as before.

We begin this section by exploring that idea, which will naturally point at a connection to EFT methods. We will see that the sound speed counterterm naturally emerges from the theory even without EFT ingredients, but we will ultimately argue that EFT methods are strictly necessary to account for a nonzero velocity dispersion in a fully self-consistent framework.

6.5.1 *Emergence of the counterterm*

Consider a split of the fully nonlinear background distribution function,

$$\bar{f}(\eta, q) = \bar{f}_{\rm P}(\eta, q) + \bar{f}_{\rm ctr}(\eta, q)\,, \tag{6.48}$$

into a perturbative piece, $\bar{f}_{\rm P}(\eta, q)$, which we calculated to third order in perturbation theory in Sec.6.4 [see Eq. (6.41), and recall that odd order terms vanish]

$$\bar{f}_{\rm P}(\eta, q) = \bar{f}^{(0)}_{\rm P}(q) + \bar{f}^{(2)}_{\rm P}(\eta, q) + \cdots\,, \tag{6.49}$$

and another one which we refer to as a counterterm piece, $\bar{f}_{\rm ctr}(\eta, q)$, with the benefit of hindsight. It accounts for the backreactions from short distance scales that are not under perturbative control,

$$\bar{f}_{\rm ctr}(\eta, q) = \bar{f}^{(2)}_{\rm ctr}(\eta, q) + \cdots\,. \tag{6.50}$$

We take the leading contribution to $\bar{f}_{\rm ctr}(\eta, q)$ as a second order quantity in perturbation theory since it is sourced by quadratic nonlinearities [18].

We now proceed to investigate the contribution from the new counterterm piece of the background distribution function to the perturbative expansion of the fluctuation $\delta f(\eta, \vec{k}, \vec{q})$. As we have shown in the previous section, it is only at third order in perturbation theory that the background distribution function backreacts into the fluctuations as given by the first line in Eq. (6.27), and due to the presence of $\bar{f}^{(2)}_{\rm ctr}(\eta, q)$ there will be a new term contributing to the integral equation that needs to be solved at third order. Its diagrammatic representation can be found in Fig.6.7, and it evaluates to

$$\delta f^{(3)}(\eta, \vec{k}, \vec{q}) = \int_0^\eta d\eta' a^2(\eta') e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')} \phi^{(1)}(\eta', \vec{k}) i\vec{k} \cdot \left.\frac{\partial \bar{f}^{(2)}_{\rm ctr}}{\partial \vec{q}}\right|_{\eta', q} + \cdots\,, \tag{6.51}$$

[18] This is not strictly necessary as one can remain agnostic about the size of $\bar{f}_{\rm ctr}(\eta, q)$, which can be thought of as resuming the dissipative effects associated to a nonzero average velocity dispersion to all orders. In that case, deviations from an ideal pressureless fluid appear already at the linear theory level. This choice is made, for instance, in [296–298]. On the other hand, the standard EFTofLSS assumes its counterterms to have a leading second order contribution as in our framework, so that linear theory is not modified.

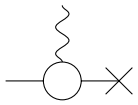

Figure 6.7: Diagram for the counterterm background distribution function contribution to fluctuations at third order in perturbation theory.

where the ellipsis represent all perturbative contributions which were accounted for in the previous subsection. Using the Poisson Eq. (6.29), Eq. (6.35) and integrating over momentum to obtain the density perturbation as in Eq. (6.3) yields

$$\begin{aligned}\delta^{(3)}(\eta,\vec{k},\vec{q}) =&\frac{1}{a^3\bar{\rho}}\delta_{\rm L}(\vec{k})\,\frac{3}{2}\Omega_{\rm m,0}H_0^2\int_0^\eta d\eta' a(\eta')D_{\rm L}(\eta')(\eta-\eta')\times\\ &\times\int\frac{d^3\vec{q}}{(2\pi)^3}e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}\bar{f}^{(2)}_{\rm ctr}(\eta',q)+\cdots\,.\end{aligned}\tag{6.52}$$

Let us take a closer look into the integral over momentum in the second line of Eq. (6.52). First recall that the comoving momentum is defined by $\vec{q}=a^2(d\vec{x}/dt)$ and hence $q\int dt/a^2\sim q\eta$ is the comoving distance traveled by cold dark matter particles, which is known to be of order $\sim 1/k_{\rm NL}$ with $k_{\rm NL}$ the scale of nonlinearities. This is the physical scale above which one expects SPT to break down, and as a consequence perturbative methods efficiently model the dynamics on scales $k\ll k_{\rm NL}$. Now note that the argument of the exponent in the second line of Eq. (6.52) scales like $\sim k/k_{\rm NL}\ll 1$, which justifies a Taylor series expansion

$$e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}=1-i\vec{k}\cdot\vec{q}(\eta-\eta')-\frac{1}{2}(\vec{k}\cdot\vec{q})^2(\eta-\eta')^2+\cdots\,.\tag{6.53}$$

Here the zeroth order term corresponds to a renormalization of the background mass density, which we know does not occur in perturbation theory from Sec. 6.4.2 (we will show this explicitly for the counterterm piece as well in Sec. 6.5.2, within the context of EFT methods). Also, the first order term is proportional to $\sim(\hat{k}\cdot\hat{q})$ which vanishes after integrating over the solid angle. We are then left with

$$\int\frac{d^3\vec{q}}{(2\pi)^3}e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}\bar{f}^{(2)}_{\rm ctr}(\eta',q)\approx-\frac{1}{2}(\eta-\eta')^2\int\frac{d^3\vec{q}}{(2\pi)^3}(\vec{k}\cdot\vec{q})^2\bar{f}^{(2)}_{\rm ctr}(\eta',q)\,,\tag{6.54}$$

After integrating over the solid angle, the substitution of Eq. (6.54) into Eq. (6.52) gives

$$\delta^{(3)}(\eta,\vec{k},\vec{q}) = -\frac{1}{6}k^2\delta_{\rm L}(\vec{k}) \times \frac{3}{2}\Omega_{\rm m,0}H_0^2\int_0^\eta d\eta' a^3(\eta')D_{\rm L}(\eta')(\eta-\eta')^3\sigma^2_{\rm dis}(\eta') + \cdots\,, \tag{6.55}$$

where we introduced the average velocity dispersion squared

$$\bar{\rho}(\eta)\sigma^2_{\rm dis}(\eta) = a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}q^2\bar{f}^{(2)}_{\rm ctr}(\eta,q)\,, \tag{6.56}$$

and used the relation $a^3(\eta)\bar{\rho}(\eta) = a^3(\eta')\bar{\rho}(\eta')$ to bring this quantity inside the time integral.

Before moving forward, let us stop for a moment to see explicitly that Eq. (6.56) indeed corresponds to the average velocity dispersion, a quantity directly related to the averaged stress tensor $\tau(\eta)$, which reads from Eq. (6.42)

$$\bar{\rho}(\eta)\sigma^2_{\rm dis}(\eta) \equiv \tau(\eta) = a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}q^2\bar{f}(\eta,q) - \langle\rho v^2\rangle\big|_\eta\,. \tag{6.57}$$

However, from Eq. (6.48) we have that $\bar{f}(\eta,q) = \bar{f}_{\rm P}(\eta,q) + \bar{f}_{\rm ctr}(\eta,q)$, and the contributions from $\bar{f}_{\rm P}(\eta,q)$ exactly cancel those from $\langle\rho v^2\rangle\big|_\eta$ according to the arguments made around Eq. (6.43) in Sec. 6.4.2 (also see a discussion about this result at the end of that section). We are then left with Eq. (6.56).

The diagram in Fig. 6.7, which evaluates to Eq. (6.55), introduces an extra contribution, $c^{(3)}_{\rm ctr}(a)h^{(3)}_{\rm ctr}(\vec{k})$, to the decomposition of the third order density contrast into a sum of separable terms, Eq. (6.46). The scale dependent part can be read off to be,

$$h_{\rm ctr}(\vec{k}) = -\frac{1}{2}k^2\delta_{\rm L}(\vec{k})\,. \tag{6.58}$$

The time-dependent coefficient, $c^{(3)}_{\rm ctr}(a)$, can be computed using the machinery developed in Appendices 6.7 and 6.8. We summarize here the steps involved for completeness. We first need to compute the source function $s^{(3)}_{\rm ctr}(a)$, associated to $c^{(3)}_{\rm ctr}(a)$ via Eq. (6.87). As explained in Appendix 6.8, this can be obtained from Eq. (6.55) by differentiating it twice with respect to superconformal time followed by a division of the result by a factor of a^6H^2. This reads [also factoring out $h^{(3)}_{\rm ctr}(\vec{k})$ from Eq. (6.55)]:

$$s^{(3)}_{\rm ctr} = \frac{2}{a^6H^2}\,\frac{3}{2}\Omega_{\rm m,0}H_0^2\int_0^\eta d\eta' a^3(\eta')D_{\rm L}(\eta')(\eta-\eta')\sigma^2_{\rm dis}(\eta')\,. \tag{6.59}$$

The time-dependent coefficient $c^{(3)}_{\rm ctr}(a)$ can then be computed from Eq. (6.93) with the source term Eq. (6.59), assuming knowledge of $\bar{f}^{(2)}_{\rm ctr}(q,\eta)$ and hence of $\sigma^2_{\rm dis}(\eta)$. It then follows from Eq. (6.58) that this new contribution to the density contrast at third order, coming from the backreaction of the counterterm background distribution function into its fluctuations, adds a new term to the one-loop power spectrum

$$\Delta P(a,k) = P_{\text{1-loop}}(a,k) - P_{\text{1-loop,SPT}}(a,k) = -D_{\rm L}(a)c^{(3)}_{\rm ctr}(a)k^2 P_{\rm L}(k)\,. \tag{6.60}$$

This accounts for a nonzero average velocity dispersion, and it has the exact same form as the effective sound speed counterterm in the EFTofLSS. In fact, our framework qualitatively captures the same physical effects as the EFT: The backreation of short distance fluctuations into the background, and its impact on the dynamics of long-wavelength fluctuations [42, 43].

Since $\bar{f}^{(2)}_{\rm ctr}(q,\eta)$ is not a priori known, the same is true of $\sigma^2_{\rm dis}(\eta)$ and hence of $c^{(3)}_{\rm ctr}(a)$ as well. Instead, one can think of it as a free parameter to be determined by matching to either N-body simulations or observations. This is analogous to the case of chiral perturbation theory in QCD [337], where a nonzero quark condensate emerges due to nonperturbative effects (i.e. confinement) and the underlying EFT is build from an apriori unknown value for this quantity. Also note that the scale dependence in Eq. (6.60) is just right so that the free coefficient $c^{(3)}_{\rm ctr}(a)$ can absorb UV divergences present in the perturbation theory loop integrals, as is well-known in the EFTofLSS [19]. One can then proceed to renormalize cosmological perturbation theory in the exact same way one would any other field theory, for cosmologies that suffer from those UV divergences [338].

So long as we think of $c^{(3)}_{\rm ctr}(a)$ as a free parameter, our formalism leads to a model for the one-loop power spectrum witch is identical to the EFTofLSS in its simplest form. This is not, however, the end of the story. The reason for this is the fact that we do happen to have a really good handle on the nonperturbative gravitational dynamics from first principles via N-body simulations (as opposed to the more traditional EFT approach to parametrize unknown

[19] Stochastic terms are also required to absorb all divergences. We will have more to say about stochastic terms in the next section.

physics), so we should be able to extract the counterterm from simulations and, crucially, have a physical interpretation for what it represents. The analogy to chiral perturbation theory, which has the pion decay constant as a free parameter, makes sense here as well. This quantity can be extracted from first principles using QCD lattice simulations [339–342] and has a clear physical interpretation.

Previous works have pointed out that SPT can be improved by incorporating a nonzero average velocity dispersion in the formalism. It was argued in [299, 310] that a nonzero average velocity dispersion can regulate the shell-crossing singularity, and our setup is similar to the approaches taken in both [296–298] and [305], where a nonzero average velocity dispersion is incorporated in the perturbative expansion, leading to the emergence of EFT-like counterterms. We reproduce these results within a framework that solves the Vlasov-Poisson system of equations directly in phase space, and hence does not require a truncation of the Boltzmann hierarchy from the outset.

The results obtained in this section seem to provide a physical interpretation for the counterterm in terms of an average velocity dispersion. In what follows we show, however, that this is necessarily incomplete due to short distance fluctuations which are not under perturbative control. To be concrete, we will see that a nonzero average velocity dispersion is directly sourced by a short-scale gravitational binding energy which is not accounted for in our framework, but it should contribute to the counterterm. On top of that, it is not the averaged quantities that source the EFT counterterm, but rather their response to the presence of a long-wavelength mode (see [83] for a concrete realization of this within the separate universe picture). For these reasons, the EFT approach is strictly necessary to account for a nonzero velocity dispersion within a fully self-consistent framework.

6.5.2 EFTofLSS framework

We now introduce standard EFT techniques to make the connection between the formalism developed in Sec. 6.5.1 and the EFTofLSS more precise. As a starting point, let us repeat

here the collisionless Boltzmann (Vlasov) Eq. (6.14),

$$\frac{\partial f}{\partial \eta} + \vec{q}\cdot\frac{\partial f}{\partial \vec{x}} = a^2(\eta)\frac{\partial \phi}{\partial \vec{x}}\cdot\frac{\partial f}{\partial \vec{q}}\,. \tag{6.61}$$

The nonlinear term on the right-hand side of Eq. (6.61) is a contact term [20], being sensitive to nonlinearities in gravitational dynamics at arbitrarily small scales. In order to have theoretical control over scales that cannot be treated perturbatively, we adopt standard EFT techniques and split the distribution function into a long-wavelength piece, defined by smoothing over a (distance) scale $1/\Lambda$,

$$f_l(\eta,\vec{x},\vec{q}) = \int d^3\vec{x}'\, W_\Lambda(|\vec{x}-\vec{x}'|) f(\eta,\vec{x}',\vec{q})\,, \tag{6.62}$$

and a short-wavelength piece $f_s = f - f_l$ [21]. The evolution equation for the long-wavelength distribution function follows from Eqs. (6.61) and (6.62)

$$\frac{\partial f_l}{\partial \eta} + \vec{q}\cdot\frac{\partial f_l}{\partial \vec{x}} = a^2(\eta)\int d^3\vec{x}'\, W_\Lambda(|\vec{x}-\vec{x}'|)\frac{\partial \phi(\eta,\vec{x}',\vec{q})}{\partial \vec{x}'}\cdot\frac{\partial f(\eta,\vec{x}',\vec{q})}{\partial \vec{q}}\,. \tag{6.63}$$

By following standard methods (e.g., see [42]) one obtains the result that mixed terms involving a product of long and short modes are suppressed by powers of $(k/\Lambda)^2 \ll 1$, and we arrive at

$$\frac{\partial f_l}{\partial \eta} + \vec{q}\cdot\frac{\partial f_l}{\partial \vec{x}} = a^2(\eta)\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}} + a^2(\eta)\left[\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right]_\Lambda + O\left(\frac{k^2}{\Lambda^2}\right)\,, \tag{6.64}$$

where,

$$\left[\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right]_\Lambda = \int d^3\vec{x}'\, W_\Lambda(|\vec{x}-\vec{x}'|)\frac{\partial \phi_s(\eta,\vec{x}',\vec{q})}{\partial \vec{x}'}\cdot\frac{\partial f_s(\eta,\vec{x}',\vec{q})}{\partial \vec{q}}\,. \tag{6.65}$$

To make direct contact with the formalism developed in Sec. 6.4 we first split the Boltzmann equation into two coupled set of equations, one for the background distribution function and another for its fluctuation. The first term on the right-hand side of Eq. (6.64) can be decomposed as before

$$a^2(\eta)\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}} = a^2(\eta)\,{:}\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}}{:} + a^2(\eta)\left\langle\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}}\right\rangle\,, \tag{6.66}$$

[20] A product of two fields evaluated at the same point in space.

[21] A similar decomposition, $\phi = \phi_l + \phi_s$, also follows for the gravitational potential, since it is linearly related to the distribution function by Eqs. (6.3) and (6.11).

where the normal ordering symbol subtracts averages as defined in Eq. (6.17). A similar expression holds for the second term on the right-hand side of Eq. (6.64)

$$a^2(\eta)\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda = a^2(\eta):\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda: + a^2(\eta)\left\langle\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda\right\rangle . \tag{6.67}$$

From Eq. (6.65), the second term on the right-hand side of Eq. (6.67) simplifies to

$$\left\langle\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda\right\rangle = \left\langle\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right\rangle . \tag{6.68}$$

We can also rewrite the first term on the right-hand side of Eq. (6.67) as follows. First define the stochastic term,

$$\epsilon_\Lambda \equiv \left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda - \left\langle\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda\right\rangle_{\delta f_l(\eta,\vec{x},\vec{q})} \tag{6.69}$$

where the subscript $\delta f_l(\eta,\vec{x},\vec{q})$ on the second term in the previous expression indicates that the average is to be taken on the presence of a long-wavelength fluctuation to the distribution function. The quantity ϵ_Λ is entirely analogous to the stochastic term usually introduced in the context of the standard EFTofLSS framework [42, 43]. Upon expanding the short scale fluctuations in terms of the long-wavelength mode, we obtain

$$\left\langle\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda\right\rangle_{\delta f_l(\eta,\vec{x},\vec{q})} - \left\langle\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right\rangle = \frac{\partial}{\partial\delta f_l}\left\langle\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda\right\rangle_{\delta f_l}\Bigg|_{\delta f_l=0}\delta f_l(\eta,\vec{x},\vec{q}) + \cdots , \tag{6.70}$$

where the ellipsis represent terms of higher order in δf_l and derivatives thereof. Combining Eqs. (6.67), (6.69) and (6.70) now yields

$$:\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda: = \frac{\partial}{\partial\delta f_l}\left\langle\left[\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right]_\Lambda\right\rangle_{\delta f_l}\Bigg|_{\delta f_l=0}\delta f_l(\eta,\vec{x},\vec{q}) + \cdots + \epsilon_\Lambda \tag{6.71}$$

Going back to the Vlasov Eq. (6.64), its ensemble average reads [22]

$$\frac{\partial\bar{f}_l}{\partial\eta} = a^2(\eta)\left\langle\frac{\partial\phi_l}{\partial\vec{x}}\cdot\frac{\partial f_l}{\partial\vec{q}}\right\rangle + a^2(\eta)\left\langle\frac{\partial\phi_s}{\partial\vec{x}}\cdot\frac{\partial f_s}{\partial\vec{q}}\right\rangle , \tag{6.72}$$

[22] Note that we drop higher derivative terms of $O(k^2/\Lambda^2)$ from now on. However, it is important to keep in mind that such terms are necessary to accurately model the nonlinear power spectrum for any finite value of the cutoff Λ [42, 43].

and subtracting Eq. (6.72) from Eq. (6.64) leads to

$$\frac{\partial \delta f_l}{\partial \eta} + \vec{q}\cdot\frac{\partial \delta f_l}{\partial \vec{x}} = a^2(\eta)\left[:\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}}: + \frac{\partial}{\partial \delta f_l}\left\langle\left[\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right]_\Lambda\right\rangle_{\delta f_l}\Bigg|_{\delta f_l=0}\delta f_l(\eta,\vec{x},\vec{q}) + \cdots + \epsilon_\Lambda\right]. \tag{6.73}$$

Let us reinforce that since $f_l = \bar{f}_l + \delta f_l$ appears on the right-hand side of Eqs. (6.72) and (6.73), these are two coupled set of equations. Note that Eq. (6.73) contains additional terms when compared to Eq. (6.16). They parametrize the dependence of short-wavelength fluctuations on the presence of a long mode and the effects of stochasticity, and were completely ignored in Sec. 6.4. From this we can already see that the framework developed in Sec. 6.4 is necessarily incomplete, but we will come back to this point in Sec. 6.5.3. However, in Eq. (6.72) short-scale fluctuations act as an additional source to the background distribution function $\bar{f}_l(\eta, q)$, which was exactly the basis for including a nonzero average velocity dispersion in Sec. 6.5.1, and naturally arises in the EFT framework.

Let us recall from Sec. 6.4 that at zeroth order in the perturbative expansion $\bar{f}_l^{(0)}(\eta, q) \propto \delta^{(3)}(\vec{q})$ for cold dark matter, and $\delta f_l^{(0)}(\eta, \vec{k}, \vec{q}) = 0$. Also, it is only at second order that the background distribution function picks up corrections (the first order contribution vanishes), in which case we can split as before

$$\bar{f}_l^{(2)}(\eta, q) = \bar{f}_{l,\mathrm{P}}^{(2)}(\eta, q) + \bar{f}_{l,\mathrm{ctr}}^{(2)}(\eta, q)\,, \tag{6.74}$$

where the perturbative piece satisfies the equation we solved in Sec. 6.4.2

$$\frac{\partial \bar{f}_{l,\mathrm{P}}^{(2)}}{\partial \eta} = a^2(\eta)\left\langle\frac{\partial \phi_l^{(1)}}{\partial \vec{x}}\cdot\frac{\partial \delta f_l^{(1)}}{\partial \vec{q}}\right\rangle, \tag{6.75}$$

and the counterterm piece is sourced by short-scale fluctuations

$$\frac{\partial \bar{f}_{l,\mathrm{ctr}}^{(2)}}{\partial \eta} = a^2(\eta)\left\langle\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right\rangle, \tag{6.76}$$

where we treat the right-hand side in Eq. (6.76) as an external source with leading contribution to second order in the expansion (see [18] for additional comments on this). We can then

immediately write

$$\bar{f}^{(2)}_{l,\rm ctr}(\eta,q;\Lambda)=\int_0^\eta d\eta' a^2(\eta')\left\langle\frac{\partial\phi_s(\eta',\vec{x})}{\partial\vec{x}}\cdot\frac{\partial f_s(\eta',\vec{x},\vec{q})}{\partial\vec{q}}\right\rangle\,, \tag{6.77}$$

where here we include explicitly the cutoff dependence of the counterterm contribution to the long-wavelength background distribution function, which will be omitted in what follows to simplify the notation. Note that Eq. (6.77) is a total derivative with respect to momentum, and hence it does not renormalize the background (mass) density, according to Eq. (6.3). We previously used this result without proof in Eq. (6.54), and it also holds true for the perturbative piece of the background distribution function as we showed in Eq. (6.41).

We are now ready to obtain an expression for the average velocity dispersion squared in terms of short scale fluctuations. It will be convenient to first rephrase it as a short scale kinetic energy per unit mass, $\kappa(\eta)$, where [essentially repeating Eq. (6.56)]

$$\bar{\rho}(\eta)\kappa(\eta)\equiv\frac{1}{2}\bar{\rho}(\eta)\sigma^2_{\rm dis}(\eta)=\frac{1}{2}a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}q^2\bar{f}^{(2)}_{l,\rm ctr}(\eta,q)\,. \tag{6.78}$$

This can be obtained from Eq. (6.77) after multiplying it by q^2 followed by an integration over momentum. A detailed derivation can be found in Appendix 6.9, but the final result is

$$\kappa(\eta)=-a^{-2}(\eta)\int_0^\eta d\eta' a(\eta')\frac{d}{d\eta'}\left[a(\eta')u(\eta')\right]\,, \tag{6.79}$$

where

$$\bar{\rho}(\eta)u(\eta)=\frac{1}{2}\langle\phi_s(\eta,\vec{x})\rho_s(\eta,\vec{x})\rangle\,, \tag{6.80}$$

defines the short scale gravitational binding energy per unit mass, $u(\eta)$. A more familiar form of Eq. (6.79) can be obtained by differentiating it with respect to superconformal time

$$\frac{d}{d\eta}(\kappa+u)+\mathcal{H}(2\kappa+u)=0\,, \tag{6.81}$$

where we recall that $\mathcal{H}=d\log a/d\eta=a^2H$. This is nothing but the Layzer-Irvine equation [343], which generalizes the notion of energy conservation to an expanding background [23].

[23] The expansion breaks time-translation symmetry, so energy is not conserved in general. From Eq. (6.81) we see that energy is conserved only for virialized scales, for which $2\kappa+u=0$.

Let us also emphasize that both κ and u appearing in Eq. (6.81) are short-scale quantities as defined in Eqs. (6.78) and (6.80), while the Layzer-Irvine equation is often phrased in terms of the total kinetic and potential energies with contributions from all scales. We present a derivation of Eq. (6.79), and hence of Eq. (6.81), from the Boltzmann equation in Appendix 6.7, which holds true because contributions from long-wavelengths separately satisfy the very same equation, and can hence be subtracted.

The integral form of the cosmic energy Eq. (6.79) shows that a nonzero average velocity dispersion is effectively sourced by a short-scale gravitational binding energy, which is not accounted for in perturbation theory, and in fact should contribute to the counterterm as we will now argue in Sec. 6.5.3.

6.5.3 The inevitableness of EFT methods

In Sec. 6.5.1 we have seen that an EFT-like counterterm, Eq. (6.60), naturally arises within an old-fashioned cosmological perturbation theory approach based directly on a perturbative expansion for the distribution function fluctuations in full phase space. In that framework the counterterm is sourced by a nonzero average velocity dispersion, which is itself linked to the short-scale gravitational binding energy via the Layzer-Irvine Eq. (6.79) as we showed in Sec. 6.5.2.

We will now argue that this short-scale gravitational binding energy should directly contribute to the counterterm as well to ensure the self-consistency of the theory. This implies that the framework developed in Sec. 6.5.1 is not self-consistent since only the velocity dispersion directly contributes to the counterterm, although it is formally identical to the EFT as long as one takes the counterterm to be a free parameter. Another way of phrasing this is to say that the value obtained for the counterterm $c^{(3)}_{\rm ctr}(a)$, from directly matching Eq. (6.60) to simulations or observations, would not agree with the one calculated via Eqs. (6.59) and (6.93), with an average velocity dispersion squared $\sigma^2_{\rm dis}(a)$ equally extracted from simulations or observations.

The statement that the short-scale gravitational binding energy should also contribute to

the counterterm can be seen from the well-understood decoupling of virialized scales [344]. For convenience, let us repeat here the Euler Eq. (6.6)

$$\Pi_i' + 4\mathcal{H}\Pi_i + 2a\partial^j K_{ij} + a\rho\partial_i\phi = 0\,, \tag{6.82}$$

where

$$K_{ij} = \frac{1}{2}a^{-5}\int \frac{d^3\vec{q}}{(2\pi)^3}\, q_i q_j\, f\,, \tag{6.83}$$

is the kinetic energy density tensor. But we can write from the Poisson Eq. (6.11),

$$\rho\partial_i\phi = (\rho-\bar{\rho})\partial_i\phi + \bar{\rho}\partial_i\phi = \frac{\nabla^2\phi\partial_i\phi}{4\pi G a^2} + \bar{\rho}\partial_i\phi = \partial^j W_{ij} + \bar{\rho}\partial_i\phi\,, \tag{6.84}$$

where

$$W_{ij} = \frac{1}{4\pi G a^2}\left[\partial_i\phi\partial_j\phi - \frac{1}{2}\delta_{ij}(\vec{\nabla}\phi)^2\right]\,, \tag{6.85}$$

is the potential energy density tensor. The Euler Eq. (6.82) can then be written in the following form

$$\Pi_i' + 4\mathcal{H}\Pi_i + a\bar{\rho}\partial_i\phi + a\partial^j(2K_{ij} + W_{ij}) = 0\,. \tag{6.86}$$

Note that $K = \langle K_i^i\rangle$ is just the kinetic energy density, and similarly from Eq. (6.85), $W = \langle W_i^i\rangle = \langle\phi\rho\rangle/2$ is the gravitational binding energy density, where we integrate by parts inside the spatial average [24], use Eq. (6.11) once again and assume $\langle\phi\rangle = 0$. For virialized scales, $2K + W = 0$ and all nonlinear terms in the Euler Eq. (6.86) vanish exactly. Of course this simplified argument ignores the tensor structure of these quantities, but a full tensorial virial decoupling theorem can be derived from the collisionless Boltzmann equation (see [345] for example). It is important to emphasize that the decoupling of virialized scales is exact [42]. This contrasts with, being significantly more constraining than, the expectation that contributions from virialized scales are parametrically suppressed but not exactly vanishing.

It is now straightforward to argue that the old-fashioned cosmological perturbation theory framework developed in Sec. 6.5.1 is necessarily incomplete, since the average velocity

[24] The ensemble average can also be thought of as a volume average, which involves integrating over position. This allows an integration by parts under the average sign.

dispersion in Eq. (6.57) is the quantity that sources the counterterm $c^{(3)}_{\rm ctr}(a)$ [see Eq. (6.60)] via Eqs. (6.59) and (6.93). However, we also know that virial velocities within halos [335, 336] give a significant (if not dominant) contribution to the average velocity dispersion, and hence to the counterterm as well. This is inconsistent with the decoupling of virialized scales.

The resolution to this paradox comes from the realization that the short scale gravitational binding energy should also directly contribute to the counterterm, in such a way that it vanishes whenever $2K + W = 0$. Additionally, just averages of the short-scale kinetic and potential energies are not sufficient to accurately model the sound speed counterterm. Instead, one needs to consider how such averages respond to the presence of a long-wavelength fluctuation [42, 43, 83]. We can explicitly identify the origin of the deficiencies in the framework developed in Sections 6.4.2 and 6.5.1 by comparing Eqs. (6.16) and (6.73) . The latter equation includes additional terms that are neglected in the former, which parametrize stochasticity and the response of short-wavelength fluctuations to the presence of a long mode. This shows that short scale nonlinearities directly backreact into the background and fluctuations to the distribution function alike, and it is not self-consistent to only model the backreactions to the background distribution function. We can overcome these deficiencies by keeping the effective sound speed counterterm as a free parameter in the model, to be fitted by observations or full cosmological simulations. In that sense, EFT methods emerge as a necessary framework to self-consistently model a non-zero velocity dispersion.

All of that being said, one way in which one may be able to improve on the standard EFTofLSS framework is to solve Eq. (6.73) perturbatively for the distribution function fluctuations while assuming knowledge of the fully nonlinear background distribution function, which can be measured from simulations. This is equivalent to resumming the effects of the background distribution function backreacting into its fluctuations to all orders in perturbation theory, while working with an EFT perturbative expansion for the fluctuations. In practice, this would look a lot like the standard EFTofLSS framework augmented with modified perturbation theory kernels. A similar idea was proposed recently in [298].

6.6 Conclusion

Effective field theory methods for large scale structure significantly improve on Standard Perturbation Theory (SPT) techniques by modeling deviations from an ideal fluid [42, 43, 273–281], and are now ubiquitous in analysis pipelines of large scale structure surveys [52–54, 248–261].

In its simplest form [25], the EFTofLSS comes at the cost of adding one free nuisance parameter, the effective sound speed. This is arguably not a desirable feature, after all N-body simulations have no free parameters, and we are entering a new era where efficient emulators are available to interpolate the predictions from simulations in broad regions of parameter (and even theory) space [46, 286–295]. This naturally raises the question of whether or not the additional free parameter is really necessary (and this becomes even more relevant in light of recent studies on prior volume effects in the EFTofLSS [53, 346–352]).

The starting point of Standard Perturbation Theory (SPT) methods for large scale structure is the assumption of negligible stress tensor [41]. This leaves open the possibility that a perturbative approach can account for these effects by avoiding to truncate the Boltzmann hierarchy at the level of the Euler equation [296–305, 308, 309]. In this work we revisit this issue in light of a framework for large scale structure perturbation theory, that directly solves for the distribution function in full phase space, expanding upon earlier work [307]. This approach circumvents the need to artificially truncate the Boltzmann hierarchy and hence relaxes the usual *apriori* assumption of a negligible velocity dispersion.

In Sec. 6.4 we introduce the framework underlying the cosmological perturbation theory in phase space. It is based on the coupled set of Boltzmann Eqs. (6.15) and (6.16) for the (ensemble) average distribution function and its fluctuations, respectively. We show that perturbatively solving this coupled set of equations directly in phase space reproduces the familiar SPT kernels, which underscores the statement that a negligible velocity dispersion

[25] Modeling the power spectrum of the matter field, in real space (as opposed to redshift space), at one-loop in perturbation theory. This is assumed in the discussion that follows.

and vorticity should be viewed as a consequence of the perturbative expansion, rather than an assumption [307].

We begin Sec. 6.5 by showing that a nonzero average velocity dispersion can be easily incorporated in the framework by perturbatively solving for the fluctuations in the distribution function, while assuming knowledge of its fully nonlinear ensemble average. This leads to a new contribution to the one-loop power spectrum with the exact same form as the effective sound speed counterterm in the EFTofLSS [see Eq. (6.60)]. This result corroborates similar previous findings in the literature [297, 304, 305]. We then proceed to make a direct connection to EFT methods, and argue that our framework is necessarily incomplete because it misses contributions to the counterterm from the short-scale gravitational binding energy, and the response of short-wavelength fluctuations to the presence of a long mode. EFT methods then arise as an inevitable framework to self-consistently model a nonzero velocity dispersion. This is a practical example of the importance in having theoretical control over short distance fluctuations, in order to write a sensible theory.

Beyond the results summarized above, we intend this manuscript to be used as a useful pedagogical reference for numerical calculations in SPT with the full time-dependence of ΛCDM kernels, up to third order in the perturbative expansion. For that purpose, simple analytic formulas for the calculation of time-dependent coefficients can be found in Appendix 6.7. We also include explicit formulas for the one-loop power spectrum and tree-level bispectrum in Appendices 6.10 and 6.11, respectively.

To conclude, we land at a picture of the microphysics behind the EFTofLSS that suggests a close analogy to chiral perturbartion theory in QCD [337]: The stress tensor acts like an order parameter (analogous to the quark condensate) which vanishes in the perturbative regime, but picks up a nonzero value in the strongly coupled regime which follows after gravitational collapse (corresponding to the QCD phase transition). Moreover, the triaxial nature of gravitational collapse implies that an approximate symmetry of the perturbative regime, i.e. isotropy, is spontaneously broken as gravitational collapse first happens along a given axis. This is analogous to the breakdown of the approximate chiral symmetry in QCD. Fur-

thermore, just as in QCD lattice simulations [339] are available to compute nonperturbative processes of interest without the need to introduce any additional free parameters, the same is true in LSS with N-body simulations. Nonetheless, in the same way that chiral perturbation theory has proved itself to be a very valuable tool in QCD, recent progress over the past decade or so has established the EFTofLSS as a great framework to interpret both simulated and real data pertaining the evolution of LSS within an analytic perturbative framework. In this context, the additional free parameter is simply the price one needs to pay in order to parametrize intrinsically nonperturbative effects within a perturbative framework.

6.7 Appendix: Analytic solution to time-dependent coefficients

In this section we derive analytic formulas for the time-dependent coefficients in SPT, entering Eqs. (6.44) and (6.46), within the full ΛCDM cosmology. To accomplish this, we will first need to find an analytic solution to the differential equation

$$\frac{d^2c}{da^2}+\frac{1}{a}\left(3+\frac{d\log H}{d\log a}\right)\frac{dc}{da}-\frac{3}{2}\Omega_{\rm m,0}H_0^2\frac{c(a)}{a^5H(a)^2}=s(a)\,, \tag{6.87}$$

for the coefficient $c(a)$, given some source function $s(a)$. To achieve this, we apply the same trick that works in the $s(a)=0$ case. The first step is to note that the Hubble expansion rate $H(a)$ satisfies the homogeneous equation in ΛCDM (one can even add a nonzero curvature), that is:

$$\frac{d^2H}{da^2}+\frac{1}{a}\left(3+\frac{d\log H}{d\log a}\right)\frac{dH}{da}-\frac{3}{2}\Omega_{\rm m,0}H_0^2\frac{1}{a^5H(a)}=0\,. \tag{6.88}$$

To take advantage of this, we define a new function $g(a)=c(a)/H(a)$, and combine Eqs. (6.87) and (6.88) to arrive at

$$\frac{d}{da}\left[\frac{dg}{da}a^3H^3(a)\right]=a^3H^2(a)s(a)\,, \tag{6.89}$$

and the solution to this can be simply obtained by integrating the equation twice with respect

to the scale factor, which reads in terms of $c(a) = H(a)g(a)$

$$\begin{aligned} c(a) =& H(a)\Bigg[g(a_i) + g'(a_i)a_i^3 H^3(a_i)\int_{a_i}^{a}\frac{da'}{(a')^3H^3(a')} + \\ &+ \int_{a_i}^{a}\frac{da'}{(a')^3H^3(a')}\int_{a_i}^{a'}da''(a'')^3H^2(a'')s(a'')\Bigg]\,. \end{aligned} \tag{6.90}$$

In the case of $s(a) = 0$, Eq. (6.87) reduces to Eq. (6.38) satisfied by the linear growth factor. According to Eq. (6.90), the general solution is then a linear combination of the growing and decaying modes, $c_+(a)$ and $c_-(a)$ respectively, where

$$\begin{aligned} c_+(a) &= H(a)\int_0^a\frac{da'}{(a')^3H^3(a')} \\ c_-(a) &= H(a)\,, \end{aligned} \tag{6.91}$$

and we set $a_i = 0$. If we drop the decaying mode and normalize the linear growth according to $D_{\rm L}(a = 1) = 1$, we arrive at:

$$D_{\rm L}(a) = \frac{H(a)}{H_0}\left[\int_0^1\frac{da'}{(a')^3H^3(a')}\right]^{-1}\int_0^a\frac{da'}{(a')^3H^3(a')}\,, \tag{6.92}$$

hence reproducing a familiar result, which we used in Sec. 6.4. However, in the more general case of a nonzero source the particular solution typically grows faster than the homogeneous solutions and eventually dominates. We then obtain

$$c(a) = H(a)\int_0^a\frac{da'}{(a')^3H^3(a')}\int_0^{a'}da''(a'')^3H^2(a'')s(a'')\,. \tag{6.93}$$

With this basic ingredient we can write down formulas for the time-dependent coefficients in SPT. Starting at second order in the perturbative expansion, we show in Appendix 6.8 that the coefficients $c_i^{(2)}(a)$ are solutions to the differential Eq. (6.87), with source functions

$$\begin{aligned} s_1^{(2)} &= \frac{D_{\rm L}^2 f}{a^2}\left[2(1+f) + \frac{d\log H}{d\log a} + \frac{d\log f}{d\log a}\right] \\ s_2^{(2)} &= \frac{D_{\rm L}^2 f^2}{a^2}\,, \end{aligned} \tag{6.94}$$

where the dependence on scale factor is implicit in Eq. (6.94). The solutions are then obtained from a direct application of Eq. (6.93), which can be easily evaluated numerically. For an

Einstein-de Sitter (EdS) universe with $\Omega_{\rm m}(a) = 1$, $D_{\rm L}(a) = a$ which implies $f(a) = 1$. In this case the source terms in Eq. (6.94) simplify to $s_1^{(2)}(a) = 5/2$ and $s_2^{(2)}(a) = 1$, and it becomes straightforward to derive the time-dependent coefficients: $c_1^{(2)}(a) = 5a^2/7$ and $c_2^{(2)}(a) = 2a^2/7$. This motivates the familiar EdS approximation to time-dependent coefficients in the general ΛCDM cosmology:

$$\begin{aligned} c_{1,\rm EdS}^{(2)}(a) &\approx \frac{5}{7} D_{\rm L}^2(a) \\ c_{2,\rm EdS}^{(2)}(a) &\approx \frac{2}{7} D_{\rm L}^2(a)\,. \end{aligned} \tag{6.95}$$

In Fig. 6.8 we plot these time-dependent coefficients in our fiducial cosmology. The exact solutions from Eqs. (6.93) and (6.94) are shown as solid lines, and the approximated ones from Eq. (6.95) as dashed lines. The EdS approximation works extremely well, with $< 1\%$ errors.

The third order time-dependent coefficients also satisfy the same differential Eq. (6.87), with new source functions

$$\begin{aligned} s_1^{(3)} &= \frac{f D_{\rm L}}{a}\frac{dc_1^{(2)}}{da} + \frac{f D_{\rm L}}{a^2}\left(2 + f + \frac{d\log H}{d\log a} + \frac{d\log f}{d\log a}\right) c_1^{(2)} \\ s_2^{(3)} &= \frac{f D_{\rm L}}{a}\frac{dc_2^{(2)}}{da} + \frac{f D_{\rm L}}{a^2}\left(2 + f + \frac{d\log H}{d\log a} + \frac{d\log f}{d\log a}\right) c_2^{(2)} \\ s_3^{(3)} &= \frac{3}{2}\Omega_{\rm m,0} H_0^2 \frac{D_{\rm L} c_1^{(2)}}{a^5 H^2} + \frac{f D_{\rm L}}{a}\frac{dc_1^{(2)}}{da} - \frac{f^2 D_{\rm L}^3}{a^2} \\ s_4^{(3)} &= \frac{3}{2}\Omega_{\rm m,0} H_0^2 \frac{D_{\rm L} c_2^{(2)}}{a^5 H^2} + \frac{f D_{\rm L}}{a}\frac{dc_2^{(2)}}{da} + \frac{f^2 D_{\rm L}^3}{a^2} \\ s_5^{(3)} &= 2\frac{f D_{\rm L}}{a}\frac{dc_1^{(2)}}{da} - 2\frac{f^2 D_{\rm L}^3}{a^2} \\ s_6^{(3)} &= 2\frac{f D_{\rm L}}{a}\frac{dc_2^{(2)}}{da}\,. \end{aligned} \tag{6.96}$$

As before, solutions are obtained from a direct application of Eq. (6.93). One can derive the EdS approximation to the third order coefficients in the exact same way as done in the

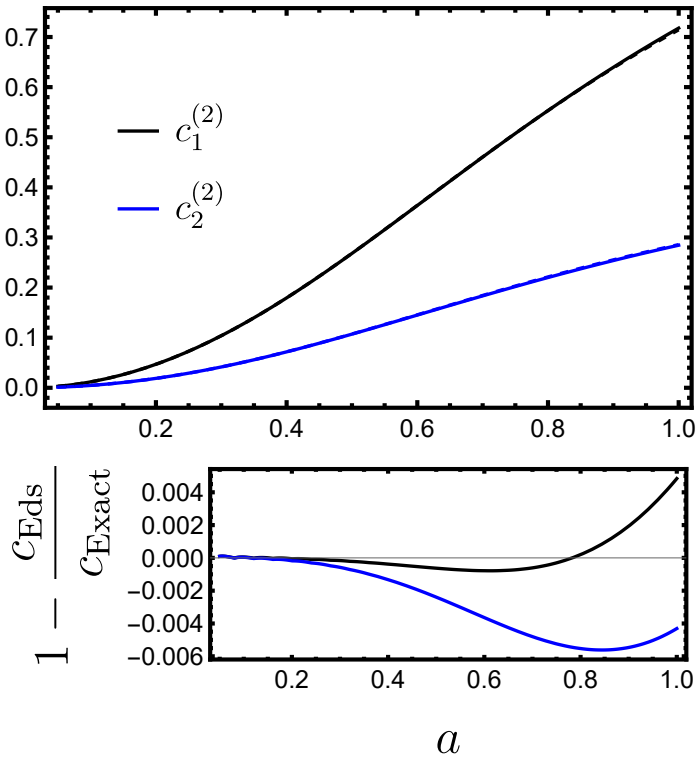

Figure 6.8: Time-dependent coefficients for the second order perturbation theory kernels, in our fiducial cosmology, as a function of the scale factor. Solid and dashed lines (almost indistinguishable) correspond to exact solutions from Eqs. (6.93) and (6.94), and the EdS approximation in Eq. (6.95), respectively. The lower plot shows the relative difference between the two, which is always at the sub-percent level.

previous case of second order coefficients. This reads

$$\begin{aligned}
c^{(3)}_{1,\rm EdS}(a) &\approx \frac{5}{18} D^3_{\rm L}(a) \\
c^{(3)}_{2,\rm EdS}(a) &\approx \frac{1}{9} D^3_{\rm L}(a) \\
c^{(3)}_{3,\rm EdS}(a) &\approx \frac{1}{6} D^3_{\rm L}(a) \\
c^{(3)}_{4,\rm EdS}(a) &\approx \frac{2}{9} D^3_{\rm L}(a) \\
c^{(3)}_{5,\rm EdS}(a) &\approx \frac{2}{21} D^3_{\rm L}(a) \\
c^{(3)}_{6,\rm EdS}(a) &\approx \frac{8}{63} D^3_{\rm L}(a)\,.
\end{aligned} \tag{6.97}$$

In Fig. 6.9 we plot the time-dependent coefficients for the third order perturbation theory kernels, $c_i^{(3)}(a)$ with $i = \overline{1,6}$. The exact solutions from Eqs. (6.93) and (6.96) are shown as solid lines, and the approximated ones from Eq. (6.97) as dashed lines. The EdS approximation still works quite well, but slightly worse than in the second order case shown in Fig.6.8, with $\lesssim 1.5\%$ errors overall.

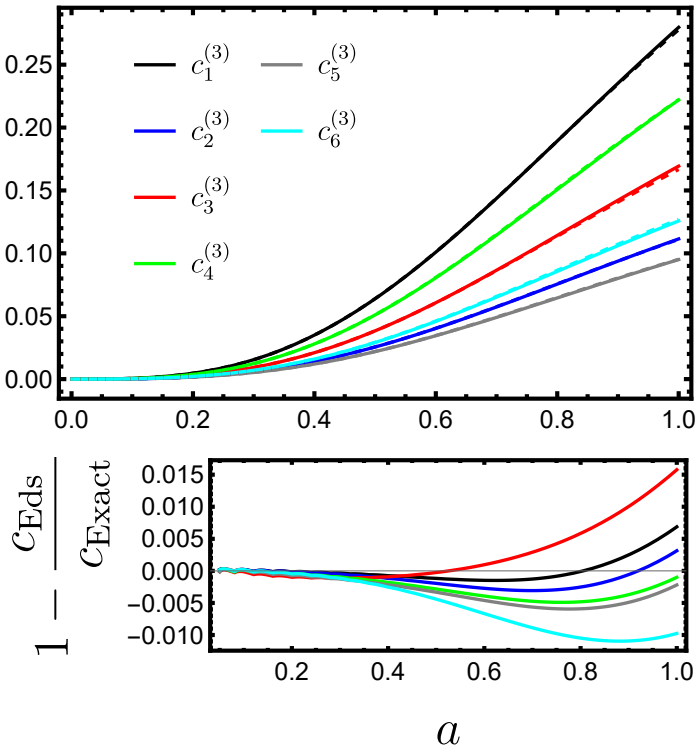

Figure 6.9: Time-dependent coefficients for the third order perturbation theory kernels, in our fiducial cosmology, as a function of the scale factor. Solid and dashed lines (almost indistinguishable) correspond to exact solutions from Eqs. (6.93) and (6.96), and the EdS approximation in Eq. (6.97), respectively. The lower plot shows the relative difference between the two, which is at the percent level.

The formulas presented here are fully equivalent to the many different approaches to

computing time-dependent coefficients in SPT, that can be found on the extensive literature on this subject (see, e.g., [323–334]). Nevertheless, we find Eq. (6.93) to be particularly simple and easy to implement in practice.

6.8 Appendix: Computing second order diagrams

We will now compute all relevant diagrams at second order in perturbation theory, starting with the diagram in Fig. 6.5 for the background distribution function at second order in the perturbative expansion. It represents two insertions of $\phi^{(1)}$ into the second order iterative solution. To compute this, we first substitute the second line of Eq. (6.23) into the first line of Eq. (6.25). This yields

$$\begin{aligned}\bar{f}^{(2\mathrm{nd})}(\eta,q) = &\int_0^{\eta} d\eta' a^2(\eta') \int_0^{\eta'} d\eta'' a^2(\eta'') \int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}} \langle \phi(\eta',\vec{k}_1)\phi(\eta'',\vec{k}_2)\rangle \times \\ &\times i\vec{k}_1 \cdot \frac{\partial}{\partial \vec{q}} \left[i\vec{k}_2 \cdot \frac{\partial \bar{f}^{(0)}}{\partial \vec{q}} e^{-i\vec{k}_2\cdot\vec{q}(\eta'-\eta'')} \right] .\end{aligned} \tag{6.98}$$

We now insert two copies of $\phi^{(1)}$, using Eqs. (6.29), (6.35) and (6.36) to obtain

$$\begin{aligned}\bar{f}^{(2)}(\eta,q) = &-\left(\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\right)^2 \int_0^{\eta} d\eta' a(\eta') D_{\mathrm{L}}(\eta') \int_0^{\eta'} d\eta'' a(\eta'') D_{\mathrm{L}}(\eta'') \times \\ &\times \int \frac{d^3\vec{k}'}{(2\pi)^3} \frac{P_{\mathrm{L}}(k')}{(k')^4} i\vec{k}' \cdot \frac{\partial}{\partial \vec{q}} \left[i\vec{k}' \cdot \frac{\partial \bar{f}^{(0)}}{\partial \vec{q}} e^{i\vec{k}'\cdot\vec{q}(\eta'-\eta'')} \right] .\end{aligned} \tag{6.99}$$

In what follows we omit time dependencies for simplicity. We evaluate next

$$\begin{aligned}&a^{-5} \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}(q) = \\ &= -\left(\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\right)^2 \int_0^{\eta} d\eta' a(\eta') D_{\mathrm{L}}(\eta') \int_0^{\eta'} d\eta'' a(\eta'') D_{\mathrm{L}}(\eta'') \times \\ &\times \int \frac{d^3\vec{k}'}{(2\pi)^3} \frac{P_{\mathrm{L}}(k')}{(k')^4} \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \; i\vec{k}' \cdot \frac{\partial}{\partial \vec{q}} \left[i\vec{k}' \cdot \frac{\partial \bar{f}^{(0)}}{\partial \vec{q}} e^{i\vec{k}'\cdot\vec{q}(\eta'-\eta'')} \right] .\end{aligned} \tag{6.100}$$

The integral over momentum can be simplified via integrating by parts twice, after which we can use $\bar{f}^{(0)}(q) \propto \delta^{(3)}(\vec{q})$ and Eq. (6.3) to arrive at

$$\begin{aligned}&a^{-5} \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}(q) = \\ &= 2a^{-2}\bar{\rho} \int_0^{\infty} \frac{dk'}{2\pi^2} P_{\mathrm{L}}(k') \times \left(\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\right)^2 \int_0^{\eta} d\eta' a(\eta') D_{\mathrm{L}}(\eta') \int_0^{\eta'} d\eta'' a(\eta'') D_{\mathrm{L}}(\eta'') .\end{aligned} \tag{6.101}$$

The final step is to evaluate the time integrals. This is straightforward to do using Eq. (6.37) since the integrands can be written as total derivatives, and the final result reads:

$$a^{-5}\int\frac{d^3\vec{q}}{(2\pi)^3}q^2\bar{f}^{(2)}(q)=\bar{\rho}\,a^2H^2f^2D_{\rm L}^2\int_0^\infty\frac{dk'}{2\pi^2}P_{\rm L}(k')\,. \tag{6.102}$$

This equation was used in the main text to show that the averaged stress tensor indeed vanishes as a result of the perturbative expansion.

We next compute the two diagrams in Fig. 6.4, for the distribution function fluctuation at second order in perturbation theory, to derive Eqs. (6.87) and (6.94). Denoting the diagram on the left of Fig. 6.4 by $\mathcal{D}_1$, and the diagram on the right by $\mathcal{D}_2$, we have:

$$\delta f^{(2)}=\mathcal{D}_1+\mathcal{D}_2\,. \tag{6.103}$$

We start with $\mathcal{D}_1$, where we need to insert one $\phi^{(2)}$ into the first order iterative solution, i.e., the second line in Eq. (6.23). This reads using Eq. (6.29),

$$\mathcal{D}_1=-\frac{3}{2}\Omega_{\rm m,0}H_0^2\int_0^\eta d\eta' a^2(\eta')e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}\delta^{(2)}(\eta',\vec{k})\frac{i\vec{k}}{k^2}\cdot\frac{\partial\bar{f}^{(0)}}{\partial\vec{q}}\,. \tag{6.104}$$

For $\mathcal{D}_2$, we need two insertions of $\phi^{(1)}$ into the second order iterative solution for the distribution function fluctuation, i.e., second line in Eq. (6.25). Combining it with Eqs. (6.23), (6.29) and (6.35) we obtain

$$\begin{aligned}\mathcal{D}_2=&\left(\frac{3}{2}\Omega_{\rm m,0}H_0^2\right)^2\int_0^\eta d\eta' a(\eta')D_{\rm L}(\eta')\int_0^{\eta'}d\eta'' a(\eta'')D_{\rm L}(\eta'')\times\\&\times\int_{\{\vec{k}_1,\vec{k}_2\}}^{\vec{k}}\frac{1}{k_1^2k_2^2}:\!\delta_{\rm L}(\vec{k}_1)\delta_{\rm L}(\vec{k}_2)\!:e^{-i\vec{k}\cdot\vec{q}(\eta-\eta')}i\vec{k}_1\cdot\frac{\partial}{\partial\vec{q}}\left[i\vec{k}_2\cdot\frac{\partial\bar{f}^{(0)}}{\partial\vec{q}}e^{-i\vec{k}_2\cdot\vec{q}(\eta'-\eta'')}\right]\,.\end{aligned} \tag{6.105}$$

To arrive at an integral equation for the density contrast, we need to integrate over momentum according to Eq. (6.3). That is,

$$\delta^{(2)}=\hat{\mathcal{D}}_1+\hat{\mathcal{D}}_2\,, \tag{6.106}$$

where

$$\hat{\mathcal{D}}_i=\frac{1}{a^3\bar{\rho}}\int\frac{d^3\vec{q}}{(2\pi)^3}\mathcal{D}_i\,. \tag{6.107}$$

Since $\bar{f}^{(0)}(q) \propto \delta^{(3)}(\vec{q})$, the momentum integrals can be easily evaluated upon integrating by parts a few times to obtain

$$\hat{\mathcal{D}}_1 = \frac{3}{2}\Omega_{\mathrm{m},0}H_0^2 \int_0^\eta d\eta' a(\eta')(\eta-\eta')\delta^{(2)}(\eta',\vec{k})\,, \tag{6.108}$$

and

$$\begin{aligned}\hat{\mathcal{D}}_2 = &\left(\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\right)^2 \int_0^\eta d\eta' a(\eta') D_{\mathrm{L}}(\eta')(\eta-\eta') \int_0^{\eta'} d\eta'' a(\eta'') D_{\mathrm{L}}(\eta'')(\eta-\eta'') \times h_1^{(2)}(\vec{k}) \\ &+ \left(\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\right)^2 \int_0^\eta d\eta' a(\eta') D_{\mathrm{L}}(\eta')(\eta-\eta')^2 \int_0^{\eta'} d\eta'' a(\eta'') D_{\mathrm{L}}(\eta'') \times h_2^{(2)}(\vec{k})\,,\end{aligned} \tag{6.109}$$

where the scale dependent functions $h_i^{(2)}(\vec{k})$ are defined in Eq. (6.45). The time integrals in Eq. (6.109) can be greatly simplified by using Eq. (6.37), followed by a few integration by parts. The final result is

$$\hat{\mathcal{D}}_2 = \frac{1}{2}D_{\mathrm{L}}^2(\eta)\; h_1^{(2)}(\vec{k}) + \int_0^\eta d\eta' \left(\frac{dD_{\mathrm{L}}}{d\eta'}\right)^2 (\eta-\eta')\; h_2^{(2)}(\vec{k})\,. \tag{6.110}$$

We are now ready to combine Eqs. (6.106), (6.108) and (6.110) to write down the integral equation satisfied by the second order density contrast:

$$\begin{aligned}\delta^{(2)}(\eta,\vec{k}) = &\frac{3}{2}\Omega_{\mathrm{m},0}H_0^2 \int_0^\eta d\eta' a(\eta')(\eta-\eta')\delta^{(2)}(\eta',\vec{k}) + \frac{1}{2}D_{\mathrm{L}}^2(\eta)\; h_1^{(2)}(\vec{k}) \\ &+ \int_0^\eta d\eta' \left(\frac{dD_{\mathrm{L}}}{d\eta'}\right)^2 (\eta-\eta')\; h_2^{(2)}(\vec{k})\,.\end{aligned} \tag{6.111}$$

This can be mapped into a second order differential equation by taking two derivatives with respect to superconformal time,

$$\frac{d^2\delta^{(2)}}{d\eta^2} - \frac{3}{2}\Omega_{\mathrm{m},0}H_0^2 a(\eta)\delta^{(2)}(\eta,\vec{k}) = \frac{1}{2}\frac{d^2}{d\eta^2}D_{\mathrm{L}}^2(\eta)\; h_1^{(2)}(\vec{k}) + \left(\frac{dD_{\mathrm{L}}}{d\eta}\right)^2 \; h_2^{(2)}(\vec{k})\,. \tag{6.112}$$

The final step is to change the time variable from superconformal time η to the scale factor a, to arrive at the differential Eq. (6.87). In terms of the separable ansatz of Eq. (6.44), the source terms are found to be

$$\begin{aligned}s_1^{(2)} &= \frac{1}{a^6H^2}\;\frac{1}{2}\frac{d^2}{d\eta^2}D_{\mathrm{L}}^2 = \frac{D_{\mathrm{L}}^2 f}{a^2}\left[2(1+f) + \frac{d\log H}{d\log a} + \frac{d\log f}{d\log a}\right] \\ s_2^{(2)} &= \frac{1}{a^6H^2}\left(\frac{dD_{\mathrm{L}}}{d\eta}\right)^2 = \frac{D_{\mathrm{L}}^2 f^2}{a^2}\,.\end{aligned} \tag{6.113}$$

The steps involved in this calculation can be summarized as follows: First integrate over momentum to obtain $\hat{\mathcal{D}}$, according to Eq. (6.107), from a given diagram $\mathcal{D}$. Then differentiate it twice with respect to superconformal time followed by a division by a^6H^2 to obtain the source term upon factoring out the scale dependent piece. This prescription was used in the main text to derive Eq. (6.59). However, note that this recipe cannot be applied to diagrams such as the leftmost ones in Figs. 6.4 and 6.6, i.e. with all wiggly lines attached to a single circle, because they contribute to the homogeneous part of the differential equation [for example the left-hand side of Eq. (6.112)], and do not appear as source terms.

We omit explicit calculations of the third order diagrams in Fig. 6.6 since they rely on the exact same tools, but are more tedious to write down as one would expect from a higher order calculation. There are two basic tricks one has to keep in mind, which appeared already at the second order level. First, to integrate by parts over momentum to leverage the fact that $\bar{f}^{(0)}(q) \propto \delta^{(3)}(\vec{q})$. Second, to simplify the integrals over superconformal time by using Eq. (6.37) followed by a few integration by parts.

6.9 Appendix: Layzer-Irvine equation

Here we present a derivation of the cosmic energy Eq. (6.79). Our starting point is Eq. (6.77) for the counterterm contribution to the background distribution function, sourced by short-wavelength fluctuations,

$$\bar{f}_{l,\mathrm{ctr}}(\eta,q) = \int_0^\eta d\eta' a^2(\eta') \left\langle \frac{\partial\phi_s(\eta',\vec{x})}{\partial\vec{x}} \cdot \frac{\partial f_s(\eta',\vec{x},\vec{q})}{\partial\vec{q}} \right\rangle , \tag{6.114}$$

The first step is to multiply Eq. (6.114) by q^2 and integrate over momentum

$$\int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}_{l,\mathrm{ctr}}(\eta,q) = -2\int_0^\eta d\eta' a^6(\eta') \left\langle \frac{\partial\phi_s}{\partial\vec{x}}\Big|_{\eta'} \cdot \vec{\Pi}_s(\eta',\vec{x}) \right\rangle , \tag{6.115}$$

where we integrated by parts once and used Eq. (6.5). Next we integrate by parts inside the spatial average (see [24]), and use Eqs. (6.4) and (6.11) to obtain [26]

$$\int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}_{l,\mathrm{ctr}}(\eta,q) = -2\int_0^\eta d\eta' a^2(\eta') \left\langle \phi_s(\eta',\vec{x}) \frac{\partial}{\partial\eta} \left[a(\eta') \frac{\nabla^2\phi_s(\eta',\vec{x})}{4\pi G} \right] \right\rangle . \tag{6.116}$$

[26] All equations which are linear in the distribution function are satisfied separately by both long and short-wavelength pieces.

Integrating by parts inside the spatial average once again leads to,

$$\int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}_{l,\mathrm{ctr}}(\eta, q) = \int_0^\eta d\eta' a(\eta') \frac{d}{d\eta'} \left\{ a^2(\eta') \left\langle \left[\frac{\vec{\nabla}\phi_s(\eta', \vec{x})}{4\pi G} \right]^2 \right\rangle \right\} . \tag{6.117}$$

Now note that, using Eqs. (6.11) and (6.80),

$$\left\langle \left[\frac{\vec{\nabla}\phi_s(\eta', \vec{x})}{4\pi G} \right]^2 \right\rangle = - \left\langle \frac{\phi_s(\eta', \vec{x}) \nabla^2 \phi_s(\eta', \vec{x})}{4\pi G} \right\rangle = -2a^2(\eta') \bar{\rho}(\eta') u(\eta') \, , \tag{6.118}$$

which yields

$$\int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}^{(2)}_{l,\mathrm{ctr}}(\eta, q) = -2a^3(\eta) \bar{\rho}(\eta) \int_0^\eta d\eta' a(\eta') \frac{d}{d\eta'} \left[a(\eta') u(\eta') \right] \, , \tag{6.119}$$

where we used the relation $a^3(\eta')\bar{\rho}(\eta') = a^3(\eta)\bar{\rho}(\eta)$ to bring this quantity out of the integral. This may now be combined with the definition for the short scale kinetic energy per unit mass, Eq. (6.78), to produce the desired result

$$\kappa(\eta) = -a^{-2}(\eta) \int_0^\eta d\eta' a(\eta') \frac{d}{d\eta'} \left[a(\eta') u(\eta') \right] \, . \tag{6.120}$$

6.10 Appendix: One-loop power spectrum

In Appendix 6.7 we include explicit formulas for the full time-dependent coefficients entering SPT kernels in ΛCDM. These can be combined with the scale dependencies shown in Eqs. (6.44) and (6.45) at second order, and Eqs. (6.46) and (6.47) at third order, to compute perturbative predictions for cosmological observables of interest. Here we provide explicit formulas, for the one-loop power spectrum, for the reader's convenience.

The (equal-time) power spectrum, $P(a, k)$ is defined by:

$$\langle \delta(a, \vec{k}) \delta(a, \vec{k}') \rangle = (2\pi)^3 \delta^{(3)}(\vec{k} + \vec{k}') P(a, k) \, . \tag{6.121}$$

From the perturbative expansion Eq. (6.30), the one-loop power spectrum in SPT reads,

$$P_{\text{1-loop}}(a, k) = D_{\mathrm{L}}^2(a) P_{\mathrm{L}}(k) + P_{13}(a, k) + P_{22}(a, k) \, . \tag{6.122}$$

We recall that $P_{\rm L}(k)$ is the linear power spectrum today and $D_{\rm L}(a=1)=1$ sets the linear growth factor normalization [see Eq. (6.40)]. The other two contributions will be flashed out in what follows. We first have, from Eqs. (6.35) and (6.46),

$$P_{13}(a,k) = 2D_{\rm L}(a)\sum_{i=1}^{6} c_i^{(3)}(a)\langle\delta_{\rm L}(\vec{k})h_i^{(3)}(-\vec{k})\rangle' \equiv 2D_{\rm L}(a)\sum_{i=1}^{6} c_i^{(3)}(a)\Gamma_i(k)\,, \tag{6.123}$$

where $\Gamma_i(k) = \langle\delta_{\rm L}(\vec{k})h_i^{(3)}(-\vec{k})\rangle'$ and the primed correlation function has the momentum conserving Dirac delta stripped off, following standard convention. Formulas for the time-dependent coefficients $c_i^{(3)}(a)$ can be found in Appendix 6.7, and from Eq. (6.47) we obtain:

$$\begin{aligned}
\Gamma_1(k) &= \Gamma_2(k) = -\frac{1}{3}k^3P_{\rm L}(k)\int_0^\infty \frac{dx}{2\pi^2}P_{\rm L}(kx)(1+x^2)\\
\Gamma_3(k) &= -\frac{1}{8}k^3P_{\rm L}(k)\int_0^\infty \frac{dx}{2\pi^2}P_{\rm L}(kx)\left[2(x^4-4x^2-1)+\frac{(x^2-1)^3}{x}\log\left|\frac{x-1}{x+1}\right|\right]\\
\Gamma_4(k) &= \frac{1}{3}k^3P_{\rm L}(k)\int_0^\infty \frac{dx}{2\pi^2}P_{\rm L}(kx)x^2\\
\Gamma_5(k) &= -\frac{1}{16}k^3P_{\rm L}(k)\int_0^\infty \frac{dx}{2\pi^2}P_{\rm L}(kx)\frac{1}{x^3}\left[2x(x^4+4x^2-1)+(x^2-1)^3\log\left|\frac{x-1}{x+1}\right|\right]\\
\Gamma_6(k) &= -\frac{1}{6}k^3P_{\rm L}(k)\int_0^\infty \frac{dx}{2\pi^2}P_{\rm L}(kx)\,.
\end{aligned} \tag{6.124}$$

In the EdS approximation the time-dependent coefficients simplify to Eq. (6.97), and from Eq. (6.123) we arrive at

$$\begin{aligned}
P_{13}(a,k) \approx &\frac{1}{84\pi^2}D_{\rm L}^4(a)k^3P_{\rm L}(k)\int_0^\infty dx\frac{P_{\rm L}(kx)}{x^2}\left[1-\frac{79}{6}x^2+\frac{25}{3}x^4\right.\\
&\left.-\frac{7}{2}x^6-\frac{1}{2x}(x^2-1)^3(1+\frac{7}{2}x^2)\log\left|\frac{x-1}{x+1}\right|\right]\,.
\end{aligned} \tag{6.125}$$

We similarly need, from Eq. (6.44),

$$P_{22}(a,k) = \sum_{i=1}^{2}\sum_{j=1}^{2} c_i^{(2)}(a)c_j^{(2)}(a)\langle h_i^{(2)}(\vec{k})h_j^{(2)}(-\vec{k})\rangle' \equiv \sum_{i=1}^{2}\sum_{j=1}^{2} c_i^{(2)}(a)c_j^{(2)}(a)\Sigma_{ij}(k)\,, \tag{6.126}$$

where $\Sigma_{ij}(k) = \langle h_i^{(2)}(\vec{k})h_j^{(2)}(-\vec{k})\rangle'$ [see Appendix 6.7 for the time-dependent coefficients

$c_i^{(2)}(a)]$. We obtain from Eq. (6.45)

$$\begin{aligned}
\Sigma_{11}(k) &= \frac{1}{2}k^3\int_0^\infty \frac{dx}{2\pi^2}\int_{-1}^1 \frac{dt}{2}\frac{(t+x-2xt^2)^2}{(1-2xt+x^2)^2}P_{\rm L}(kx)P_{\rm L}\left(k\sqrt{1-2xt+x^2}\right)\\
\Sigma_{12}(k) &= \Sigma_{21}(k) = \frac{1}{2}k^3\int_0^\infty \frac{dx}{2\pi^2}\int_{-1}^1 \frac{dt}{2}\frac{(t+x-2xt^2)(t-x)}{(1-2xt+x^2)^2}P_{\rm L}(kx)P_{\rm L}\left(k\sqrt{1-2xt+x^2}\right)\\
\Sigma_{22}(k) &= \frac{1}{2}k^3\int_0^\infty \frac{dx}{2\pi^2}\int_{-1}^1 \frac{dt}{2}\frac{(t-x)^2}{(1-2xt+x^2)^2}P_{\rm L}(kx)P_{\rm L}\left(k\sqrt{1-2xt+x^2}\right)\,.
\end{aligned} \tag{6.127}$$

In the EdS approximation the time-dependent coefficients simplify to Eq. (6.95), and from Eq. (6.126) we arrive at

$$P_{22}(a,k) \approx \frac{1}{392\pi^2}D_{\rm L}^4(a)k^3\int_0^\infty dx\int_{-1}^1 dt\frac{(7t+3x-10t^2x)^2}{(1-2xt+x^2)^2}P_{\rm L}(kx)P_{\rm L}\left(k\sqrt{1-2xt+x^2}\right)\,. \tag{6.128}$$

6.11 Appendix: Tree-level bispectrum

In this section we explicitly write down the relevant formulas for the tree-level bispectrum in SPT, with the full time dependencies of ΛCDM kernels. The bispectrum $B(k_1,k_2,k_3)$ is defined by

$$\langle\delta(\vec{k_1})\delta(\vec{k_2})\delta(\vec{k_3})\rangle = (2\pi)^3\delta^{(3)}(\vec{k_1}+\vec{k_2}+\vec{k_3})B(k_1,k_2,k_3)\,. \tag{6.129}$$

From Eqs. (6.30), (6.35), and (6.44) the tree-level bispectrum in SPT reads

$$B(k_1,k_2,k_3) = D_{\rm L}^2(a)\sum_{i=1}^2 c_i^2(a)\langle h_i^{(2)}(\vec{k}_1)\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3)\rangle' + \dots\,, \tag{6.130}$$

where the ellipsis represent two additional terms corresponding to permutations of the three wavenumbers. The scale dependent part $\langle h_i^{(2)}(\vec{k}_1)\delta_{\rm L}(\vec{k}_2)\delta_{\rm L}(\vec{k}_3)\rangle'$ can be computed from Eq. (6.45), and Eq. (6.130) becomes

$$B(k_1,k_2,k_3) = 2D_{\rm L}^2(a)\left[c_1^{(2)}(a)\alpha^{(s)}(\vec{k}_1,\vec{k}_2) + c_2^{(2)}(a)\beta(\vec{k}_1,\vec{k}_2)\right]P_{\rm L}(k_1)P_{\rm L}(k_2) + \dots\,, \tag{6.131}$$

where

$$\begin{aligned}\alpha^{(s)}(\vec{k}_1,\vec{k}_2) &= \frac{1}{2}\frac{\vec{k}_1\cdot\vec{k}_{12}}{k_1^2}+\frac{1}{2}\frac{\vec{k}_2\cdot\vec{k}_{12}}{k_2^2}\\ \beta(\vec{k}_1,\vec{k}_2) &= \frac{k_{12}^2(\vec{k}_1\cdot\vec{k}_2)}{2k_1^2k_2^2}\,,\end{aligned} \tag{6.132}$$

with $\vec{k}_{12}=\vec{k}_1+\vec{k}_2$, and formulas for the time-dependent coefficients $c_i^{(2)}(a)$ can be found in Appendix 6.7. In the EdS approximation they simplify to Eq. (6.95), and from Eq. (6.131) we arrive at

$$B(k_1,k_2,k_3)\approx 2D_{\rm L}^4(a)F_2(\vec{k}_1,\vec{k}_2)P_{\rm L}(k_1)P_{\rm L}(k_2)+\dots\,, \tag{6.133}$$

where

$$\begin{aligned}F_2(\vec{k}_1,\vec{k}_2) &= \frac{5}{7}\alpha^{(s)}(\vec{k}_1,\vec{k}_2)+\frac{2}{7}\beta(\vec{k}_1,\vec{k}_2)\\ &= \frac{5}{7}+\frac{1}{2}\frac{\vec{k}_1\cdot\vec{k}_2}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)+\frac{2}{7}\frac{(\vec{k}_1\cdot\vec{k}_2)^2}{k_1^2k_2^2}\,.\end{aligned} \tag{6.134}$$

Chapter 7

A SEMI-ANALYTIC ESTIMATE FOR THE EFFECTIVE SOUND SPEED COUNTERTERM IN THE EFTOFLSS

7.1 Abstract

The Effective Field Theory of Large Scale Structure (EFTofLSS) has found tremendous success as a perturbative framework for the evolution of large scale structure, and it is now routinely used to compare theoretical predictions against cosmological observations. The model for the total matter field includes one nuisance parameter at 1-loop order, the effective sound speed, which can be extracted by matching the EFT to full N-body simulations. In this work we first leverage the Layzer-Irvine cosmic energy equation to show that the equation of state can be exactly computed with knowledge of the fully nonlinear power spectrum. When augmented with separate universe methods, we show one can estimate the effective sound speed. This estimate is in good agreement with simulation results, with errors at the few tens of percent level. We apply our method to investigate the cosmology dependence of the effective sound speed and to shed light on what cosmic structures shape its value.

7.2 Introduction

In the past fifteen years the Effective Field Theory of Large Scale Structure (EFTofLSS) [42, 43, 273–276] has proved itself to be a very powerful framework, enabling one to push the regime of validity of standard perturbation theory (SPT) methods [41, 133, 235–238] for the large-scale structure evolution to smaller scales [279–281]. The EFTofLSS is now an integral part of theory modeling involved in the analysis of real data [53, 248–256].

While cosmological simulations are strictly necessary to ensure accuracy of modeling on sufficiently small scales [244] where the nonlinear gravitational evolution becomes strongly

coupled, perturbative methods offer a number of advantages and are thus complementary to simulations. They provide the kind of intuition that only an analytic model can offer, with flexibility to go beyond standard scenarios [239–243], and circumvent the need for extensive computational resources that simulations require [37]. Additionally, semi-analytic methods can provide insights on the nonperturbative regime of nonlinear structure formation as well (e.g. [353–357]).

The EFTofLSS was created to fix two major shortcomings of SPT: To accommodate deviations from the assumption of a pressureless fluid [1] and to address the issue of sensitivity to uncontrolled short distance physics [2]. At one-loop, this comes at the cost of introducing a new free parameter, the effective sound speed counterterm, which is to be determined either by observations or through a matching to N-body simulations. The one-loop power spectrum in the EFT reads

$$P_{\text{1-loop}}(a,k;\Lambda) = P_{\text{1-loop,SPT}}(a,k;\Lambda) - D_{\text{L}}(a)c_{\text{ctr}}(a;\Lambda)k^2W_\Lambda^2(k)P_{\text{L}}(k)\,, \tag{7.1}$$

where $P_{\text{1-loop,SPT}}(a,k;\Lambda)$ is the one-loop power spectrum in SPT, $D_{\text{L}}(a)$ is the linear growth factor, $P_{\text{L}}(k)$ is the linear theory power spectrum, $W_\Lambda(k)$ is a window function introduced to smooth the density field over some distance scale Λ^{-1}, and $c_{\text{ctr}}(a;\Lambda)$ is the counterterm [which is directly related to the effective sound speed of dark matter when treated as a fluid, $c^2_{\text{eff}}(a;\Lambda)$, as detailed in Appendix 7.8, see Eq. (7.63)]. One eventually takes the $\Lambda \to \infty$ limit, which removes the sensitivity to a choice of window function.

In this paper we focus on the effective sound speed parameter, which enters in all applications of the EFTofLSS. However, additional free parameters become necessary when considering higher order terms in the perturbative expansion [282], when modeling biased tracers [44] and also for higher point correlation functions (at the same loop precision) [283–285]. We will consider the simplest case of one-loop perturbation theory to describe the

[1] Even collisionless systems have a nonzero pressure induced by short-wavelength fluctuations. This is exactly the subject of study in this manuscript.

[2] As can be seen from the fact that modes with arbitrarily high frequencies are running on the loops in SPT.

two-point clustering of matter. This is sufficient as a perturbative description of weak gravitational lensing on sufficiently large scales, but not of galaxy clustering which also requires modeling of galaxy formation physics (even on large scales via a bias expansion). We will also work in real space, i.e., ignoring redshift space distortions [358].

The effective sound speed is expected to account for the backreaction on large scales of nonlinear effects associated to gravitational collapse and the subsequent formation of the cosmic web in the form of cosmic sheets, filaments and halos [310, 317, 318, 359]. The counterterm has been successfully extracted from matching EFT predictions to simulations in a number of previous works [43, 49, 356, 360–367], but such measurements do not offer any clear physical interpretation for the underlying microphysics responsible for the effective sound speed. In particular, a simple estimate of the numerical value of that quantity from some semi-analytic model for the short scale fluctuations is still lacking (beyond just dimensional analysis, e.g. see [359]). This illustrates the point that the microphysics associated to the EFT counterterm is yet to be fully understood. Such an understanding would shed light on the cosmology dependence of the sound speed parameter, and on what cosmic structures shape its value.

In this work we will make progress in that direction by providing a simple semi-analytic estimate for the effective sound speed, which we show is in very good agreement with simulation results. We first show that we can leverage the Layzer-Irvine cosmic energy equation to exactly compute the equation of state, i.e. the ratio between background pressure and energy density, assuming knowledge of the fully nonlinear matter power spectrum. Since the sound speed is related to the response of the equation of state to the presence of a long-wavelength mode, this introduces the necessary ingredients to estimate the effective sound speed using separate universe techniques. The estimate reproduces the results from high resolution simulations with errors at the tens of percent level.

We will work with a ΛCDM universe, particularly its late time dynamics (redshifts $z \lesssim 100$) on subhorizon scales $k \gg aH$ [where $a(t)$ is the cosmological scale factor, $H = d\log a/dt$ is the Hubble expansion rate and t is cosmic time]. In this regime one can write Newtonian

equations of motion in an expanding cosmological background, and the collective dynamics of cold dark matter particles is encoded in the collisionless Boltzmann equation coupled to the Poisson equation, the so-called Vlasov-Poisson system.

Unless explicitly stated otherwise, for numerical calculations we assume a fiducial cosmology $\Lambda CDM_{\rm DarkSky}$ with $\Omega_{\rm m} = 0.295$, $\Omega_\Lambda = 0.705$, $h = 0.688$, $n_s = 0.9676$ and $\sigma_8 = 0.835$, matching the choice made in [360], which adopts the Dark Sky simulations [368]. All power spectra are computed using the Boltzmann solver CLASS [47], in combination with HMcode [201] for the nonlinear power spectrum when needed (without baryonic feedback effects), at 16 equally spaced redshift values in the interval $0 \leq z \leq 5$. We interpolate in-between and for $z > 5$ assume a linear-theory evolution in all cases, i.e., $P(z,k) = [D_{\rm L}(z)/D_{\rm L}(z = 5)]^2 P(z = 5, k)$, where $D_{\rm L}(a)$ is the linear growth factor (we use the fitting formula provided in [322]) [3]. Although we only numerically compute quantities of interest at $z \lesssim 2$, integrals over the scale factor appear that require the nonlinear power spectrum out to arbitrarily high redshift. However, we have checked that when computing the equation of state and the effective sound speed, contributions from $z > 5$ are essentially negligible, which justifies our simple linear-theory scaling.

The paper is organized as follows: In Sec. 7.3 we review the standard EFTofLSS framework to set notation and collect some important results. In Sec. 7.4 we use the Layzer-Irvine equation to compute the equation of state assuming knowledge of the fully nonlinear power spectrum. In Sec. 7.5 we apply the tools developed in the previous section, in combination with separate universe methods, to derive an estimate for the effective sound speed which we compare to simulation results. In Sec. 7.6 we summarize our findings. Appendices 7.7, 7.8 and 7.9 supplement the main text with relevant discussions and derivations.

[3] HMcode is calibrated against simulations at $z \leq 2$. In that redshift range, it is accurate out to $k = 10h{\rm Mpc}^{-1}$ at the few percent level [201]. We also used the EuclidEmulator2 (EE2), which is accurate at the percent level in the same range of scales and redshifts [46], to cross-check our results. We obtained nearly identical results using either HMcode or EE2.

7.3 EFTofLSS framework

In this section we review the EFTofLSS framework, mostly following the phase-space approach of [43], to establish notation and recollect some well-known results that will be used in later sections.

The collective behavior of particles interacting only via the gravitational force in an expanding universe is governed by the collisionless Boltzmann (or Vlasov) equation:

$$\frac{\partial f}{\partial \eta} + \frac{d\vec{x}}{d\eta} \cdot \frac{\partial f}{\partial \vec{x}} + \frac{d\vec{q}}{d\eta} \cdot \frac{\partial f}{\partial \vec{q}} = 0 \,, \tag{7.2}$$

for the phase space distribution function $f(\eta, \vec{x}, \vec{q})$. We work with the superconformal time defined by $d\eta = dt/a^2(t)$ (sometimes we also use the scale factor as the time variable). Additionally, $\vec{x}$ are comoving coordinates and $\vec{q}$ is the comoving momentum such that $(d\vec{x}/d\eta) = a^2(d\vec{x}/dt) = (\vec{q}/m) \equiv \vec{q}$, with m the cold dark matter (CDM) particle mass which we set to unity from this point forward. We then have $(d\vec{q}/d\eta) = -ma^2\vec{\nabla}\phi = -a^2\vec{\nabla}\phi$, with $\phi(\eta, \vec{x})$ the Newtonian gravitational potential. Eq. (7.2) now reads,

$$\frac{\partial f}{\partial \eta} + \vec{q} \cdot \frac{\partial f}{\partial \vec{x}} = a^2(\eta) \frac{\partial \phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{q}} \,. \tag{7.3}$$

The gravitational potential is not independent from the distribution function, but rather determined by the Poisson equation

$$\nabla^2 \phi = 4\pi G a^2 (\rho - \bar{\rho}) \,, \tag{7.4}$$

where

$$\rho(\eta, \vec{x}) = a^{-3}(\eta) \int \frac{d^3\vec{q}}{(2\pi)^3} f(\eta, \vec{x}, \vec{q}) \,, \tag{7.5}$$

is the energy density. [4] Note that only the fluctuations around the (ensemble) average density, $\bar{\rho}(\eta) = \langle \rho(\eta, \vec{x}) \rangle \propto a^{-3}(\eta)$, contribute to the gravitational potential.

The nonlinear term on the right-hand side of Eq. (7.3) is a contact term, [5] being sensitive to nonlinearities in gravitational dynamics at arbitrarily small scales. In order to have

[4] In our convention of unity CDM particle mass, this becomes the number density. Similar statements can be made on the other fluid quantities as well.

[5] By contact term we mean a product of two fields evaluated at the same point in space.

theoretical control over the backreaction from the uncertain short scale modes, we follow the standard EFT approach and split the distribution function into a long-wavelength piece defined by smoothing over some (distance) scale $1/\Lambda$

$$f_l(\eta,\vec{x},\vec{q}) = \int d^3\vec{x}'\, W_\Lambda(|\vec{x}-\vec{x}'|) f(\eta,\vec{x}',\vec{q})\,, \tag{7.6}$$

and a short-wavelength piece $f_s = f - f_l$, via a window function $W_\Lambda(x)$ [6]. A similar decomposition, $\phi = \phi_l + \phi_s$, also follows for the gravitational potential since it is linearly related to the distribution function by Eqs. (7.4) and (7.5).

The evolution equation for the long-wavelength piece now follows from Eqs. (7.3) and (7.6)

$$\frac{\partial f_l}{\partial \eta} + \vec{q}\cdot\frac{\partial f_l}{\partial \vec{x}} = a^2(\eta)\int d^3\vec{x}'\, W_\Lambda(|\vec{x}-\vec{x}'|)\frac{\partial \phi(\eta,\vec{x}')}{\partial \vec{x}'}\cdot\frac{\partial f(\eta,\vec{x}',\vec{q})}{\partial \vec{q}}\,. \tag{7.7}$$

Following standard methods (e.g., see [42]) one obtains the result that mixed terms involving a product of long and short modes are suppressed, and hence

$$\frac{\partial f_l}{\partial \eta} + \vec{q}\cdot\frac{\partial f_l}{\partial \vec{x}} = a^2(\eta)\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}} + a^2(\eta)\left[\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right]_\Lambda + O\Big(\frac{k^2}{\Lambda^2}\Big)\,, \tag{7.8}$$

where,

$$\left[\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right]_\Lambda = \int d^3\vec{x}'\, W_\Lambda(|\vec{x}-\vec{x}'|)\frac{\partial \phi_s(\eta,\vec{x}')}{\partial \vec{x}'}\cdot\frac{\partial f_s(\eta,\vec{x}',\vec{q})}{\partial \vec{q}}\,. \tag{7.9}$$

From this point forward we will drop the higher derivative corrections of order $\mathcal{O}(k^2/\Lambda^2)$. They vanish in the $\Lambda\to\infty$ limit, and do not contribute to the effective sound speed even at finite Λ. However, it is important to keep in mind that their presence become necessary to accurately model the nonlinear power spectrum for any finite value of the cutoff [42, 43]. In the remainder of this section we delve into the effective stress tensor in the EFTofLSS.

---

[6] For numerical calculations we follow the standard convention in the EFTofLSS literature and consider a Gaussian filter, i.e., $W_\Lambda(k) = e^{-k^2/2\Lambda^2}$ in Fourier space. We assume a time-independent filter. One could think of working with a time-dependent filter, since more modes are perturbative at higher redshift, as follows from the time-dependence of the scale of nonlinearities. However, this adds unnecessary technical complications to the model since the ultimate goal is to take the $\Lambda\to\infty$ limit.

7.3.1 Effective stress tensor

By taking the first two moments of Eq. (7.8), one can derive the dynamical equations for the long-wavelength effective fluid. This calculation can be found in the literature [43, 365] and leads to the continuity and Euler equations. The former corresponds to conservation of mass, and reads

$$\rho_l' + 3\mathcal{H}\rho_l + a\vec{\nabla}\cdot\vec{\Pi}_l = 0\,, \tag{7.10}$$

where prime denotes a derivative with respect to superconformal time, $\mathcal{H} = d\log a/d\eta = a^2 H$, and

$$\vec{\Pi}_l(\eta,\vec{x}) = a^{-4}(\eta)\int \frac{d^3\vec{q}}{(2\pi)^3}\,\vec{q}\,f_l(\eta,\vec{x},\vec{q})\,, \tag{7.11}$$

is the fluid momentum. The Euler equation is

$$v_{l,i}' + \mathcal{H}v_{l,i} + av_{l,j}\partial^j v_{l,i} + a\partial_i\phi_l + \frac{a}{\rho_l}\partial^j\tau_{ij}^{\rm eff} = 0\,, \tag{7.12}$$

where

$$\vec{\Pi}_l(\eta,\vec{x}) = \rho_l(\eta,\vec{x})\vec{v}_l(\eta,\vec{x})\,, \tag{7.13}$$

defines the long-wavelength velocity field $\vec{v}_l(\eta,\vec{x})$. The effective stress tensor is a sum of kinetic and potential terms:

$$\tau_{ij}^{\rm eff}(\eta,\vec{x}) = 2K_{ij}(\eta,\vec{x}) + U_{ij}(\eta,\vec{x})\,. \tag{7.14}$$

The kinetic part is

$$\begin{aligned} K_{ij}(\eta,\vec{x}) &= \frac{1}{2}a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}f_l(\eta,\vec{x},\vec{q})\left[q_i - a(\eta)v_{li}(\eta,\vec{x})\right]\left[q_j - a(\eta)v_{lj}(\eta,\vec{x})\right] \\ &= \frac{1}{2}\left[a^{-5}(\eta)\int\frac{d^3\vec{q}}{(2\pi)^3}f_l(\eta,\vec{x},\vec{q})q_iq_j - \rho_l(\eta,\vec{x})v_{li}(\eta,\vec{x})v_{lj}(\eta,\vec{x})\right]\,, \end{aligned} \tag{7.15}$$

where the second equality follows from Eqs. (7.11) and (7.13). The potential part reads

$$U_{ij}(\eta,\vec{x}) = \frac{1}{4\pi G a^2(\eta)}\left\{\left[\partial_i\phi_s\partial_j\phi_s\right]_\Lambda - \frac{1}{2}\left[\partial_k\phi_s\partial^k\phi_s\right]_\Lambda\delta_{ij}\right\}\,. \tag{7.16}$$

The kinetic contribution in Eq. (7.15) is sourced by the velocity dispersion while the potential contribution in Eq. (7.16) derives from the short scale gravitational potential. Let us

introduce some notation for the average of these quantities. The short-scale kinetic energy per unit mass, $\kappa(\eta)$, is defined as

$$\bar{\rho}(\eta)\kappa(\eta) \equiv \langle K_i^i(\eta, \vec{x})\rangle \,. \tag{7.17}$$

Similarly the short-scale gravitational binding energy per unit mass, $u(\eta)$, is defined as

$$\bar{\rho}(\eta)u(\eta) \equiv \langle U_i^i(\eta, \vec{x})\rangle \,, \tag{7.18}$$

in terms of ensemble averages of kinetic and potential contributions to the effective stress. The traceless parts of $K_{ij}(\eta, \vec{x})$ and $U_{ij}(\eta, \vec{x})$ vanish upon averaging due to statistical isotropy. Note from the second line in Eq. (7.15), as well as Eq. (7.17), that the kinetic energy from the bulk flow is subtracted from the total kinetic energy in the definition of $\kappa(\eta)$, with only the contribution from velocity dispersion remaining, hence we refer to it as the short-scale kinetic energy per unit mass. Similarly, from Eqs. (7.16) and (7.18) it follows that

$$\bar{\rho}u = -\frac{1}{8\pi G a^2}\langle \vec{\nabla}\phi_s \cdot \vec{\nabla}\phi_s\rangle = \frac{1}{2}\langle \phi_s \rho_s\rangle \implies u = \frac{1}{2}\langle \phi_s \delta_s\rangle \,, \tag{7.19}$$

where we integrate by parts inside the ensemble average [7], using the Poisson Eq. (7.4), and from now on we often omit time and scale dependencies to simplify the notation. This shows that $u(\eta)$ is indeed the short-scale gravitational binding energy. Before showing how to compute this quantity we first introduce the Friedmann equation:

$$H^2 = \frac{8\pi G}{3}\rho_{\rm cri} = \frac{8\pi G}{3}\frac{\bar{\rho}}{\Omega_{\rm m}(a)} \,, \tag{7.20}$$

where $\Omega_{\rm m}(a) = \bar{\rho}(a)/\rho_{\rm cri}(a)$ is the fractional contribution of matter to the total (critical) energy density. Evaluating Eq. (7.20) at the present time, and combining it with Eq. (7.4), allows us to write the Poisson equation in the following form in Fourier space ($\vec{\nabla} \to i\vec{k}$)

$$k^2\phi = -\frac{3}{2}\Omega_{\rm m,0}H_0^2\frac{\delta}{a} \,, \tag{7.21}$$

[7] The ensemble average can always be thought of as a volume average, which involves integrating over position. This justifies an integration by parts under the average sign. Also note that since these are global averages, the integral over positions are defined over the entire volume and hence boundary terms can be assumed negligible with appropriate boundary conditions at spatial infinity.

where we used $\bar{\rho} \propto a^{-3}$, and quantities evaluated at the present time carry a subscript 0, such as the Hubble expansion rate today $H_0 = 100h(\mathrm{km/s})/\mathrm{Mpc}$. We can now substitute Eq. (7.21) into Eq. (7.19) to arrive at

$$u(a) = -\frac{3}{4}\Omega_{\mathrm{m},0}H_0^2\frac{1}{a}\int\frac{dk}{2\pi^2}P(a,k)\left[1-W_\Lambda(k)\right]^2\,, \tag{7.22}$$

where

$$\langle\delta(a,\vec{k})\delta(a,\vec{k'})\rangle = (2\pi)^3\delta^{(3)}(\vec{k}+\vec{k'})P(a,k)\,, \tag{7.23}$$

defines the fully nonlinear matter power spectrum $P(a,k)$, which Eq. (7.22) shows is all one needs to compute the short-scale gravitational binding energy per unit mass. In Fig. 7.1 we plot $kP(k)$, which is integrated over $\log k$ in Eq. (7.22), for a few redshift values. Solid and dashed lines correspond to the fully nonlinear and linear theory power spectra, respectively. It is clear that at $z = 0$, $kP(k)$ has a second peak at a small scale of $k \sim 2h\mathrm{Mpc}^{-1}$ in the nonlinear case. This second peak is less pronounced at higher redshift and eventually disappears as the linear and nonlinear power spectra start to converge toward one another. For that reason, the integral in Eq. (7.22) picks up significant contributions from scales of a few $\sim h\mathrm{Mpc}^{-1}$ at low redshift. This shows that the gravitational binding energy picks-up large contributions from virialized halos.

In Fig. 7.2 we plot the short-scale gravitational binding energy as a function of the smoothing scale $R = \Lambda^{-1}$, for a few redshift values. The round data points were extracted from averaging the two phase-reversed runs of the dark-matter-only MilleniumTNG simulations [369]. The simulations contain 2160^3 dark matter particles in a $L_{\mathrm{box}} = 500$ Mpc h^{-1} box [8]. Since we do not have access to many realizations of the simulation, we do not include errors bars due to cosmic variance (which is a source of uncertainty on the largest scales, but is mitigated by averaging over the two phase-reversed runs). Solid lines are the theory

[8] We compute the binding energies from the simulation snapshots using an Eulerian density mesh on high-resolution 3D grid of 2048^3 voxels, constructed using the cubic spline assignment scheme. We Fourier transform the density field, solve the Poisson equation for the modes of the potential, and inverse Fourier transform the potential modes back to real space. Then we smooth both the original density field and the Newtonian potential with Gaussian filters. By subtracting the smoothed fields from the original fields, we can compute the mean binding energy by averaging over the grid, using Eq. (7.19).

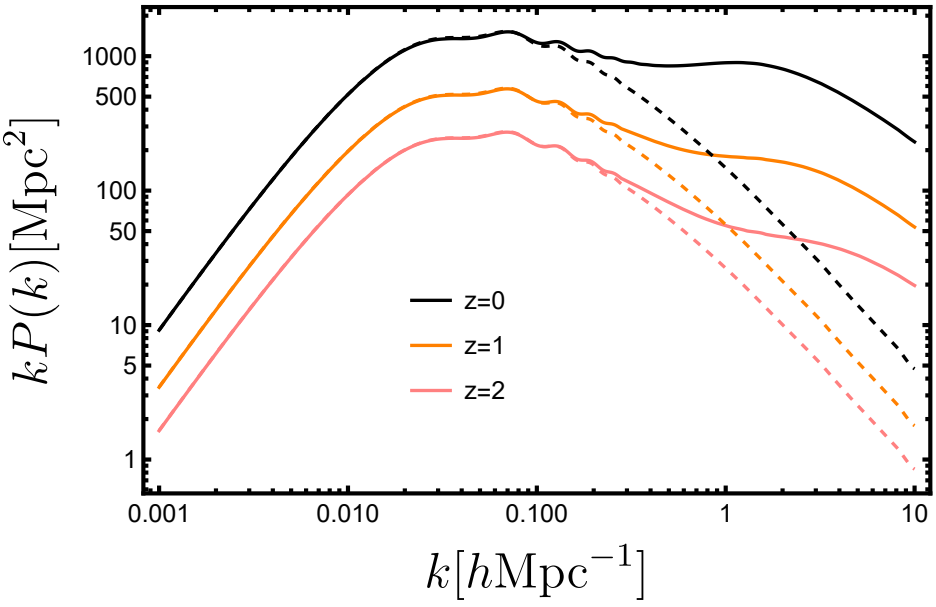

Figure 7.1: Plot of $kP(k)$, which is integrated over $\log k$ in Eq. (7.22), as a function of the comoving wavenumber k. Solid lines show the fully nonlinear matter power spectrum and dashed lines correspond to the linear theory matter power spectrum. Note the presence of a second peak at $k \sim 1h\mathrm{Mpc}^{-1}$, which causes the short-scale gravitational binding energy in Eq. (7.22) to pick up sizable contributions from very small scales of a few $\sim h\mathrm{Mpc}^{-1}$, typical of virialized halos.

predictions, from Eq. (7.22), using HMcode to compute the nonlinear power spectrum, and also adding a hard infrared cutoff at the fundamental mode of the simulation $k_{\rm F} = 2\pi/L_{\rm box}$. The dashed line is the theory prediction integrated over all wavenumber, and the dot-dashed line uses Halofit [370] to compute the nonlinear power spectrum, and also adds a hard cutoff at the fundamental mode. The theory and simulation results are in excellent agreement, with Halofit perfoming slightly better than HMcode.

We will see next that the short-scale kinetic energy can be computed from $u(a)$, which is hence the basic ingredient needed in the calculation of the average effective stress tensor, and also of the effective sound speed as we will see in Sec.7.5.

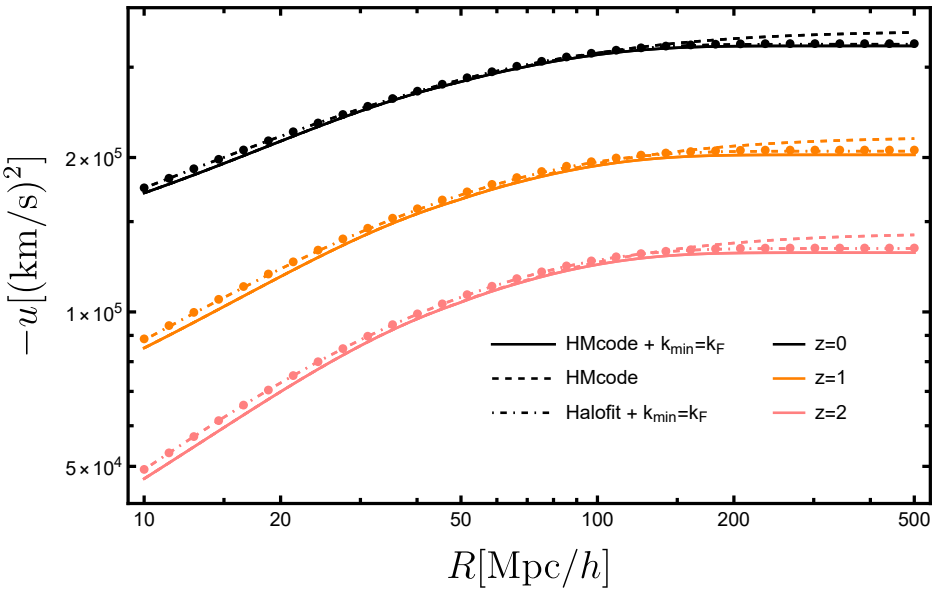

Figure 7.2: Short-scale gravitational binding energy per unit mass as a function of the smoothing scale $R = \Lambda^{-1}$, for varying redshift. Round data points correspond to results from the dark-matter-only runs of the MilleniumTNG simulations [369]. Solid, dashed and dot-dashed lines correspond to different theory predictions from Eq. (7.22), differing on the choice of code used to model the nonlinear power spectrum and whether or not a hard cutoff is introduced at the fundamental mode of the simulation (as indicated in the legends). There is excellent agreement between theory and simulation results, with Halofit performing slightly better than HMcode.

7.3.2 Cosmic energy equation

In the previous section we established that the short-scale gravitational binding energy can be straightforwardly calculated assuming knowledge of the fully nonlinear power spectrum. On the other hand, one could *apriori* suspect that the short-scale kinetic energy is much harder to extract and requires simulations. However, by using energy conservation arguments we can expect to relate the kinetic and potential terms [371, 372]. This is achieved by the

Layzer-Irvine cosmic energy equation [42, 343]:

$$\frac{d}{d\eta}(\kappa + u) + \mathcal{H}(2\kappa + u) = 0\,, \tag{7.24}$$

which generalizes the concept of energy conservation to an expanding space that breaks time-translation symmetry, so that energy is not conserved in general. From Eq. (7.24) we see that energy conservation applies only for virialized scales for which $2\kappa + u = 0$.

Let us also emphasize that both κ and u appearing in Eq. (7.24) are short-scale quantities as defined in Eqs. (7.17) and (7.18), while the Layzer-Irvine equation is often phrased in terms of the total kinetic and potential energies with contributions from all scales. We present a derivation of Eq. (7.24) from the Boltzmann equation in Appendix 7.7, which holds true because contributions from long-wavelengths separately satisfy the very same equation, and can hence be subtracted.

We will be primarily interested in the integral form of Eq. (7.24) [9],

$$\kappa(\eta) = -a^{-2}(\eta) \int_0^\eta d\eta' a(\eta') \frac{d}{d\eta'} \left[a(\eta')u(\eta')\right] \,. \tag{7.25}$$

This formula shows that a nonzero average velocity dispersion is effectively sourced by a short-scale gravitational binding energy, and is hence a short-wavelength quantity (also see Appendix A in [43]).

Integrating Eq. (7.25) by parts leads to

$$\kappa(a) = -u(a) + a^{-2} \int_0^a da' a' u(a')\,, \tag{7.26}$$

where we now use the scale factor as our clock. This can be applied to compute the short-scale kinetic energy per unit mass from Eq. (7.22), and hence also the root mean square velocity dispersion $\sigma_{\rm dis} = \sqrt{2\kappa}$.

At this point it is instructive to see how Eq. (7.26) simplifies in the case where we use the linear theory power spectrum, $P(a,k) = D_{\rm L}^2(a) P_{\rm L}(k)$, to evaluate the gravitational binding

[9] We drop integration constants since they can be made negligible by choosing the initial time to be sufficiently early, and simply set $\eta_i = 0$. The resulting integral is dominated by its upper limit.

energy in Eq. (7.22). In this case, Eq. (7.22) becomes

$$u = -\frac{3}{4}\Omega_{\rm m,0}H_0^2\frac{D_{\rm L}^2}{a}\int\frac{dk}{2\pi^2}P_{\rm L}(k)\left[1-W_\Lambda(k)\right]^2\,. \tag{7.27}$$

Substituting Eq. (7.27) into Eq. (7.25) yields

$$\begin{aligned}\kappa &= \int\frac{dk}{2\pi^2}P_{\rm L}(k)\left[1-W_\Lambda(k)\right]^2\times\frac{3}{2}\Omega_{\rm m,0}H_0^2a^{-2}\int_0^\eta d\eta' a(\eta')D_{\rm L}(\eta')\frac{dD_{\rm L}}{d\eta'}\\ &= \frac{1}{2}a^2H^2f^2D_{\rm L}^2\int\frac{dk}{2\pi^2}P_{\rm L}(k)\left[1-W_\Lambda(k)\right]^2\,,\end{aligned} \tag{7.28}$$

where we introduced the linear growth rate $f(a) = d\log D_{\rm L}/d\log a$, and to arrive at the second line in Eq. (7.28) we used the differential equation satisfied by the linear growth factor

$$\frac{d^2D_{\rm L}}{d\eta^2} = \frac{3}{2}\Omega_{\rm m,0}H_0^2a(\eta)D_{\rm L}(\eta)\,, \tag{7.29}$$

to do the remaining time integral in the first line of Eq. (7.28). The final formula Eq. (7.28) corresponds to the linear theory velocity field integrated over short scales $k > \Lambda$, precisely as one would expect.

To verify the validity of the cosmic energy equation in the nonlinear regime we extract the root mean square velocity dispersion, smoothed on scales $R = \Lambda^{-1}$, from averaging the two phase-reserved runs of the dark-matter-only MilleniumTNG simulations (something similar was previously done in [373] as well). In Fig. 7.3 we plot the short-scale kinetic energy as a function of the smoothing scale $R = \Lambda^{-1}$, for varying redshift. Simulation results are shown as round data points [10]. Solid lines correspond to the theoretical expectation based on the Layzer-Irvine Eq. (7.26), computed using HMcode for the nonlinear power spectrum and also adding a hard cutoff at the fundamental mode of the simulation $k_{\rm F} = 2\pi/L_{\rm box}$. The

[10] We computed the N-body velocity dispersion by taking the particle-wise root-mean-square (RMS) of the simulation velocities. To subtract the long-range contribution, we first estimate the momentum field by distributing the particle velocities to a high-resolution 3D grid of 2048^3 voxels using the cubic spline assignment scheme. We also construct the density field the same way. We then divide the momentum field by the density field to obtain the velocities on the grid. After applying a Gaussian filter of radius R, we interpolate the smoothed velocities back to the N-body particles and subtract the smoothed velocities from the original particle velocities. Note, the particle-wise average is automatically density weighted. Finally, the short-scale kinetic energy follows from $\kappa = \sigma^2_{\rm dis}/2$.

dashed line is the theory prediction integrated over all wavenumber, and the dot-dashed line uses Halofit to compute the nonlinear power spectrum, and also adds a hard cutoff at the fundamental mode.

There is good agreement between theory and simulation results, confirming the validity of the cosmic energy Eq. (7.26). Once again Halofit perfoms slightly better than HMcode, but there are residual differences to the simulation results, which prefer slightly larger velocity dispersion, and increase on the smaller scales at later times. We were unable to pinpoint what is causing the discrepancies. Shot noise would increase the simulation measurements, but this leads to the question of why velocities are more sensitive to the shot noise than the densities, and we would expect it to be worse at early times. Another potential source of systematics are empty cells, which are problematic since we need to divide by density to get the velocities [11]. It is important to emphasize that cosmic variance cannot explain any discrepancies on the largest scales, since Fig. 7.2 shows excellent agreement even on such scales, and the actual realization of the kinetic energy (i.e. the round data points) can be reconstructed exactly from the actual realization of the gravitational energy.

7.4 The equation of state

In Sec. 7.3.1 we reviewed that short-wavelength fluctuations, which are not under perturbative control, contribute to the dynamics of long-wavelength fluctuations via an effective stress tensor. In this section we compute its ensemble average, which is fully specified by the pressure due to statistical isotropy. After dividing by the background density this produces the equation of state, $\omega(a)$, defined as follows

$$\omega(a) \equiv \frac{1}{3\bar{\rho}(a)} \langle \delta^{ij} \tau_{ij}^{\rm eff}(a, \vec{x}) \rangle = \frac{1}{3} \left[2\kappa(a) + u(a) \right] , \tag{7.30}$$

[11] This is dealt with by averaging an empty cell with all of its nearest neighbors. This smooths out the density and velocity fields only for these underdense cells. This could potentially slightly over-predict their velocities, and the effect would be more significant at later times and at smaller scales since a stronger clustering implies a larger abundance of empty cells.

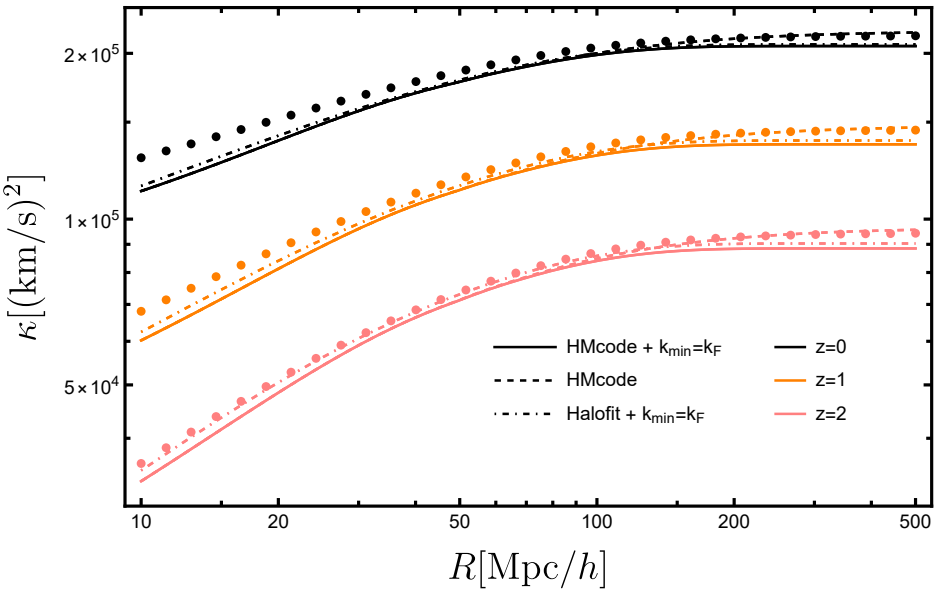

Figure 7.3: Short-scale kinetic energy per unit mass as a function of the smoothing scale $R = \Lambda^{-1}$, for varying redshift. Round data points correspond to results from the dark-matter-only runs of the MilleniumTNG simulations [369]. Solid, dashed and dot-dashed lines correspond to different theory predictions based on the Lyzer-Irvine Eq. (7.26), differing on the choice of code used to model the nonlinear power spectrum and whether or not a hard cutoff is introduced at the fundamental mode of the simulation (as indicated in the legends). There is good agreement between theory and simulation results, with Halofit performing slightly better than HMcode (see discussion on the main text for comments on the small discrepancies). This confirms the validity of the cosmic energy equation.

where we combined Eqs. (7.14),(7.17) and (7.18). Note that with this definition the cosmic energy Eq. (7.24) can be written in a more familiar form:

$$\bar{\rho}'_{\rm total} + 3\mathcal{H}\bar{\rho}_{\rm total}(1+\omega) = 0\,, \tag{7.31}$$

where $\bar{\rho}_{\rm total} = \bar{\rho}(1+\kappa+u)$ is the total background energy density, including contributions from the backreaction of short distance fluctuations, and we used $\bar{\rho} \propto a^{-3}$.

From Eq. (7.30) we also see that virialized scales, for which $2\kappa + u = 0$, exactly decouple

and do not contribute to the averaged effective stress [344]. This is much stronger than the statement that contributions from virialized structures are parametrically suppressed, which perhaps would be more natural from an EFT point of view. Instead, the decoupling of virialized scales should be thought of as a non-renormalization theorem [42].

With the ingredients developed in Sec. 7.3, particularly Eqs. (7.22) and (7.26), we can now easily compute the equation of state. This is plotted in Fig. 7.4, as a function of the cutoff scale Λ, in our fiducial $\Lambda\text{CDM}_{\text{DarkSky}}$ cosmology.

Note that the value of $\omega \approx 2 \times 10^{-7}$, at $R = \Lambda^{-1} = 10h^{-1}\text{Mpc}$ and $z = 0$, translates to a velocity of $\sqrt{2\omega} \approx 190\text{km/s}$, which is about a factor of two smaller than the value of $\sigma_{\text{dis}} \approx 450\text{km/s}$ observed in Fig. 7.3 at the same scale and redshift. This is due to (aside from the small difference in cosmologies) a partial cancellation between kinetic and potential contributions to the equation of state, which is exact for virialized scales.

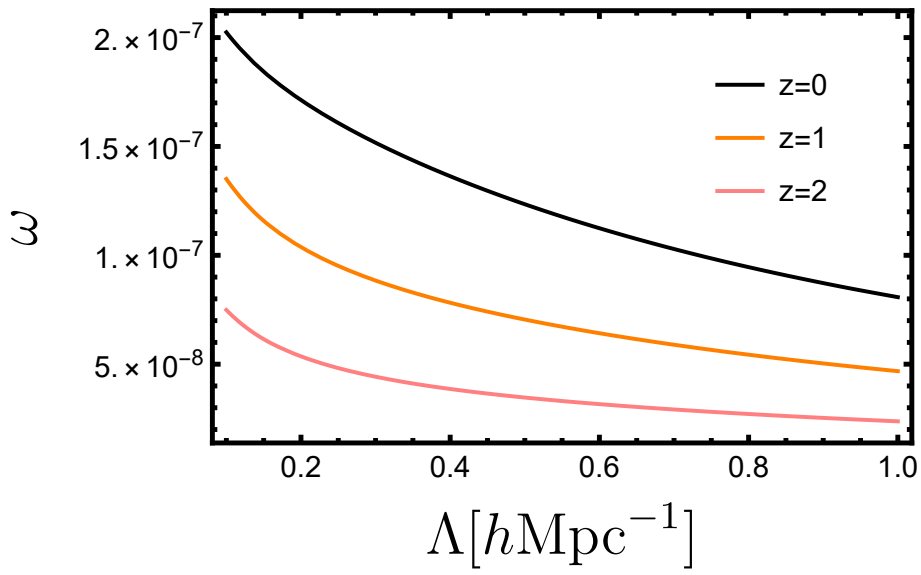

Figure 7.4: Equation of state as a function of the cutoff scale Λ in our fiducial cosmology.

Let us emphasize that our simple calculation for the equation of state makes no approximations, and only assumes accurate knowledge of the fully nonlinear matter power spectrum

as a function of redshift [12]. This is possible due to the fact that at this level all one needs to compute is the averaged effective stress tensor, which reduces to its trace due to statistical isotropy and its two contributions, i.e. kinetic and potential, are actually tied together by the cosmic energy Eq. (7.24). The situation is more complicated for the effective sound speed (the only 1-loop EFTofLSS free parameter), since in that case what matters is how the averaged stress responds to the presence of a long-wavelength mode. However, we will now see that it is still possible to meaningfully estimate the effective sound speed based on separate universe techniques.

7.5 Effective sound speed

In Sec.7.3 we wrote down Eqs. (7.14)-(7.16) for the effective stress tensor, $\tau_{ij}^{\rm eff}(a,\vec{x})$, in the EFTofLSS. In principle this fully specifies the effective sound speed in terms of short-wavelength fluctuations, but in practice one traditionally follows a bottom-up approach that consists in writing down an expansion including all operators made of long-wavelength fluctuations that are consistent with the symmetries [374–376], and with arbitrary time-dependent coefficients. This reads [13]

$$\frac{\tau_{ij}^{\rm eff}}{\bar{\rho}} = \omega\delta_{ij} + \left(c_s^2\delta_l - \frac{c_{bv}^2}{aH}\vec{\nabla}\cdot\vec{v}_l\right)\delta_{ij} - \frac{3}{4}\frac{c_{sv}^2}{aH}\left(\partial_j v_{li} + \partial_i v_{lj} - \frac{2}{3}\delta_{ij}\vec{\nabla}\cdot\vec{v}_l\right) + \Delta\tau_{ij} + \dots\,, \quad (7.32)$$

where ω is the equation of state calculated in Sec.7.4, $c_s^2(a)$ is the speed of sound, $c_{bv}^2(a)$ and $c_{sv}^2(a)$ are the coefficients of bulk and shear viscosity respectively, $\Delta\tau_{ij}$ represents stochastic terms and there are additional contributions of higher order in long-wavelength fields and derivatives thereof.

The relevant quantity, which appears in the Euler Eq. (7.12) (the divergence of that

[12] To compute the equation of state at $z = 0$ one only needs to know the matter power spectrum out to $z \lesssim 2$. This is because the time integral in Eq. (7.26) is dominated by its contributions from $z \ll 1$.

[13] This expansion is local in time, which is sufficient at 1-loop order. See [377] for the more general case of time non-locality, and references therein for recent discussions on the role of time non-locality in cosmological perturbation theory [378].

equation to more precise), is

$$\partial^i\left(\frac{1}{\rho_l}\partial^j\tau_{ij}^{\rm eff}\right)=\frac{1}{\bar\rho}\partial^i\partial^j\tau_{ij}^{\rm eff}+\cdots=c_s^2\nabla^2\delta_l-\frac{c_{bv}^2+c_{sv}^2}{aH}\nabla^2(\vec\nabla\cdot\vec v_l)=c_{\rm eff}^2\nabla^2\delta_l\,, \tag{7.33}$$

where

$$c_{\rm eff}^2(a)=c_s^2(a)+f(a)\left[c_{bv}^2(a)+c_{sv}^2(a)\right]\,, \tag{7.34}$$

is the effective sound speed. In the second equality in Eq. (7.33) we dropped higher order and stochastic terms, and used Eq. (7.32), and finally the third equality holds at one-loop level in perturbation theory since in that case we can apply the linear theory relation $\vec\nabla\cdot\vec v_l/aH=-f\delta_l$, with $f(a)=d\log D_{\rm L}/d\log a$ the linear growth rate.

The effective sound speed is then to be determined by matching the EFT perturbative power spectrum to the fully nonlinear power spectrum from simulations or observations. However, we will now show that the effective sound speed can be estimated directly from Eqs. (7.14)-(7.16) using tools developed to compute the equation of state in Sec.7.4, and in combination with separate universe methods. Let us also mention that the sound speed has been computed exactly from Eqs. (7.14)-(7.16) via cross correlations using N-body simulations through a much more complicated procedure (see Section 2.4 in [43]).

Now it no longer suffices to consider the average effective stress tensor since the quantity of interest, the left-hand side of Eq. (7.33), is proportional to derivatives of that tensor (i.e., these are tidal effects). To proceed, we exploit the separation of scales between long-wavelength and short-wavelength fluctuations to consider a local average of the effective stress tensor in the presence of a long-wavelength fluctuation $\delta_l(a,\vec x)$, $\langle\tau_{ij}^{\rm eff}\rangle_{\delta_l(a,\vec x)}$. This can be estimated using standard separate universe techniques that take the limit of an infinitely long-wavelength homogeneous fluctuation $\delta_l(a)$, which can be absorbed in the background expansion [379, 380]. This implicitly makes the assumption of a large separation of scales, with $k_s>\Lambda\gg k_l$. This is equivalent to dropping higher derivative corrections, which are not the subject of our investigation in any case.

The averaged stress tensor will only have a trace component in the separate universe due to statistical isotropy (i.e. δ_l is spherically symmetric). We will have more to say about this

at the end of this section, but for now we simply consider the trace's contribution to the effective stress

$$\langle \tau_{ij}^{\rm eff}\rangle_{\rm SU} = \frac{1}{3}\delta_{ij}\langle \tau^{\rm eff}\rangle_{SU}\,. \tag{7.35}$$

where we introduce the notation $\tau^{\rm eff} \equiv (\tau^{\rm eff})^k{}_k$, and quantities evaluated on the separate universe will carry a SU subscript, such as $\langle \tau^{\rm eff}\rangle_{SU}$. The first step to compute this is to expand in the long-wavelength fluctuation

$$\langle \tau^{\rm eff}\rangle_{\rm SU} = \langle \tau^{\rm eff}\rangle + \frac{\partial\langle \tau^{\rm eff}\rangle_{\rm SU}}{\partial\delta_l}\Big|_{\delta_l=0}\delta_l + \mathcal{O}(\delta_l^2)\,, \tag{7.36}$$

It will prove useful to define the equation of state in the presence of a homogeneous fluctuation, generalizing Eq. (7.30). That is,

$$\omega_{\rm SU} = \frac{\langle \tau^{\rm eff}\rangle_{\rm SU}}{3\bar\rho_{\rm SU}}\,. \tag{7.37}$$

The background evolution in the separate universe differs from the original cosmology. We review some results that will be needed shortly, and refer the reader to [379, 381] for additional details. The matter density in the separate universe is given by $\bar\rho_{\rm SU} = \bar\rho(1+\delta_l)$. From $\bar\rho_{\rm SU} \propto a_{\rm SU}^{-3}$ with $a_{\rm SU}$ the scale factor in the separate universe, and the Friedmann Eq. (7.20) evaluated at the present time, it follows that

$$\frac{\Omega_{\rm m,0}H_0^2}{a^3}(1+\delta_l) = \frac{\Omega_{\rm m,0}^{\rm SU}(H_0^{\rm SU})^2}{a_{\rm SU}^3}\,. \tag{7.38}$$

If at sufficiently early times we switch off the homogeneous fluctuation, $\delta_l \to 0$, while also demanding that $a_{\rm SU} \to a$ in the same limit, it then follows that $\Omega_{\rm m,0}H_0^2 = \Omega_{\rm m,0}^{\rm SU}(H_0^{\rm SU})^2$. This implies

$$a_{\rm SU} = a(1+\delta_l)^{-\frac{1}{3}} \implies \frac{da_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} = -\frac{1}{3}a\,. \tag{7.39}$$

These are all the results we need, and we are ready to substitute Eq. (7.37) into Eq. (7.36) to obtain

$$\langle \tau^{\rm eff}\rangle_{\rm SU} = 3\bar\rho\left[\omega + \left(\omega + \frac{d\omega_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0}\right)\delta_l + \mathcal{O}(\delta_l^2)\right]\,, \tag{7.40}$$

where ω is the equation of state in the original cosmology, and we used the relation $\bar{\rho}_{\rm SU} = \bar{\rho}(1+\delta_l)$. Combining Eqs. (7.33), (7.35) and (7.40) finally yields

$$c^2_{\rm eff} = \omega + \frac{d\omega_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} . \tag{7.41}$$

We can now use the results of Sec.7.4 for the equation of state to estimate the effective sound speed. From Eq. (7.30),

$$\frac{d\omega_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} = \frac{1}{3}\left(2\frac{d\kappa_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} + \frac{du_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0}\right) . \tag{7.42}$$

Taking a derivative of Eqs. (7.22) and (7.26) with respect to δ_l gives [14],

$$\frac{du_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} = -\frac{3}{4}\Omega_{\rm m,0}H_0^2\frac{1}{a}\int\frac{dk}{2\pi^2}P(a,k)\left[\frac{1}{3}+R(a,k)\right]\left[1-W_\Lambda(k)\right]^2 , \tag{7.43}$$

and

$$\frac{d\kappa_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} = -\frac{1}{3}u - \frac{du_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} + a^{-2}\int_0^a da' a'\left[\frac{du_{\rm SU}}{d\delta_l}\Big|_{\delta_l=0} + \frac{2}{3}u(a')\right] , \tag{7.44}$$

respectively, where

$$R(a,k) = \frac{d\log P_{\rm SU}(a,k)}{d\delta_l}\Big|_{\delta_l=0} , \tag{7.45}$$

is the response function. This quantity is directly related to the angle-averaged squeezed limit bispectrum [382–384], and was extracted from separate universe simulations in [380, 381, 385]. Here we use the semi-analytic halo model (HM) [39] to calculate the response, with explicit formulas provided in Appendix 7.9. Despite the limitations of the halo model, its predictions for the response function are in excellent agreement with the simulation results (see Fig.4 in [385] for a comparison at redshifts $z = 0$ and $z = 2$).

We now have all the ingredients needed to compute the estimated effective sound speed from separate universe techniques using Eq. (7.41), and we do so for two distinct cosmologies, matching the different choices made in a couple of previous independent measurements of the effective sound speed from N-body simulations. The first cosmology, here denoted by

[14] The combination $\Omega_{\rm m,0}H_0^2$ is independent of the long mode as we showed above, but the scale factor is not, and we need to use Eq. (7.39) when taking derivatives with respect to the long mode.

$\Lambda CDM_{\rm Consuelo}$, is defined by: $\Omega_{\rm m} = 0.25$, $\Omega_\Lambda = 0.75$, $h = 0.7$, $\sigma_8 = 0.8$ and $n_s = 1$, matching the choice made in [43] which uses the Consuelo simulations [386]. The second one is our fiducial cosmology, $\Lambda CDM_{\rm DarkSky}$: $\Omega_{\rm m} = 0.295$, $\Omega_\Lambda = 0.705$, $h = 0.688$, $\sigma_8 = 0.835$ and $n_s = 0.9676$, matching the choice made in [360] which uses the Dark Sky simulations [368].

In Fig. 7.5 we plot the effective sound speed as a function of the cutoff scale Λ, at redshift $z = 0$. Solid lines show the separate universe estimate and dashed lines are obtained from matching to simulations in the renormalized limit of $\Lambda \to \infty$, with a cutoff dependence fixed by the Renormalization Group (RG) flow (see Appendix 7.8 for details). The left and right panels compare our estimate to the simulation results of [43] and [360] respectively, [15] with details on these comparisons regarding different conventions in the literature explained in Appendix 7.8.

It is important to note that [43] was one of the very first efforts to measure the effective sound speed, using low resolution simulations and matching at a high renormalized scale of $k_{\rm ren} = 0.16h{\rm Mpc}^{-1}$. On the other hand, the follow-up work of [360] used higher resolution simulations and aimed for precision, managing to match at much lower scales of $k_{\rm ren} \lesssim 0.05h{\rm Mpc}^{-1}$. One of the conclusions reached in [361] was precisely that the effective sound speed can be overestimated by as much as a factor of two when matching at a high scale of $k_{\rm ren} \approx 0.2h{\rm Mpc}^{-1}$. This is apparent in the left plot of Fig. 7.5, where the estimated effective sound speed is about a factor of two lower than the low resolution simulation result, which is likely overestimated by the aforementioned reasons. On the other hand, as shown in the right panel in Fig. 7.5, there is a much better agreement between estimated and higher resolution simulation effective sound speeds.

When interpreting Fig. 7.5 it is important to keep in mind, as we will expand upon at the end of this section, that our framework is limited in that it does not account for the traceless part of the effective stress tensor. We expect its contribution to be comparable in

[15] Many other works have measured the sound speed counterterm in different scenarios, including a scaling universe [362], simplified initial conditions in Einstein-de Sitter [363], Lagrangian perturbation theory [364], in 1+1 dimensions [365] and with an IR resummed theoretical template [49, 356]. Also see [366, 367] and some works on the EFT at the field level, e.g. [332, 387–390].

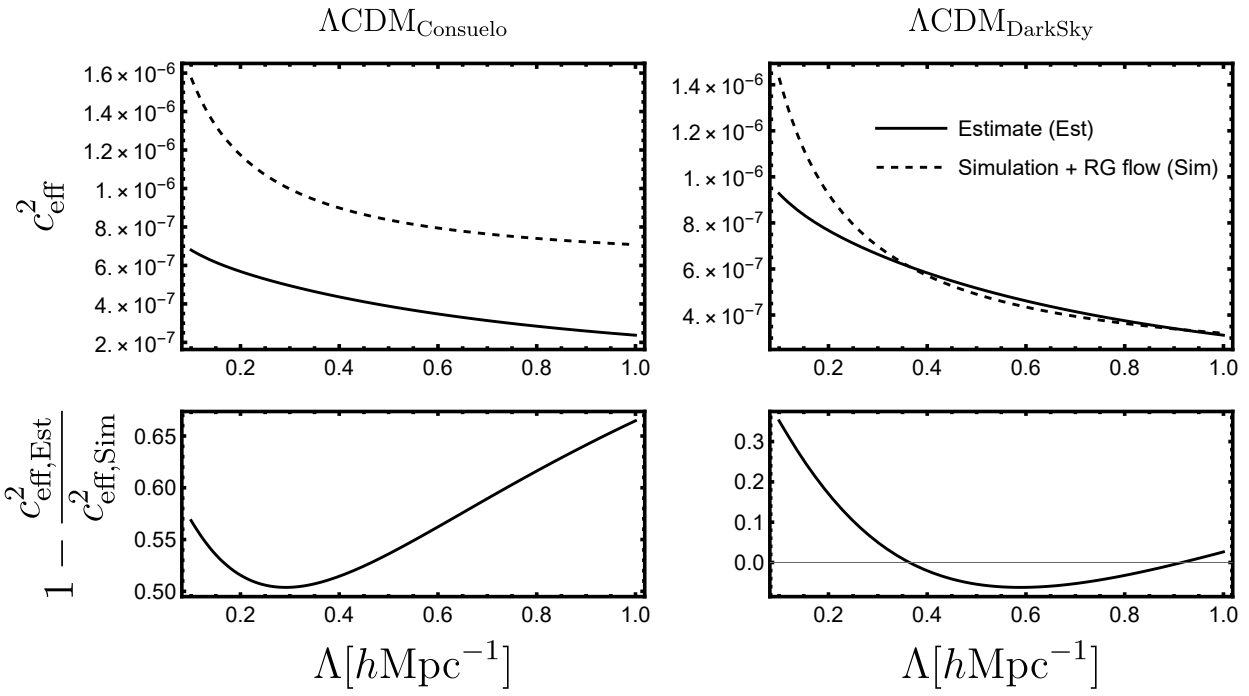

Figure 7.5: Effective sound speed as a function of the cutoff scale Λ, at redshift $z = 0$. Solid lines correspond to the estimate from separate universe techniques, and dashed lines are obtained from matching to simulations with a cutoff dependence fixed by the RG flow. Lower plots show the relative difference between estimated and simulation results. Left and right panels compare our estimate to the simulation results of [43] and [360]. There is better agreement (with differences at the few tens of percent level) between the estimated sound speed and the higher resolution simulation result of [360] (right panel), when compared to the lower resolution simulation result of [43] (left panel). Additional comments on the discrepancy between estimated and lower resolution simulation results, as seen on the left panel, can be found in the main text.

size to the one from the trace, and hence the estimates provided to be correct up to a factor of two. In particular, we do not expect the contribution from the trace to separately track

the RG flow.

In Fig. 7.6 we plot the estimated and simulation effective sound speeds as a function of the scale factor in our fiducial cosmology, for two different values of the cutoff, $\Lambda = (1/6)h\text{Mpc}^{-1}$ and $\Lambda = (1/3)h\text{Mpc}^{-1}$. Let us mention that the interpretation of simulation results (here from [360]) require an assumption about the time dependence of the renormalized effective sound speed counterterm, which we take to be that expected from a scaling universe [338] (and has been found to be consistent with the time dependence from N-body simulations [283, 367]). We elaborate on this in Appendix 7.8. Fig. 7.6 shows that the time dependence of the estimated sound speed is approximately consistent with the assumption of a scaling universe.

Let us stop for a moment to address a potential source of confusion related to the comparison of our semi-analytic estimate to simulation results. A naive direct application of our formalism in the limit $\Lambda \to \infty$ produces a vanishing effective sound speed, since in that limit there are no short modes contributing to the counterterm, while one does find a nonzero value for the renormalized counterterm from simulations in the same limit. The reason for this is that in our approach it does not make sense to directly push the cutoff far beyond the scale of nonlinearities, since then one would lose the perturbative control over the long-wavelength modes. On the other hand, theoretical control over all scales can be achieved with simulations, so a similar issue does not arise in this case [16].

More concretely, in the EFTofLSS the sound speed counterterm accomplishes two things at the same time. It accounts for the backreactions of short-distance fluctuations into the long modes, but it also absorbs the UV dependence of perturbation theory loop integrals. By working with a finite cutoff, which is below the scale of nonlinearities, one can ensure that the latter contributions to the counterterm are effectively vanishing since only perturbative modes are running in the loops. Then we only need to account for the former, physical

[16] In reality simulations have a finite resolution and make approximations such as force softening, so theoretical control can only be achieved up to some cutoff scale Λ_{sim}. This is of no practical importance as long as the observables of interest are well converged with respect to this cutoff.

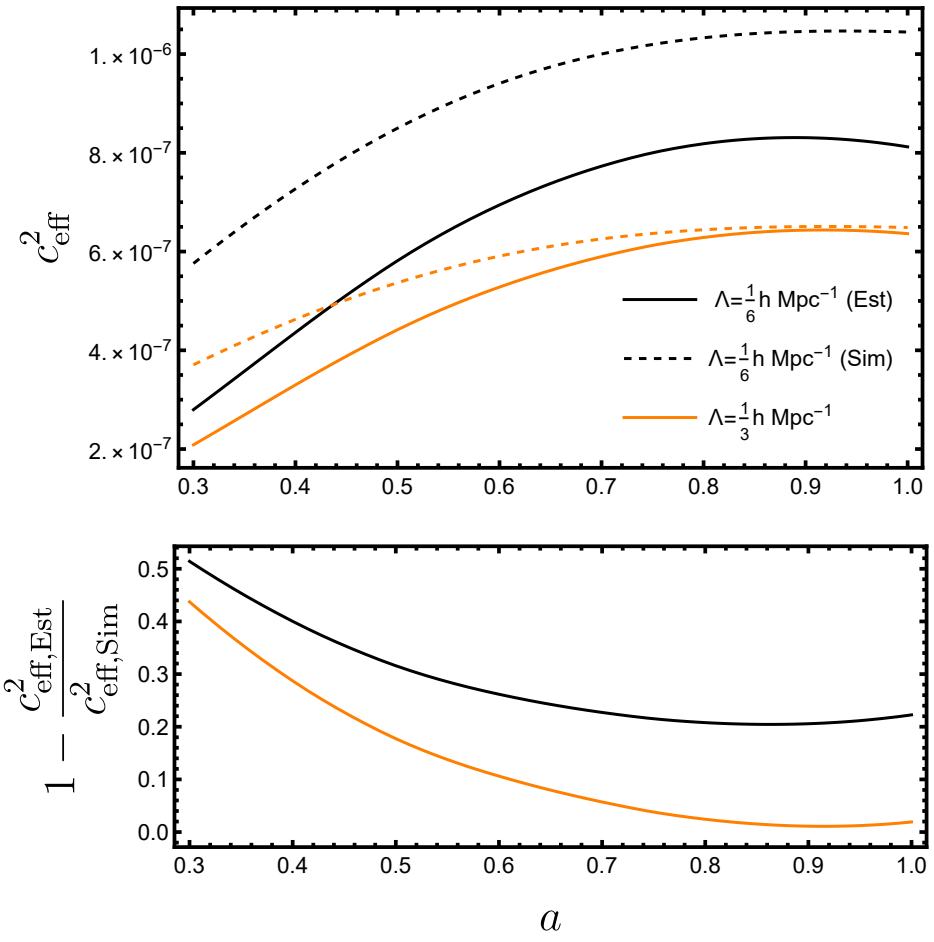

Figure 7.6: Effective sound speed as a function of the scale factor in our fiducial cosmology. Solid lines show the separate universe estimate while dashed lines correspond to the simulation results of [360] when augmented with the RG flow plus a scaling universe ansatz for the time dependence of the renormalized counterterm (see Appendix 7.8 for more details). Black and orange lines show results for $\Lambda = (1/6)h\text{Mpc}^{-1}$ and $\Lambda = (1/3)h\text{Mpc}^{-1}$, respectively. The time dependence of the estimated sound speed is approximately consistent with the assumption of a scaling universe, which was verified to accurately capture the time-dependence from full simulations.

contribution due to the coupling of short and long wavelength modes which is exactly the quantity we aim to estimate in this work. Afterwards, the unphysical contribution from the spurious UV behavior of loop integrals can be encapsulated by RG flowing towards $\Lambda = \infty$.

Note that due to the arguments made in the previous paragraph, in Fig. 7.5 we choose to compare the estimated and simulation sound speeds at a finite cutoff. An alternative approach would be to RG flow the estimated sound speed towards $\Lambda = \infty$, to compare with the renormalized counterterms directly extracted from simulations. However, since the estimates do not exactly follow the RG flow as shown in Fig. 7.5, this procedure would result in an estimated renormalized sound speed that depends on a spurious reference scale $\Lambda_{\rm ini}$ (the inital scale one is running from, towards $\Lambda \to \infty$). For example, in the Dark Sky cosmology at $z = 0$, the choices of $\Lambda_{\rm ini,1} = 0.3h{\rm Mpc}^{-1}$ and $\Lambda_{\rm ini,2} = 0.5h{\rm Mpc}^{-1}$ produce $c^2_{\rm eff,1}(\Lambda = \infty) \approx 1.4 \times 10^{-7}$ and $c^2_{\rm eff,2}(\Lambda = \infty) \approx 2 \times 10^{-7}$. We avoid this ambiguity by RG flowing the simulation results to finite Λ to compare with our estimates.

We find it remarkable that a semi-analytic calculation can reproduce the effective sound speed extracted from simulations so well. Let us emphasize that the short scale gravitational binding energy in the separate universe is the basic ingredient that goes into estimating the effective sound speed, or exactly computing the equation of state, and hence these calculations facilitate the interpretability of the EFT counterterm. For example, let us introduce the quantity $dc^2_{\rm eff}/d\log k$ according to

$$c^2_{\rm eff} = \int_{k>\Lambda} d\log k \frac{dc^2_{\rm eff}}{d\log k}\,, \tag{7.46}$$

which carries information about what scales contribute the most to the effective sound speed, and can be easily extracted by differentiating the latter with respect to the cutoff. This is plotted in Fig. 7.7. The features in the plot directly trace the shape of the power spectrum times the response function, i.e. the derivative of the power spectrum with respect to a long mode evaluated at $\delta_l = 0$. The curves for $dc^2_{\rm eff}/d\log k$ show a peak at typical scales of sheets and filaments, $k \sim (0.5 - 1)h{\rm Mpc}^{-1}$. This suggests that these are the cosmic structures which contribute the most to the effective sound speed. While we expect the contributions

from the traceless part of the stress tensor to make quantitative changes to the shape of the curves in Fig. 7.7 (for example, it may somewhat shift the peak locations), we expect the overall qualitative features to remain the same.

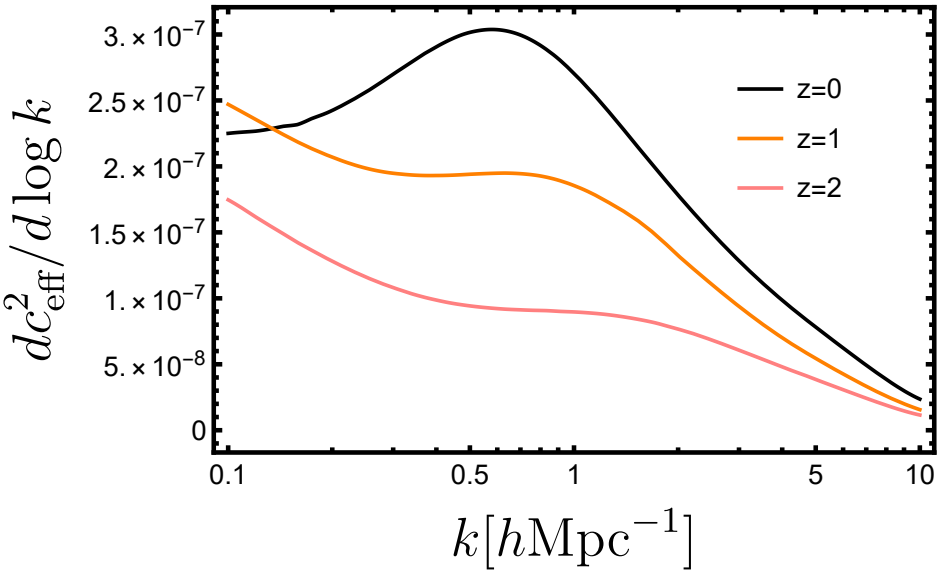

Figure 7.7: Derivative of the effective sound speed with respect to wavenumber, as defined in Eq. (7.46). The features in the plot directly trace the shape of the derivative of the power spectrum with respect to a long mode evaluated at $\delta_l = 0$, and show that scales as small as a few $\sim h\mathrm{Mpc}^{-1}$ have a sizable contribution to the effective sound speed. Additionally, the peaks at $k \sim (0.5-1)h\mathrm{Mpc}^{-1}$ suggest that cosmic sheets and filaments are the structures which contribute the most to the effective sound speed.

Additionally, the EFT counterterm carries information about the short scale gravitational dynamics, and is hence sensitive to the cosmology. This can be easily investigated within our semi-analytic framework. In Fig. 7.8 we show the response of the effective sound speed to varying σ_8, which corresponds roughly to a scaling of $c^2_{\rm eff} \sim \sigma_8^{2.2-2.5}$. Fig. 7.9 shows the response of the effective sound speed to varying $\Omega_{\rm m,0}$, corresponding roughly to a scaling of $c^2_{\rm eff} \sim \Omega_{\rm m,0}^{1.1-1.2}$.

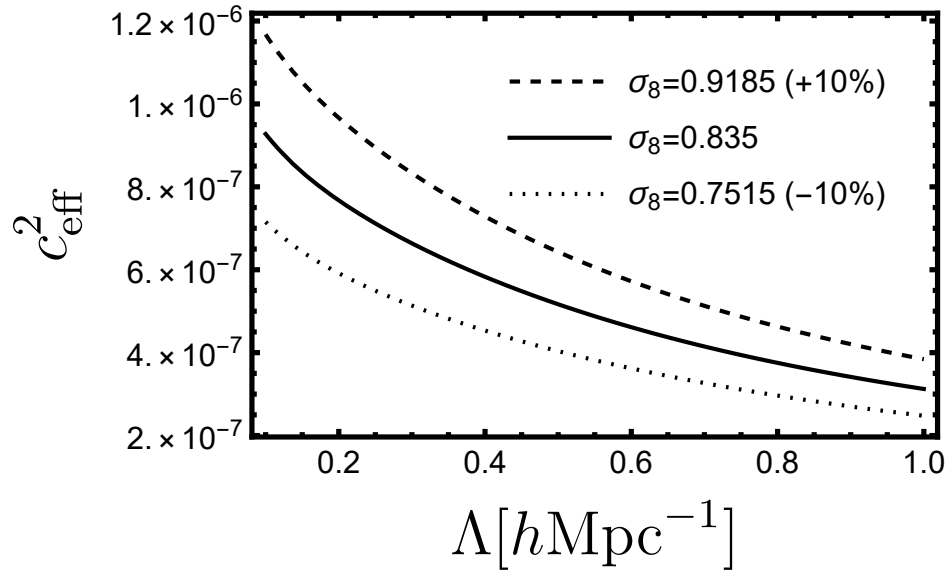

Figure 7.8: Effective sound speed as a function of the cutoff scale for varying σ_8. Here $z = 0$ and we consider 10% variations on σ_8 away from its value in our fiducial cosmology. This corresponds to an approximate scaling of $c^2_{\rm eff} \sim \sigma_8^{2.2-2.5}$.

A strong dependence on σ_8 was previously pointed out in [367] based on simulations, where they report a $c_{\rm ctr} \sim \sigma_8^{3.5}$ scaling for the counterterm (see footnote 13 in that reference). However, their comparison is not done while holding the other cosmological parameters fixed (for example, the cosmology with higher σ_8 also has a value of $\Omega_{\rm m,0}$ which is $\sim 5\%$ larger). Additionally, they extract the counterterm in the renormalized limit ($\Lambda \to \infty$) while we compute the sound speed at a finite cutoff, so a direct comparison is not appropriate.

We finish this section with a discussion on the main limitations of the separate universe approach pursued here. It cannot account for higher derivative corrections due to the underlying assumption of a large separation of scales, i.e., $k_s > \Lambda \gg k_l$. It does not account for stochastic terms either since it is only upon averaging that we can connect the kinetic and potential terms of the effective stress tensor, via the cosmic energy equation. These are not relevant limitations if one is only interested in the effective sound speed.

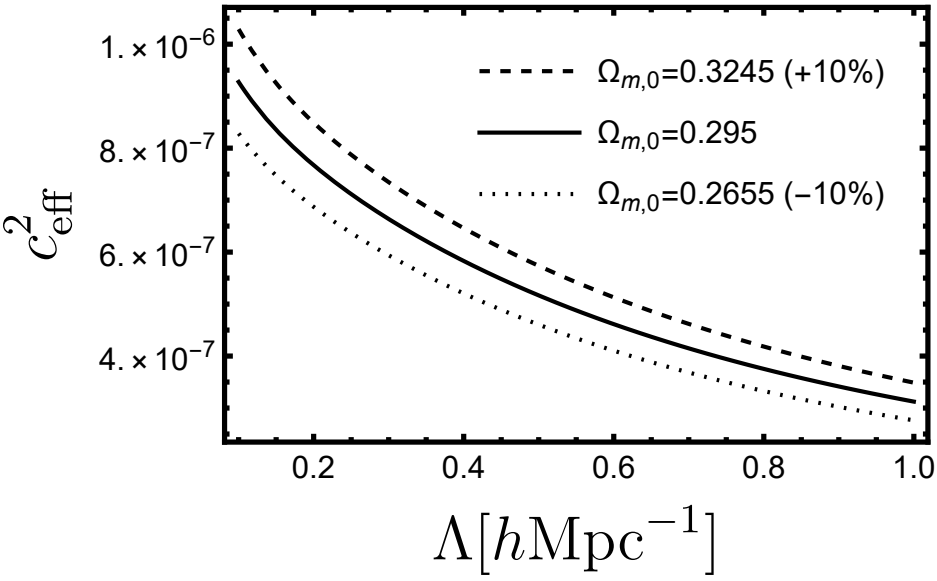

Figure 7.9: Effective sound speed as a function of the cutoff scale for varying $\Omega_{\rm m,0}$. Here $z = 0$ and we consider 10% variations on $\Omega_{\rm m,0}$ away from its value in our fiducial cosmology. This corresponds to an approximate scaling of $c^2_{\rm eff} \sim \Omega^{1.1-1.2}_{\rm m,0}$.

Most importantly, our approach misses contributions from the traceless part of the effective stress tensor, see Eq. (7.35). The local averaged traceless stress is nonzero precisely because at each point in space the preferred vector $\vec{\nabla}\delta_l$ locally breaks isotropy. In fact, previous measurements from simulations [43] and analytic estimates from perturbation theory considerations (see Appendix D in [42]) indicate the contribution from the traceless part of the effective stress tensor and from its trace to be of the same order of magnitude. We can then expect errors that may be as large as order unity, which is reflected in Figs. 7.5 and 7.6. This issue can potentially be addressed with anisotropic separate universe simulations, which account for the presence of a large-scale tidal field [391–394], and may allow for an exact calculation of the effective sound speed using separate universe techniques (as first suggested in [395]).

7.6 Conclusion

Effective field theory methods for large-scale structure have proven themselves to be extremely powerful in pushing the regime of applicability of cosmological perturbation theory to smaller scales, and are now a central piece of the theory modeling involved in the analysis of real data. Even in its simplest form [17], the EFTofLSS comes at the cost of adding a new nuisance parameter to the theory, the effective sound speed, which can be extracted from cosmological simulations by matching EFT predictions to the fully nonlinear power spectrum.

While very effective, this procedure does not provide a clear physical interpretation for the counterterm in that it does not shed light on the microphysics that is ultimately responsible for the effective sound speed. This is to be contrasted, for example, with Chiral perturbation theory in QCD [337] which has the pion decay constant as a free parameter (a quantity that has a physical interpretation). In this work we fill that gap by providing a semi-analytic estimate of the effective sound speed from a simplified model of the short scale fluctuations based on separate universe techniques. Our method reproduces the results from simulations with errors at the few tens of percent level at redshifts $z \lesssim 2$, and across a range of cutoffs $0.1h\text{Mpc}^{-1} \leq \Lambda \leq 1h\text{Mpc}^{-1}$. This is summarized in Figs. 7.5 and 7.6.

In Sec.7.3 we review the EFT approach to large scale structure and emphasize the power of the Layzer-Irvine Eq. (7.24), which enables a calculation of the average velocity dispersion, smoothed on a given scale, assuming knowledge of the fully nonlinear power spectrum, see Fig. 7.3. We build upon this to first compute the equation of state in Sec.7.4, as summarized in Fig. 7.4. Finally, in Sec.7.5 we use the separate universe approach to semi-analytically estimate the effective sound speed. The basic ingredient needed in that calculation is the short-scale gravitational binding energy in a separate universe, which provides a physical interpretation for the counterterm. The effective sound speed carries information about

[17] In the context of modeling the power spectrum of the matter field, in real space, at 1-loop in perturbation theory. This set-up is assumed in the discussion that follows.

the nonlinear gravitational dynamics of short-scale fluctuations, revealed for example by its dependence on cosmological parameters, see Figs. 7.8 and 7.9. This is underscored in our approach which is ultimately based on integrals of the fully nonlinear power spectrum with contributions from very small scales, as illustrated in Fig. 7.1 and 7.7.

In a broader note, methods that shed light on and/or constrain nuisance parameters in our EFT based perturbartive models of large scale structure can be helpful in providing well informed theoretical priors for such parameters. This improves the extraction of cosmological parameters [396–398], and helps to alleviate prior volume effects in Bayesian analyses [53, 346–352].

We have some ideas for future directions of investigation. It would be interesting to further elucidate the physical interpretation of the equation of state and effective sound speed. A field level analysis could help to provide a visual picture of what type of nonlinear structures contribute the most to these two quantities, such as sheets or filaments [18], and directly reveal the decoupling of fully virialized structures in an instructive way. Additionally, it would be worthwhile to test the limits of our approach in a thorough comparison against different methods to extract the sound speed from simulations in different cosmologies, potentially even beyond ΛCDM. Such investigations hold the potential to reveal interesting aspects of the short scale gravitational dynamics underlying the EFTofLSS.

7.7 Appendix: A derivation of the Layzer-Irvine equation

Our goal is to derive the Layzer-Irvine Eq. (7.24) directly from the Boltzmann Eq. (7.8), which we repeat here for convenience

$$\frac{\partial f_l}{\partial \eta} + \vec{q}\cdot\frac{\partial f_l}{\partial \vec{x}} = a^2(\eta)\frac{\partial \phi_l}{\partial \vec{x}}\cdot\frac{\partial f_l}{\partial \vec{q}} + a^2(\eta)\left[\frac{\partial \phi_s}{\partial \vec{x}}\cdot\frac{\partial f_s}{\partial \vec{q}}\right]_\Lambda + O\Big(\frac{k^2}{\Lambda^2}\Big)\,. \tag{7.47}$$

We first introduce the background long-wavelength distribution function, defined by $\bar{f}_l(\eta, q) = \langle f_l(\eta, \vec{x}, \vec{q})\rangle$, and its fluctuations $\delta f_l = f_l - \bar{f}_l$. Now take the ensemble average of Eq. (7.47)

[18] This type of analysis may however be hindered by the fact that differentiating these kinds of structures is still an open area of research (see, e.g., [399–405] for a few recent papers on the subject).

to obtain

$$\frac{\partial \bar{f}_l}{\partial \eta} = a^2 \left\langle \frac{\partial \phi_l}{\partial \vec{x}} \cdot \frac{\partial \delta f_l}{\partial \vec{q}} \right\rangle + a^2 \left\langle \frac{\partial \phi_s}{\partial \vec{x}} \cdot \frac{\partial f_s}{\partial \vec{q}} \right\rangle + \dots \,, \tag{7.48}$$

where we omit cutoff and time dependences for simplicity of notation, and the ellipsis denote higher derivative corrections. Each term on the right-hand side of Eq. (7.48) act as a source term, whose contribution to $\bar{f}_l$ can be found by integrating it over superconformal time. This leads to a solution of the form

$$\bar{f}_l(\eta, q) = \bar{f}_{l,\mathrm{pt}}(\eta, q) + \bar{f}_{l,\mathrm{ctr}}(\eta, q) + \dots \,, \tag{7.49}$$

where $\bar{f}_{l,\mathrm{pt}}(\eta, q)$ and $\bar{f}_{l,\mathrm{ctr}}(\eta, q)$ are sourced by the long and short wavelength fluctuations respectively, as given by the first and second terms on the right-hand side of Eq. (7.48). The mixed long-short terms, i.e. higher derivative corrections, vanish in the renormalized limit of $\Lambda \to \infty$ but are otherwise necessary to accurately model observables of interest [42, 43]. We neglect higher derivative corrections, but note that averages of such terms are always small due to the separation of scales implied by the EFT [19].

The first equation we need to solve is then,

$$\frac{\partial \bar{f}_{l,\mathrm{pt}}}{\partial \eta} = a^2 \left\langle \frac{\partial \phi_l}{\partial \vec{x}} \cdot \frac{\partial \delta f_l}{\partial \vec{q}} \right\rangle . \tag{7.50}$$

Since this only involves long-wavelength fluctuations, Eq. (7.50) can be solved perturbatively when coupled to the Boltzmann equation for the distribution function fluctuations as well. A subset of us pursue this in detail elsewhere [82], and we find that Standard Perturbation Theory (SPT) can be formulated directly at the distribution function level in phase-space, circumventing the need to truncate the Boltzmann hierarchy. Here we only need the following result

$$a^{-5} \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}_{l,\mathrm{pt}} = \langle \rho_l v_l^2 \rangle \,, \tag{7.51}$$

which can be understood as follows: SPT does not account for velocity dispersion, hence Eq. (7.51) must hold to ensure that the ensemble average of Eq. (7.15) vanishes when there are no contributions from short-wavelength fluctuations.

[19] For example, $\langle \phi_l \phi_s \rangle$ is given by an integral over wavenumber which contains a factor of $W_\Lambda(k)[1 - W_\Lambda(k)]$ in the integrand. This is zero when both $k \ll \Lambda$ and $k \gg \Lambda$.

Moving on to the counterterm contribution to the long-wavelength background distribution function, we need to solve the following equation,

$$\frac{\partial \bar{f}_{l,\rm ctr}}{\partial \eta} = a^2 \left\langle \frac{\partial \phi_s}{\partial \vec{x}} \cdot \frac{\partial f_s}{\partial \vec{q}} \right\rangle . \tag{7.52}$$

Now multiply Eq. (7.52) by q^2, followed by integrating over momentum $\vec{q}$. The left-hand side of Eq. (7.52) becomes

$$\int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \frac{\partial \bar{f}_{l,\rm ctr}}{\partial \eta} = \frac{d}{d\eta} \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \bar{f}_{l,\rm ctr} = 2\frac{d}{d\eta}\left(a^5 \bar{\rho}\kappa\right) , \tag{7.53}$$

where the third equality follows from taking the ensemble average of Eq. (7.15), using Eqs. (7.17), (7.49) and (7.51). The right-hand side of Eq. (7.52) leads to

$$\int \frac{d^3\vec{q}}{(2\pi)^3} q^2 a^2 \left\langle \frac{\partial \phi_s}{\partial \vec{x}} \cdot \frac{\partial f_s}{\partial \vec{q}} \right\rangle = a^2 \left\langle \frac{\partial \phi_s}{\partial \vec{x}} \cdot \int \frac{d^3\vec{q}}{(2\pi)^3} q^2 \frac{\partial f_s}{\partial \vec{q}} \right\rangle = -2a^6 \left\langle \frac{\partial \phi_s}{\partial \vec{x}} \cdot \vec{\Pi}_s \right\rangle , \tag{7.54}$$

where after integrating by parts we used the short-wavelength analogue of Eq. (7.11). One can now use the short-wavelength analogs of Eqs. (7.4) and (7.10) [20] to derive the following result, through a series of integration by parts under the average sign (see [7])

$$-2a^6 \left\langle \frac{\partial \phi_s}{\partial \vec{x}} \cdot \vec{\Pi}_s \right\rangle = -2a \frac{d}{d\eta}\left(a^4 \bar{\rho}\omega\right) , \tag{7.55}$$

where the short scale gravitational binding energy, $v(\eta)$, is defined in Eq. (7.19). It is now straightforward to show that combining Eqs. (7.53) and (7.55) leads to the desired result:

$$\frac{d}{d\eta}(\kappa + v) + \mathcal{H}(2\kappa + v) = 0 , \tag{7.56}$$

using $\bar{\rho} \propto a^{-3}$ and $\mathcal{H} = d\log a/d\eta$.

7.8 Appendix: The renormalization group flow

In Figs. 7.5 and 7.6 we compare the semi-analytic estimate for the effective sound speed, based on separate universe methods, to the results obtained from matching the EFT to full

[20] All equations which are linear in the distribution function are satisfied separately for both long and short wavelength parts.

N-body simulations. In the latter case we read off the renormalized effective sound speed, corresponding to the $\Lambda \to \infty$ limit, directly from the literature [43, 360, 361] and the cutoff dependence is fixed by the RG flow which we now review.

Our first step will be to use the fluid equations in the EFTofLSS to make the connection between the effective sound speed and the EFT counterterm [21]. In terms of $\rho_l = \bar{\rho}(1+\delta_l)$, and $\vec{\Pi}_l = \rho_l \vec{v}_l = \bar{\rho}(1+\delta_l)\vec{v}_l$, the continuity Eq. (7.10) reads

$$\delta_l' + a\theta_l = -a\vec{\nabla}\cdot(\delta_l \vec{v}_l)\,, \tag{7.57}$$

where we used the relation $\bar{\rho} \propto a^{-3}$, and defined $\theta_l = \vec{\nabla}\cdot\vec{v}_l$. Time, scale and cutoff dependencies are implicit for simplicity of notation. Next we want to take the divergence of the Euler Eq. (7.12), which yields

$$\theta_l' + \mathcal{H}\theta_l + \frac{3}{2}\Omega_{\rm m,0}H_0^2\delta_l = -a\partial^i\left(v_{l,j}\partial^j v_{l,i}\right) - a\partial^i\left(\frac{1}{\rho_l}\partial^j\tau_{ij}^{\rm eff}\right)\,, \tag{7.58}$$

where we used the Poisson Eq. (7.21). Here we will focus on the contribution to the density contrast coming from the effective stress. Let us first repeat here Eq. (7.33) defining the effective sound speed:

$$\partial^i\left(\frac{1}{\rho_l}\partial^j\tau_{ij}^{\rm eff}\right) = c_{\rm eff}^2\nabla^2\delta_l\,. \tag{7.59}$$

The coefficient $c_{\rm eff}^2(a)$ is generated by short-wavelength quadratic nonlinearities and hence should be thought of as a second order quantity in perturbation theory to leading order. At one-loop precision we can then substitute $\delta_l \to \delta_l^{(1)}$ in Eq. (7.59). Keeping that in mind, and dropping all other nonlinear source terms in Eqs. (7.57) and (7.58), we arrive at the following system of equations

$$\begin{aligned}
&\frac{d}{d\eta}\delta_{l,\rm ctr} + a\theta_{l,\rm ctr} = 0\\
&\frac{d}{d\eta}\theta_{l,\rm ctr} + \mathcal{H}\theta_{l,\rm ctr} + \frac{3}{2}\Omega_{\rm m,0}H_0^2\delta_{l,\rm ctr} = ac_{\rm eff}^2k^2\delta_l^{(1)}\,.
\end{aligned} \tag{7.60}$$

[21] It is often the case in the literature that the EFT counterterm directly defines the effective sound speed. We are following the conventions of [43] where these two quantities are interconnected, but are not the same. This will be made more clear in what follows.

We can now eliminate the divergence of the velocity field to arrive at

$$\frac{d^2}{d\eta^2}\delta_{l,\rm ctr} - \frac{3}{2}\Omega_{\rm m,0}H_0^2 a(\eta)\delta_{l,\rm ctr} = 2a^2(\eta)D_{\rm L}(\eta)c^2_{\rm eff}(\eta)h_{\rm ctr}(\vec{k})\,, \tag{7.61}$$

where $\delta_l^{(1)}(\eta,\vec{k}) = D_{\rm L}(\eta)W_\Lambda(k)\delta_{\rm L}(\vec{k})$, and we define

$$h_{\rm ctr}(\vec{k}) = -\frac{1}{2}k^2 W_\Lambda(k)\delta_{\rm L}(\vec{k})\,. \tag{7.62}$$

In [82] we show that Eq. (7.61) admits an analytic solution of the form $\delta_{l,\rm ctr}(a,\vec{k}) = c_{\rm ctr}(a)h_{\rm ctr}(\vec{k})$, where:

$$c_{\rm ctr}(a;\Lambda) = 2H(a)\int_0^a \frac{da'}{(a')^3H^3(a')}\int_0^{a'} da''\frac{D_{\rm L}(a'')}{a''}c^2_{\rm eff}(a';\Lambda)\,. \tag{7.63}$$

Here we switch from superconformal time to the scale factor and restore the cutoff dependence. The counterterm contribution to the one-loop power spectrum is then

$$P_{\text{1-loop}}(a,k;\Lambda) = P_{\text{1-loop,SPT}}(a,k;\Lambda) - D_{\rm L}(a)c_{\rm ctr}(a;\Lambda)k^2W^2_\Lambda(k)P_{\rm L}(k)\,. \tag{7.64}$$

The 1-loop power spectrum in SPT has the following structure

$$P_{\text{1-loop,SPT}}(a,k;\Lambda) = D^2_{\rm L}(a)W^2_\Lambda(k)P_{\rm L}(k) + P_{22}(a,k;\Lambda) + P_{13}(a,k;\Lambda)\,, \tag{7.65}$$

where

$$P_{13}(a,k\to 0;\Lambda) \approx -\frac{61}{315}D^4_{\rm L}(a)k^2W^2_\Lambda(k)P_{\rm L}(k)\int_0^\infty \frac{dq}{2\pi^2}W^2_\Lambda(q)P_{\rm L}(q)\,, \tag{7.66}$$

in the EdS approximation to perturbation theory kernels [41]. The soft limit of $P_{22}(a,k;\Lambda)$ is subdominant, with a wavenumber scaling $\sim\mathcal{O}(k^4)$ as $k\to 0$.

Self-consistency of the EFT requires that the soft limit of the total 1-loop power spectrum in Eq. (7.64) has to be cutoff independent, up to an overall multiplicative factor of $W^2_\Lambda(k)$. This ensures that UV contributions to perturbation theory loop integrals can be absorbed by the EFT counterterm, and implies [22]

$$c^{(3)}_{\rm ctr}(a;\Lambda) = c^{(3)}_{\rm ctr}(a;\Lambda=\infty) + \frac{61}{315}D^3_{\rm L}(a)\int_0^\infty \frac{dq}{2\pi^2}P_{\rm L}(q)\left[1-W^2_\Lambda(q)\right]\,. \tag{7.67}$$

[22] The subdominant UV contributions from $P_{22}(a,k;\Lambda)$ are absorbed by stochastic terms.

This is the Renormalization Group (RG) flow, which entirely fixes the counterterm as a function of the cutoff as soon as the renormalized coefficient, $c^{(3)}_{\rm ctr}(a;\Lambda=\infty)$, is specified (see [406–408] for recent papers on the RG flow in the context of biased tracers).

For us this is not the end of the story, as we are interested in working at the level of the effective sound speed, which is related to the counterterm via Eq. (7.63). This equation can be inverted:

$$\begin{aligned} c^2_{\rm eff}(a';\Lambda) &= \frac{a}{D_{\rm L}(a)}\frac{d}{da}\left\{a^3H^3(a)\frac{d}{da}\left[\frac{c_{\rm ctr}(a;\Lambda)}{2H(a)}\right]\right\} \\ =&\frac{a^4H^2}{2D_{\rm L}}\left\{\frac{d^2c_{\rm ctr}}{da^2}+\frac{1}{a}\left(3+\frac{d\log H}{d\log a}\right)\frac{dc_{\rm ctr}}{da}-\frac{1}{a^2}\left[3\frac{d\log H}{d\log a}+\left(\frac{d\log H}{d\log a}\right)^2+\frac{a^2}{H}\frac{d^2H}{da^2}\right]c_{\rm ctr}\right\}\,, \end{aligned} \tag{7.68}$$

where in the second line we omit time and cutoff dependencies. The substitution of Eq. (7.67) into Eq. (7.68) yields

$$\begin{aligned} c^2_{\rm eff}(\Lambda) = c^2_{\rm eff}(\Lambda=\infty)+\frac{61}{630}a^2D_{\rm L}^2H^2\Bigg\{&3f\left(3f-1+\frac{d\log f}{d\log a}\right)+3f\left(3+\frac{d\log H}{d\log a}\right) \\ &-\left[3\frac{d\log H}{d\log a}+\left(\frac{d\log H}{d\log a}\right)^2+\frac{a^2}{H}\frac{d^2H}{da^2}\right]\Bigg\}\int_0^\infty\frac{dq}{2\pi^2}P_{\rm L}(q)\left[1-W^2_\Lambda(k)\right]\,. \end{aligned} \tag{7.69}$$

This is the RG flow for the effective sound speed, where $f(a)=d\log D_{\rm L}/d\log a$ is the linear growth rate. Note that Eq. (7.69) was derived while remaining completely agnostic about the time dependence of the renormalized effective sound speed. This improves upon the considerations made in [43] (while remaining consistent with it), which assume a perturbative time dependence for this quantity in order to derive its RG flow.

In the main text we directly compare our semi-analytic estimate for the effective sound speed, based on separate universe methods, with two previous measurements of the same quantity based on matching the EFT to full N-body simulations. In [43] the authors directly quote the value for the renormalized effective sound speed $c^2_{\rm eff}(\Lambda=\infty)\approx 0.6\times10^{-6}$ at $z=0$, so that Eq. (7.69) can be applied as is to compute the effective sound speed as a function of the cutoff scale.

On the other hand, in [360] the authors only extract numerical values for the renormalized counterterm $c_{\rm ctr}(\Lambda=\infty)$. It is important to keep in mind that they assume different conventions than the ones adopted here. In [360], Eq. (7.64) should be replaced by

$$P_{\text{1-loop}}(a,k;\Lambda)=P_{\text{1-loop,SPT}}(a,k;\Lambda)-2(2\pi)c_s^2(a;\Lambda)D_{\rm L}^2(a)\frac{k^2}{k_{\rm NL}^2}W_\Lambda^2(k)P_{\rm L}(k)\,, \tag{7.70}$$

with for instance $c_s^2(\Lambda\to\infty)\approx 0.53\ (k_{\rm NL}/(2h{\rm Mpc}^{-1}))^2$ at $z=0$ extracted from N-body simulations. A direct comparison of Eq. (7.70) with Eq. (7.64) allows us to obtain numerical values for $c_{\rm ctr}(\Lambda=\infty)$. Note that [360] denote by the sound speed what we have defined as the EFT counterterm, up to overall multiplicative factors. We then need to apply Eq. (7.63) to obtain numerical values for $c_{\rm eff}^2(\Lambda=\infty)$, and this entails making assumptions about the time dependence of the renormalized effective sound speed.

We assume the time dependence expected of a scaling universe [338], $P(k)\propto k^n$, with $n=-1.5$ following [356]. This is a good approximation in our fiducial cosmology for the range of scales relevant for the EFT [366, 409]. This leads to

$$c_{\rm ctr}(a,\Lambda=\infty)=c_{\rm ctr}(a=1,\Lambda=\infty)\left[D_{\rm L}(a)\right]^\alpha\,, \tag{7.71}$$

with $\alpha=4/(n+3)$. We checked explicitly that Eq. (7.71) is in good agreement with the results of Table 2 in [360] from matching the EFT to N-body simulations at redshifts $z=0$, $z=1$ and $z=2$. The substitution of Eq. (7.71) into Eq. (7.68) yields

$$\begin{aligned}c_{\rm eff}^2(a,\Lambda=\infty)&=c_{\rm ctr}(a=1,\Lambda=\infty)\frac{a^4H^2}{2D_{\rm L}}\frac{D_{\rm L}^\alpha}{a^2}\times\\&\times\left\{\alpha f\left(\alpha f-1+\frac{d\log f}{d\log a}\right)+\alpha f\left(3+\frac{d\log H}{d\log a}\right)-\left[3\frac{d\log H}{d\log a}+\left(\frac{d\log H}{d\log a}\right)^2+\frac{a^2}{H}\frac{d^2H}{da^2}\right]\right\}\,.\end{aligned} \tag{7.72}$$

This can now be combined with Eq. (7.69) to generate the effective sound speed as a function of the cutoff scale from reference [360], at any given redshift.

7.9 Appendix: Halo model response

The expression for the response function in the halo model is

$$R^{\rm HM}(a,k) = \frac{\left(1+\frac{26}{21}-\frac{1}{3}\frac{d\log P_{\rm L}}{d\log k}\right)\left[I_1^1(a,k)\right]^2 D_{\rm L}^2(a)P_{\rm L}(k) + I_2^1(a,k)}{\left[I_1^1(a,k)\right]^2 D_{\rm L}^2(a)P_{\rm L}(k) + I_2^0(a,k)}\,, \tag{7.73}$$

where

$$I_m^k(a,k) = \int dM \frac{dn}{dM}\Big|_{a,M} \left(\frac{M}{\bar{\rho}_0}\right)^m [b(a,M)]^k [u(M|k)]^m\,. \tag{7.74}$$

Here $dn/dM|_{a,M}$ is the halo mass function, $b(a,M)$ is the halo bias and $u(M|k)$ is the normalized halo density profile in Fourier space. We assume universality and adopt the halo mass function from Appendix C in [410], halo bias from [210] and a Navarro-Frenk-White (NFW) halo profile [206].

Note that the first line of Eq. (2.33) in [385] contains a typo, since it misses the denominator in Eq. (7.73) that is just the halo model expression for the power spectrum.

Chapter 8

CONCLUSION

The era of precision cosmology is upon us, as evidenced by the wealth of high quality data already available from previous and ongoing astronomical surveys [411–419]. The future of the field is also very bright, with a number of missions planned to launch in the upcoming decade [60, 420–423]. A key component of this program consists in mapping out the distribution of structures on large scales, via various tracers of the underlying dark matter field such as galaxies, weak gravitational lensing and line intensity mapping.

The ultimate goal is to use these unprecedented amounts of data to learn about fundamental physics. For instance, this may result in the discovery of the nature of dark matter and dark energy, or new ingredients at play in the very early Universe. Another very well motivated science case of cosmological surveys is to measure the sum of neutrino masses, which leaves distinctive signatures on the Universe's expansion history and its cosmological perturbations.

The success of this enterprise relies on the development of robust and accurate theoretical models for the cosmological observables, to be paired with efficient computational techniques that enable extensive data analysis. This thesis detailed many of the contributions I made to different aspects of this program as a grad student, which we now summarize.

8.1 Summary of the thesis

- Chapter 2 addressed the question of to what extent Newtonian N-body simulations of nonlinear structure formation can be applied to model the dynamics of perturbations in massive neutrino species. Massive neutrinos have significant thermal velocities at sufficiently early times, where special relativistic corrections to the equations of motion

need to be incorporated. We found that the total distance neutrino particles travel in the simulation box, the neutrino horizon, is overestimated by the Newtonian evolution (when compared to the more accurate relativistic dynamics). This produces biases on cosmological observables at large linear scales, that can be mitigated by a proper choice of initial conditions for the simulations.

- Chapter 3 introduced a novel approach to solve the exact dynamics of massive neutrino perturbations in the linear regime, the Generalized Boltzamnn Hierarchy (GBH). It effectively integrates out the momentum dependence of the system of equations, which is now defined in configuration space instead of in full phase space. We argued that the GBH is much simpler and easier to solve numerically than the standard Boltzmann hierarchy.

- Chapter 4 significantly improved the fluid approximation for massive neutrinos, by properly accounting for the scale dependence of the sound speed, which can be derived analytically in the limit of both large and small scales, when compared to the neutrino free-streaming scale. We showed that the new fluid approximation models neutrino density and velocity transfer functions with errors $< 5\%$ at resdshifts $z \lesssim 5$, for all relevant scales, $k \sim (10^{-3} - 10)h\mathrm{Mpc}^{-1}$.

- Chapter 5 develops the theoretical model for neutrino wakes, a new signature of neutrino masses in the large scale structure, that emerges due to the preferential accumulation of neutrino particles downstream of moving cold dark matter structures. We provide a first principles derivation of this effect starting from the Vlasov equation for massive neutrinos around a known cold dark matter profile, which allows a more complete characterization of its imprints on cosmological observables. We show that this signal can potentially be detected in future survey via three-point cross correlations of matter and galaxies, and it may open a window to distinguish the mass splittings (individual neutrino masses).

- Chapter 6 advanced a perturbative approach to nonlinear structure formation at late times, which directly solves the coupled system of Vlasov-Poisson equations in full phase space. In this framework one does not need to assume an ideal fluid from the outset, which is the starting point of the Standard Perturbation Theory (SPT) method. We rederived the SPT kernels from this new framework, and also applied it to shed new light on the emergence of EFT-like counterterms.

- Chapter 7 demonstrates that the sound speed counterterm, introduced in the Effective Field Theory of Large Scale Structure (EFTofLSS), can be estimated analytically via a combination of the separate universe technique and the halo model. We showed that this estimate is in very good agreement with the values extracted by matching the theory to full numerical simulations, and it carries valuable information about the microphysics underlying the sound speed, and its dependence on cosmological parameters.

8.2 Future directions

There are many natural extensions to the works summarized in Sec. 8.1, some of which I am currently pursuing with collaborators, and some others I plan to investigate at a later time. I now list a few of these directions.

- We are currently working on a reformulation of how massive neutrinos are implemented in the Boltzmann solver CLASS, based on a combination of the Generalized Boltzmann Hierarchy (described in Chapter 3) and the improved fluid approximation (described in Chapter 4). Preliminary results show improvement in both accuracy and speed when compared to existing methods.

- We are in final stages of preparing a publication that further develops the theoretical modeling of neutrino wakes, extending the work presented in Chapter 5, while also rephrasing it in terms of a general warm sub-component of the dark matter. In this

work we also perform more realistic forecasts for the observability of these warm dark matter wakes. We are also pursuing the measurement of neutrino wakes effects from simulations of nonlinear structure formation.

- I plan on extending the methodology developed in Chapter 7, to promote the analytic estimate of the sound speed into a complete calculation that accounts for all relevant physical effects. This would be valuable to produce priors for nuisance parameters in analysis of weak gravitational lensing, which result in further tightening of cosmological constraints. Additionally, there are some interesting follow-ups to the work described in Chapter 6. In the case of just one spatial dimension it is likely that the cosmological perturbation theory for large scale structure in phase space can be summed to all orders in perturbation theory. Also, explicitly working out the Effective Field Theory of Vlasov-Poisson may result in new important insights to the structure of perturbation theory methods.

- Finally, there are a few ongoing efforts which are related to, but are not direct extensions of, the work presented in this thesis. We are studying the potential of recent developments in the path-integral formulation of large scale structure [356, 424], which account for nonperturbative effects, to improve the modeling of cosmological observables. Another direction of ongoing investigation is an extension of ideas introduced in [66], to carefully assess the information content of neutrino mass constraints in galaxy clustering surveys via redshift space distortions.

The efforts detailed in this thesis, and of future investigations, are key to ensure an efficient and robust extraction of fundamental physics (including but not limited to the sum of neutrino masses) with the large number of ongoing and upcoming cosmological surveys. I hope to keep making meaningful contributions to this exciting scientific discipline in the next few decades.

BIBLIOGRAPHY

[1] Marcelo A. Alvarez, Michael Busha, Tom Abel, and Risa H. Wechsler. Connecting Reionization to the Local Universe: Draft version. *Astrophys. J. Lett.*, 703:L167–L171, 2009.

[2] Steven V. W. Beckwith et al. The Hubble Ultra Deep Field. *Astron. J.*, 132:1729–1755, 2006.

[3] Y. Akrami et al. Planck 2018 results. IV. Diffuse component separation. *Astron. Astrophys.*, 641:A4, 2020.

[4] Mark Vogelsberger, Shy Genel, Volker Springel, Paul Torrey, Debora Sijacki, Dandan Xu, Gregory F. Snyder, Dylan Nelson, and Lars Hernquist. Introducing the Illustris Project: Simulating the coevolution of dark and visible matter in the Universe. *Mon. Not. Roy. Astron. Soc.*, 444(2):1518–1547, 2014.

[5] Robert M. Wald. *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984.

[6] Giorgio Arcadi, Maíra Dutra, Pradipta Ghosh, Manfred Lindner, Yann Mambrini, Mathias Pierre, Stefano Profumo, and Farinaldo S. Queiroz. The waning of the WIMP? A review of models, searches, and constraints. *Eur. Phys. J. C*, 78(3):203, 2018.

[7] Masha Baryakhtar, Leslie Rosenberg, and Gray Rybka. Searching for the QCD Dark Matter Axion. 4 2025.

[8] Anne M. Green and Bradley J. Kavanagh. Primordial Black Holes as a dark matter candidate. *J. Phys. G*, 48(4):043001, 2021.

[9] Jerome Martin. Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask). *Comptes Rendus Physique*, 13:566–665, 2012.

[10] Kensuke Akita and Masahide Yamaguchi. A precision calculation of relic neutrino decoupling. *JCAP*, 08:012, 2020.

[11] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[12] George Efstathiou and Steven Gratton. The evidence for a spatially flat Universe. *Mon. Not. Roy. Astron. Soc.*, 496(1):L91–L95, 2020.

[13] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. Planck evidence for a closed Universe and a possible crisis for cosmology. *Nature Astron.*, 4(2):196–203, 2019.

[14] Shi-Fan Chen and Matias Zaldarriaga. It's All `Ok`: Curvature in Light of BAO from DESI DR2. 5 2025.

[15] Viatcheslav F. Mukhanov and G. V. Chibisov. Quantum Fluctuations and a Nonsingular Universe. *JETP Lett.*, 33:532–535, 1981.

[16] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Phys. Rev. D*, 23:347–356, 1981.

[17] Daniel Baumann and Daniel Green. The power of locality: primordial non-Gaussianity at the map level. *JCAP*, 08(08):061, 2022.

[18] Ivana Babić, Fabian Schmidt, and Beatriz Tucci. Forward vs Backward: Improving BAO Constraints with Field-Level Inference. 5 2025.

[19] Chung-Pei Ma and Edmund Bertschinger. Cosmological perturbation theory in the synchronous and conformal Newtonian gauges. *Astrophys. J.*, 455:7–25, 1995.

[20] Alan Kogut et al. The Primordial Inflation Explorer (PIXIE): mission design and science goals. *JCAP*, 04:020, 2025.

[21] B. Maffei et al. BISOU: A balloon project to measure the CMB spectral distortions. In *16th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories*, 10 2021.

[22] Bryce Cyr. CMB Spectral Distortions: A Multimessenger Probe of the Primordial Universe. In *58th Rencontres de Moriond on Cosmology*, 6 2024.

[23] James M. Bardeen, J. R. Bond, Nick Kaiser, and A. S. Szalay. The Statistics of Peaks of Gaussian Random Fields. *Astrophys. J.*, 304:15–61, 1986.

[24] Gabriel Jung, Michele Citran, Bartjan van Tent, Léa Dumilly, and Nabila Aghanim. Constraints on primordial non-Gaussianity from Planck PR4 data. 4 2025.

[25] Oliver H. E. Philcox. Searching for inflationary physics with the CMB trispectrum. III. Constraints from Planck. *Phys. Rev. D*, 111(12):123534, 2025.

[26] Antony Lewis and Anthony Challinor. Weak gravitational lensing of the CMB. *Phys. Rept.*, 429:1–65, 2006.

[27] R. A. Sunyaev and Ya. B. Zeldovich. The Observations of relic radiation as a test of the nature of X-Ray radiation from the clusters of galaxies. *Comments Astrophys. Space Phys.*, 4:173–178, 1972.

[28] R. K. Sachs and A. M. Wolfe. Perturbations of a cosmological model and angular variations of the microwave background. *Astrophys. J.*, 147:73–90, 1967.

[29] Craig J. Copi, Dragan Huterer, Dominik J. Schwarz, and Glenn D. Starkman. Large angle anomalies in the CMB. *Adv. Astron.*, 2010:847541, 2010.

[30] Volker Springel, Rüdiger Pakmor, Oliver Zier, and Martin Reinecke. Simulating cosmic structure formation with the gadget-4 code. *Mon. Not. Roy. Astron. Soc.*, 506(2):2871–2949, 2021.

[31] Julian Adamek, David Daverio, Ruth Durrer, and Martin Kunz. gevolution: a cosmological N-body code based on General Relativity. *JCAP*, 07:053, 2016.

[32] Cristian Barrera-Hinojosa and Baojiu Li. GRAMSES: a new route to general relativistic N-body simulations in cosmology. Part I. Methodology and code description. *JCAP*, 01:007, 2020.

[33] Hayley J. Macpherson, Daniel J. Price, and Paul D. Lasky. Einstein's Universe: Cosmological structure formation in numerical relativity. *Phys. Rev. D*, 99(6):063522, 2019.

[34] Nora Elisa Chisari and Matias Zaldarriaga. Connection between Newtonian simulations and general relativity. *Phys. Rev. D*, 83:123505, 2011. [Erratum: Phys.Rev.D 84, 089901 (2011)].

[35] Christian Fidler, Cornelius Rampf, Thomas Tram, Robert Crittenden, Kazuya Koyama, and David Wands. General relativistic corrections to N-body simulations and the Zel'dovich approximation. *Phys. Rev. D*, 92(12):123517, 2015.

[36] Christian Fidler, Thomas Tram, Cornelius Rampf, Robert Crittenden, Kazuya Koyama, and David Wands. General relativistic weak-field limit and Newtonian N-body simulations. *JCAP*, 12:022, 2017.

[37] Aurel Schneider, Romain Teyssier, Doug Potter, Joachim Stadel, Julian Onions, Darren S. Reed, Robert E. Smith, Volker Springel, Frazer R. Pearce, and Roman Scoccimarro. Matter power spectrum and the challenge of percent accuracy. *JCAP*, 04:047, 2016.

[38] Jordan Carlson, Martin White, and Nikhil Padmanabhan. Critical look at cosmological perturbation theory techniques. *Physical Review D—Particles, Fields, Gravitation, and Cosmology*, 80(4):043531, 2009.

[39] Asantha Cooray and Ravi K. Sheth. Halo Models of Large Scale Structure. *Phys. Rept.*, 372:1–129, 2002.

[40] Andrew R. Zentner. The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth. *Int. J. Mod. Phys. D*, 16:763–816, 2007.

[41] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro. Large scale structure of the universe and cosmological perturbation theory. *Phys. Rept.*, 367:1–248, 2002.

[42] Daniel Baumann, Alberto Nicolis, Leonardo Senatore, and Matias Zaldarriaga. Cosmological Non-Linearities as an Effective Fluid. *JCAP*, 07:051, 2012.

[43] John Joseph M. Carrasco, Mark P. Hertzberg, and Leonardo Senatore. The Effective Field Theory of Cosmological Large Scale Structures. *JHEP*, 09:082, 2012.

[44] Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. Large-Scale Galaxy Bias. *Phys. Rept.*, 733:1–193, 2018.

[45] Nicholas Frontiere, Katrin Heitmann, Esteban Rangel, Patricia Larsen, Adrian Pope, Imran Sultan, Thomas Uram, Salman Habib, Silvio Rizzi, and Joe Insley. Farpoint: A High-resolution Cosmology Simulation at the Gigaparsec Scale. *Astrophys. J. Supp.*, 259(1):15, 2022.

[46] M. Knabenhans et al. Euclid preparation: IX. EuclidEmulator2 – power spectrum emulation with massive neutrinos and self-consistent dark energy perturbations. *Mon. Not. Roy. Astron. Soc.*, 505(2):2840–2869, 2021.

[47] Diego Blas, Julien Lesgourgues, and Thomas Tram. The cosmic linear anisotropy

solving system (class). part ii: approximation schemes. *Journal of Cosmology and Astroparticle Physics*, 2011(07):034, 2011.

[48] Tobias Baldauf, Mehrdad Mirbabayi, Marko Simonović, and Matias Zaldarriaga. Equivalence Principle and the Baryon Acoustic Peak. *Phys. Rev. D*, 92(4):043514, 2015.

[49] Leonardo Senatore and Matias Zaldarriaga. The IR-resummed Effective Field Theory of Large Scale Structures. *JCAP*, 02:013, 2015.

[50] Zvonimir Vlah, Uroš Seljak, Man Yat Chu, and Yu Feng. Perturbation theory, effective field theory, and oscillations in the power spectrum. *JCAP*, 03:057, 2016.

[51] Diego Blas, Mathias Garny, Mikhail M. Ivanov, and Sergey Sibiryakov. Time-Sliced Perturbation Theory II: Baryon Acoustic Oscillations and Infrared Resummation. *JCAP*, 07:028, 2016.

[52] Guido D'Amico, Yaniv Donath, Matthew Lewandowski, Leonardo Senatore, and Pierre Zhang. The BOSS bispectrum analysis at one loop from the Effective Field Theory of Large-Scale Structure. *JCAP*, 05:059, 2024.

[53] Mikhail M. Ivanov, Marko Simonović, and Matias Zaldarriaga. Cosmological Parameters from the BOSS Galaxy Power Spectrum. *JCAP*, 05:042, 2020.

[54] Shi-Fan Chen, Martin White, Joseph DeRose, and Nickolas Kokron. Cosmological analysis of three-dimensional BOSS galaxy clustering and Planck CMB lensing cross correlations via Lagrangian perturbation theory. *JCAP*, 07(07):041, 2022.

[55] A. G. Adame et al. DESI 2024 VII: cosmological constraints from the full-shape modeling of clustering measurements. *JCAP*, 07:028, 2025.

[56] Maria Concepcion Gonzalez-Garcia, Michele Maltoni, and Thomas Schwetz. NuFIT: Three-Flavour Global Analyses of Neutrino Oscillation Experiments. *Universe*, 7(12):459, 2021.

[57] Max Aker et al. Direct neutrino-mass measurement based on 259 days of KATRIN data. *Science*, 388(6743):adq9592, 2025.

[58] Sunny Vagnozzi, Suhail Dhawan, Martina Gerbino, Katherine Freese, Ariel Goobar, and Olga Mena. Constraints on the sum of the neutrino masses in dynamical dark energy models with $w(z) \geq -1$ are tighter than those obtained in ΛCDM. *Phys. Rev. D*, 98(8):083501, 2018.

[59] Mariana Vargas-Magana, David D Brooks, Michael M Levi, and Gregory G Tarle. Unraveling the universe with desi. *arXiv preprint arXiv:1901.01581*, 2019.

[60] Željko Ivezić et al. LSST: from Science Drivers to Reference Design and Anticipated Data Products. *Astrophys. J.*, 873(2):111, 2019.

[61] Peter Ade et al. The Simons Observatory: Science goals and forecasts. *JCAP*, 02:056, 2019.

[62] Y. Mellier et al. Euclid. I. Overview of the Euclid mission. *Astron. Astrophys.*, 697:A1, 2025.

[63] Cora Dvorkin et al. The Physics of Light Relics. In *Snowmass 2021*, 3 2022.

[64] Julien Lesgourgues and Sergio Pastor. Massive neutrinos and cosmology. *Phys. Rept.*, 429:307–379, 2006.

[65] Jacob Brandbyge, Steen Hannestad, Troels Haugbølle, and Bjarne Thomsen. The Effect of Thermal Neutrino Motion on the Non-linear Cosmological Matter Power Spectrum. *JCAP*, 08:020, 2008.

[66] Marilena Loverde and Zachary J. Weiner. Massive neutrinos and cosmic composition. *JCAP*, 12:048, 2024.

[67] W. Elbers et al. Constraints on Neutrino Physics from DESI DR2 BAO and DR1 Full Shape. 3 2025.

[68] Erminia Calabrese et al. The Atacama Cosmology Telescope: DR6 Constraints on Extended Cosmological Models. 3 2025.

[69] E. Camphuis et al. SPT-3G D1: CMB temperature and polarization power spectra and cosmology from 2019 and 2020 observations of the SPT-3G Main field. 6 2025.

[70] M. Abdul Karim et al. DESI DR2 Results II: Measurements of Baryon Acoustic Oscillations and Cosmological Constraints. 3 2025.

[71] Eleonora Di Valentino, Olga Mena, Supriya Pan, Luca Visinelli, Weiqiang Yang, Alessandro Melchiorri, David F. Mota, Adam G. Riess, and Joseph Silk. In the realm of the Hubble tension—a review of solutions. *Class. Quant. Grav.*, 38(15):153001, 2021.

[72] Adam G Riess, Wenlong Yuan, Lucas M Macri, Dan Scolnic, Dillon Brout, Stefano Casertano, David O Jones, Yukei Murakami, Gagandeep S Anand, Louise Breuval, et al. A comprehensive measurement of the local value of the hubble constant with 1 km s- 1 mpc- 1 uncertainty from the hubble space telescope and the sh0es team. *The Astrophysical journal letters*, 934(1):L7, 2022.

[73] Wendy L. Freedman, Barry F. Madore, Taylor J. Hoyt, In Sung Jang, Abigail J. Lee, and Kayla A. Owens. Status Report on the Chicago-Carnegie Hubble Program (CCHP): Measurement of the Hubble Constant Using the Hubble and James Webb Space Telescopes. *Astrophys. J.*, 985(2):203, 2025.

[74] Nathaniel Craig, Daniel Green, Joel Meyers, and Surjeet Rajendran. No νs is Good News. *JHEP*, 09:097, 2024.

[75] K. Lodha et al. Extended Dark Energy analysis using DESI DR2 BAO measurements. 3 2025.

[76] Noah Sailer, Gerrit S. Farren, Simone Ferraro, and Martin White. Dispuτable: the high cost of a low optical depth. 4 2025.

[77] Tanisha Jhaveri, Tanvi Karwal, and Wayne Hu. Turning a negative neutrino mass into a positive optical depth. 4 2025.

[78] Caio Bastos de Senna Nascimento and Marilena Loverde. Neutrinos in N-body simulations. *Phys. Rev. D*, 104(4):043512, 2021.

[79] Caio Bastos de Senna Nascimento. Generalized boltzmann hierarchy for massive neutrinos in cosmology. *Physical Review D*, 104(8):083535, 2021.

[80] Caio Nascimento. Accurate fluid approximation for massive neutrinos in cosmology. *Phys. Rev. D*, 108(2):023505, 2023.

[81] Caio Nascimento and Marilena Loverde. Neutrino winds on the sky. *JCAP*, 11:036, 2023.

[82] Caio Nascimento and Marilena Loverde. Cosmological perturbation theory for large scale structure in phase space. *JCAP*, 06:002, 2025.

[83] Caio Nascimento, Drew Jamieson, Matthew McQuinn, and Marilena Loverde. A semi-analytic estimate for the effective sound speed counterterm in the EFTofLSS. *JCAP*, 02:023, 2025.

[84] P. F. de Salas, D. V. Forero, C. A. Ternes, M. Tortola, and J. W. F. Valle. Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity. *Phys. Lett. B*, 782:633–640, 2018.

[85] P. F. De Salas, S. Gariazzo, O. Mena, C. A. Ternes, and M. Tórtola. Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects. *Front. Astron. Space Sci.*, 5:36, 2018.

[86] M. Aker et al. Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN. *Phys. Rev. Lett.*, 123(22):221802, 2019.

[87] Patrick Stöcker et al. Strengthening the bound on the mass of the lightest neutrino with terrestrial and cosmological experiments. *Phys. Rev. D*, 103(12):123508, 2021.

[88] Arthur Loureiro et al. On The Upper Bound of Neutrino Masses from Combined Cosmological Observations and Particle Physics Experiments. *Phys. Rev. Lett.*, 123(8):081301, 2019.

[89] Hyun Min Lee. Lectures on physics beyond the Standard Model. *J. Korean Phys. Soc.*, 78(11):985–1017, 2021.

[90] Saurabh W. Jha et al. Next Generation LSST Science. 7 2019.

[91] A. Blanchard et al. Euclid preparation. VII. Forecast validation for Euclid cosmological probes. *Astron. Astrophys.*, 642:A191, 2020.

[92] Neelima Sehgal et al. CMB-HD: An Ultra-Deep, High-Resolution Millimeter-Wave Survey Over Half the Sky. *Bull. Am. Astron. Soc.*, 51(7):1–23, 2019.

[93] D. Spergel et al. Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report. 5 2013.

[94] Kyle S. Dawson et al. The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. *Astron. J.*, 151:44, 2016.

[95] L. E. H. Godfrey, H. Bignall, S. Tingay, L. Harvey-Smith, M. Kramer, S. Burke-Spolaor, J. C. A. Miller-Jones, M. Johnston-Hollitt, R. Ekers, and S. Gulyaev. Very

High Angular Resolution Science with the Square Kilometre Array. *Publ. Astron. Soc. Austral.*, 29:42, 2012.

[96] Michael E. Levi et al. The Dark Energy Spectroscopic Instrument (DESI). 7 2019.

[97] Cora Dvorkin et al. Neutrino Mass from Cosmology: Probing Physics Beyond the Standard Model. 3 2019.

[98] Thejs Brinckmann, Deanna C. Hooper, Maria Archidiacono, Julien Lesgourgues, and Tim Sprenger. The promising future of a robust cosmological neutrino mass measurement. *JCAP*, 01:059, 2019.

[99] Francesco De Bernardis, Thomas D. Kitching, Alan Heavens, and Alessandro Melchiorri. Determining the Neutrino Mass Hierarchy with Cosmology. *Phys. Rev. D*, 80:123509, 2009.

[100] Anže Slosar et al. Dark Energy and Modified Gravity. 3 2019.

[101] Daniel Green et al. Messengers from the Early Universe: Cosmic Neutrinos and Other Light Relics. *Bull. Am. Astron. Soc.*, 51(3):159, 2019.

[102] Douglas Potter, Joachim Stadel, and Romain Teyssier. PKDGRAV3: beyond trillion particle cosmological simulations for the next era of galaxy surveys. *Comput. Astrophys. Cosmol.*, 4(1):2, 2017.

[103] Matteo Viel, Martin G. Haehnelt, and Volker Springel. The effect of neutrinos on the matter distribution as probed by the Intergalactic Medium. *JCAP*, 06:015, 2010.

[104] J. D. Emberson et al. Cosmological neutrino simulations at extreme scale. *Res. Astron. Astrophys.*, 17(8):085, 2017.

[105] Adrian E. Bayer, Arka Banerjee, and Yu Feng. A fast particle-mesh simulation of non-linear cosmological structure formation with massive neutrinos. *JCAP*, 01:016, 2021.

[106] Emanuele Castorina, Carmelita Carbone, Julien Bel, Emiliano Sefusatti, and Klaus Dolag. DEMNUni: The clustering of large-scale structures in the presence of massive neutrinos. *JCAP*, 07:043, 2015.

[107] Svetlin Tassev, Matias Zaldarriaga, and Daniel Eisenstein. Solving Large Scale Structure in Ten Easy Steps with COLA. *JCAP*, 06:036, 2013.

[108] Joe Zhiyu Chen, Amol Upadhye, and Yvonne Y. Y. Wong. One line to run them all: SuperEasy massive neutrino linear response in N-body simulations. *JCAP*, 04:078, 2021.

[109] Cornelius Rampf and Gerasimos Rigopoulos. Zel'dovich Approximation and General Relativity. *Mon. Not. Roy. Astron. Soc.*, 430:L54–L58, 2013.

[110] David Daverio, Yves Dirian, and Ermis Mitsou. General relativistic cosmological N-body simulations. Part I. Time integration. *JCAP*, 10:065, 2019.

[111] Julian Adamek, Ruth Durrer, and Martin Kunz. Relativistic N-body simulations with massive neutrinos. *JCAP*, 11:004, 2017.

[112] Arka Banerjee, Devon Powell, Tom Abel, and Francisco Villaescusa-Navarro. Reducing Noise in Cosmological N-body Simulations with Neutrinos. *JCAP*, 09:028, 2018.

[113] Simeon Bird, Matteo Viel, and Martin G. Haehnelt. Massive Neutrinos and the Non-linear Matter Power Spectrum. *Mon. Not. Roy. Astron. Soc.*, 420:2551–2561, 2012.

[114] Francisco Villaescusa-Navarro, Federico Marulli, Matteo Viel, Enzo Branchini, Emanuele Castorina, Emiliano Sefusatti, and Shun Saito. Cosmology with massive neutrinos I: towards a realistic modeling of the relation between matter, haloes and galaxies. *JCAP*, 03:011, 2014.

[115] Francisco Villaescusa-Navarro, Arka Banerjee, Neal Dalal, Emanuele Castorina, Roman Scoccimarro, Raul Angulo, and David N. Spergel. The imprint of neutrinos on clustering in redshift-space. *Astrophys. J.*, 861(1):53, 2018.

[116] Graziano Rossi. The Sejong Suite: Cosmological Hydrodynamical Simulations with Massive Neutrinos, Dark Radiation, and Warm Dark Matter. *Astrophys. J. Suppl.*, 249(2):19, 2020.

[117] Jazhiel Chacon, J. Alberto Vazquez, and Ruslan Gabbasov. Dark matter with n-body numerical simulations. *Rev. Mex. Fis. E*, 17(2):241–254, 2020.

[118] Volker Springel. The Cosmological simulation code GADGET-2. *Mon. Not. Roy. Astron. Soc.*, 364:1105–1134, 2005.

[119] Derek Inman and Hao-ran Yu. Simulating the Cosmic Neutrino Background using Collisionless Hydrodynamics. *Astrophys. J. Suppl.*, 250(1):21, 2020.

[120] Matteo Zennaro, Julien Bel, Francisco Villaescusa-Navarro, Carmelita Carbone, Emiliano Sefusatti, and Luigi Guzzo. Initial Conditions for Accurate N-Body Simulations of Massive Neutrino Cosmologies. *Mon. Not. Roy. Astron. Soc.*, 466(3):3244–3258, 2017.

[121] Christian Fidler, Thomas Tram, Cornelius Rampf, Robert Crittenden, Kazuya Koyama, and David Wands. Relativistic Interpretation of Newtonian Simulations for Cosmic Structure Formation. *JCAP*, 09:031, 2016.

[122] Thomas Tram, Jacob Brandbyge, Jeppe Dakin, and Steen Hannestad. Fully relativistic treatment of light neutrinos in N-body simulations. *JCAP*, 03:022, 2019.

[123] Chi-Ting Chiang, Marilena LoVerde, and Francisco Villaescusa-Navarro. First detection of scale-dependent linear halo bias in N-body simulations with massive neutrinos. *Phys. Rev. Lett.*, 122(4):041302, 2019.

[124] Christian Partmann, Christian Fidler, Cornelius Rampf, and Oliver Hahn. Fast simulations of cosmic large-scale structure with massive neutrinos. *JCAP*, 09:018, 2020.

[125] Christian Fidler, Alexander Kleinjohann, Thomas Tram, Cornelius Rampf, and Kazuya Koyama. A new approach to cosmological structure formation with massive neutrinos. *JCAP*, 01:025, 2019.

[126] Guilherme Brando, Kazuya Koyama, and David Wands. Relativistic Corrections to the Growth of Structure in Modified Gravity. *JCAP*, 01:013, 2021.

[127] Julien Lesgourgues and Thomas Tram. The cosmic linear anisotropy solving system (class) iv: efficient implementation of non-cold relics. *Journal of Cosmology and Astroparticle Physics*, 2011(09):032, 2011.

[128] Francisco Villaescusa-Navarro et al. The Quijote simulations. *Astrophys. J. Suppl.*, 250(1):2, 2020.

[129] Masatoshi Shoji and Eiichiro Komatsu. Massive neutrinos in cosmology: analytic solutions and fluid approximation. *Physical Review D*, 81(12):123516, 2010.

[130] Martina Gerbino. Neutrino properties from cosmology. In *Prospects in Neutrino Physics*, pages 52–52, 3 2018.

[131] Julien Lesgourgues and Sergio Pastor. Neutrino cosmology and Planck. *New J. Phys.*, 16:065002, 2014.

[132] Yacine Ali-Haimoud and Simeon Bird. An efficient implementation of massive neutrinos in non-linear structure formation simulations. *Mon. Not. Roy. Astron. Soc.*, 428:3375–3389, 2012.

[133] Ya. B. Zeldovich. Gravitational instability: An Approximate theory for large density perturbations. *Astron. Astrophys.*, 5:84–89, 1970.

[134] F. R. Bouchet, S. Colombi, E. Hivon, and R. Juszkiewicz. Perturbative Lagrangian approach to gravitational instability. *Astron. Astrophys.*, 296:575, 1995.

[135] Michaël Michaux, Oliver Hahn, Cornelius Rampf, and Raul E. Angulo. Accurate initial conditions for cosmological N-body simulations: Minimizing truncation and discreteness errors. *Mon. Not. Roy. Astron. Soc.*, 500(1):663–683, 2020.

[136] Diego Blas, Mathias Garny, Thomas Konstandin, and Julien Lesgourgues. Structure formation with massive neutrinos: going beyond linear theory. *JCAP*, 11:039, 2014.

[137] Christian Fidler, Thomas Tram, Cornelius Rampf, Robert Crittenden, Kazuya Koyama, and David Wands. Relativistic initial conditions for N-body simulations. *JCAP*, 06:043, 2017.

[138] Christian Fidler and Alexander Kleinjohann. Suitable Initial Conditions for Newtonian Simulations with Massive Neutrinos. *JCAP*, 06:018, 2019.

[139] Marilena LoVerde and Matias Zaldarriaga. Neutrino clustering around spherical dark matter halos. *Phys. Rev. D*, 89(6):063502, 2014.

[140] Joe Zhiyu Chen, Amol Upadhye, and Yvonne Y. Y. Wong. The cosmic neutrino background as a collection of fluids in large-scale structure simulations. *JCAP*, 03:065, 2021.

[141] Jasjeet Singh Bagla. Cosmological N-body simulation: Techniques, scope and status. *Curr. Sci.*, 88:1088, 2005.

[142] Antony Lewis and Anthony Challinor. Camb: Code for anisotropies in the microwave background. *ascl*, pages ascl–1102, 2011.

[143] Wayne Hu. Structure formation with generalized dark matter. *The Astrophysical Journal*, 506(2):485, 1998.

[144] Antony Lewis and Anthony Challinor. Evolution of cosmological dark matter perturbations. *Phys. Rev. D*, 66:023531, 2002.

[145] Edmund Bertschinger. The Effects of Cold Dark Matter Decoupling and Pair Annihilation on Cosmological Perturbations. *Phys. Rev. D*, 74:063509, 2006.

[146] Cullan Howlett, Antony Lewis, Alex Hall, and Anthony Challinor. CMB power spectrum parameter degeneracies in the era of precision cosmology. *JCAP*, 04:027, 2012.

[147] Jeppe Dakin, Jacob Brandbyge, Steen Hannestad, Troels Haugbølle, and Thomas Tram. νconcept: Cosmological neutrino simulations from the non-linear boltzmann hierarchy. *Journal of Cosmology and Astroparticle Physics*, 2019(02):052, 2019.

[148] Maria Archidiacono and Steen Hannestad. Efficient calculation of cosmological neutrino clustering in the non-linear regime. *JCAP*, 06:018, 2016.

[149] P. F. de Salas, D. V. Forero, S. Gariazzo, P. Martínez-Miravé, O. Mena, C. A. Ternes, M. Tórtola, and J. W. F. Valle. 2020 global reassessment of the neutrino oscillation picture. *JHEP*, 02:071, 2021.

[150] Francesco Capozzi, Eleonora Di Valentino, Eligio Lisi, Antonio Marrone, Alessandro Melchiorri, and Antonio Palazzo. Addendum to "global constraints on absolute neutrino masses and their ordering". *Physical Review D*, 101(11):116013, 2020.

[151] Ivan Esteban, Maria Conceptión González-García, Michele Maltoni, Thomas Schwetz, and Albert Zhou. The fate of hints: updated global analysis of three-flavor neutrino oscillations. *Journal of High Energy Physics*, 2020(9):1–22, 2020.

[152] M Aker, M Balzer, D Batzler, A Beglarian, J Behrens, A Berlev, U Besserer, M Biassoni, B Bieringer, F Block, et al. Katrin: Status and prospects for the neutrino mass and beyond. *arXiv preprint arXiv:2203.08059*, 2022.

[153] Steen Hannestad. Neutrino masses and the dark energy equation of state: Relaxing the cosmological neutrino mass bound. *Physical Review Letters*, 95(22):221301, 2005.

[154] Miguel Escudero, Jacobo Lopez-Pavon, Nuria Rius, and Stefan Sandner. Relaxing cosmological neutrino mass bounds with unstable neutrinos. *Journal of High Energy Physics*, 2020(12):1–44, 2020.

[155] Zackaria Chacko, Abhish Dev, Peizhi Du, Vivian Poulin, and Yuhsin Tsai. Cosmological limits on the neutrino mass and lifetime. *Journal of High Energy Physics*, 2020(4):1–33, 2020.

[156] Amir Aghamousa, Jessica Aguilar, Steve Ahlen, Shadab Alam, Lori E Allen, Carlos Allende Prieto, James Annis, Stephen Bailey, Christophe Balland, Otger Ballester, et al. The desi experiment part i: science, targeting, and survey design. *arXiv preprint arXiv:1611.00036*, 2016.

[157] Paul A Abell, Julius Allison, Scott F Anderson, John R Andrew, J Roger P Angel, Lee Armus, David Arnett, SJ Asztalos, Tim S Axelrod, Stephen Bailey, et al. Lsst science book, version 2.0. *arXiv preprint arXiv:0912.0201*, 2009.

[158] Kevork N Abazajian, Peter Adshead, Zeeshan Ahmed, Steven W Allen, David Alonso, Kam S Arnold, Carlo Baccigalupi, James G Bartlett, Nicholas Battaglia, Bradford A Benson, et al. Cmb-s4 science book. *arXiv preprint arXiv:1610.02743*, 2016.

[159] Daniel Green and Joel Meyers. Cosmological implications of a neutrino mass detection. *arXiv preprint arXiv:2111.01096*, 2021.

[160] Antony Lewis and Anthony Challinor. Code for anisotropies in the microwave background. *Documentation, January. http://camb. info*, 2017.

[161] Lingyuan Ji, Marc Kamionkowski, and Jose Luis Bernal. Cosmological perturbations: non-cold relics without the boltzmann hierarchy. *arXiv preprint arXiv:2201.11129*, 2022.

[162] Derek Inman and Ue-Li Pen. Cosmic neutrinos: dispersive and non-linear. *arXiv preprint arXiv:1609.09469*, 2016.

[163] Andreas Ringwald and Yvonne YY Wong. Gravitational clustering of relic neutrinos and implications for their detection. *Journal of Cosmology and Astroparticle Physics*, 2004(12):005, 2004.

[164] Thejs Brinckmann, Jae Hyeok Chang, Peizhi Du, and Marilena LoVerde. Confronting interacting dark radiation scenarios with cosmological data. *arXiv preprint arXiv:2212.13264*, 2022.

[165] Isabel M Oldengott, Thomas Tram, Cornelius Rampf, and Yvonne YY Wong. Interacting neutrinos in cosmology: exact description and constraints. *Journal of Cosmology and Astroparticle Physics*, 2017(11):027, 2017.

[166] Francesco Capozzi, Eleonora Di Valentino, Eligio Lisi, Antonio Marrone, Alessandro Melchiorri, and Antonio Palazzo. Global constraints on absolute neutrino masses and their ordering. *Phys. Rev. D*, 95(9):096014, 2017. [Addendum: Phys.Rev.D 101, 116013 (2020)].

[167] James N. Fry. Dynamical friction and massive neutrinos. *eConf*, C801002:81–83, 1980.

[168] Chiamaka Okoli, Morag I. Scrimgeour, Niayesh Afshordi, and Michael J. Hudson. Dynamical friction in the primordial neutrino sea. *Mon. Not. Roy. Astron. Soc.*, 468(2):2164–2175, 2017.

[169] Hong-Ming Zhu, Ue-Li Pen, Xuelei Chen, and Derek Inman. Probing Neutrino Hierarchy and Chirality via Wakes. *Phys. Rev. Lett.*, 116(14):141301, 2016.

[170] Albert Stebbins, Shoba Veeraraghavan, Robert H. Brandenberger, Joseph Silk, and Neil Turok. Cosmic String Wakes. *Astrophys. J.*, 322:1–19, 1987.

[171] Hong-Ming Zhu, Ue-Li Pen, Xuelei Chen, Derek Inman, and Yu Yu. Measurement of Neutrino Masses from Relative Velocities. *Phys. Rev. Lett.*, 113:131301, 2014.

[172] Hong-Ming Zhu and Emanuele Castorina. Measuring dark matter-neutrino relative velocity on cosmological scales. *Phys. Rev. D*, 101(2):023525, 2020.

[173] Derek Inman, J. D. Emberson, Ue-Li Pen, Alban Farchi, Hao-Ran Yu, and Joachim Harnois-Déraps. Precision reconstruction of the cold dark matter-neutrino relative velocity from N-body simulations. *Phys. Rev. D*, 92(2):023502, 2015.

[174] Dmitriy Tseliakhovich and Christopher Hirata. Relative velocity of dark matter and baryonic fluids and the formation of the first structures. *Phys. Rev. D*, 82:083520, 2010.

[175] Francisco Villaescusa-Navarro, Simeon Bird, Carlos Pena-Garay, and Matteo Viel. Non-linear evolution of the cosmic neutrino background. *JCAP*, 03:019, 2013.

[176] Fabian Zimmer, Camila A. Correa, and Shin'ichiro Ando. Influence of local structure on relic neutrino abundances and anisotropies. 6 2023.

[177] Jacob Brandbyge and Steen Hannestad. Grid Based Linear Neutrino Perturbations in Cosmological N-body Simulations. *JCAP*, 05:002, 2009.

[178] Chi-Ting Chiang, Wayne Hu, Yin Li, and Marilena Loverde. Scale-dependent bias and bispectrum in neutrino separate universe simulations. *Phys. Rev. D*, 97(12):123526, 2018.

[179] Julien Lesgourgues, Sabino Matarrese, Massimo Pietroni, and Antonio Riotto. Non-linear Power Spectrum including Massive Neutrinos: the Time-RG Flow Approach. *JCAP*, 06:017, 2009.

[180] Helene Dupuy and Francis Bernardeau. Describing massive neutrinos in cosmology as a collection of independent flows. *JCAP*, 01:030, 2014.

[181] Florian Führer and Yvonne Y. Y. Wong. Higher-order massive neutrino perturbations in large-scale structure. *JCAP*, 03:046, 2015.

[182] Michele Levi and Zvonimir Vlah. Massive neutrinos in nonlinear large scale structure: A consistent perturbation theory. 5 2016.

[183] Leonardo Senatore and Matias Zaldarriaga. The Effective Field Theory of Large-Scale Structure in the presence of Massive Neutrinos. 7 2017.

[184] Derek Inman, Hao-Ran Yu, Hong-Ming Zhu, J. D. Emberson, Ue-Li Pen, Tong-Jie Zhang, Shuo Yuan, Xuelei Chen, and Zhi-Zhong Xing. Simulating the cold dark matter-neutrino dipole with TianNu. *Phys. Rev. D*, 95(8):083518, 2017.

[185] Kohji Yoshikawa, Satoshi Tanaka, Naoki Yoshida, and Shun Saito. Cosmological Vlasov–Poisson Simulations of Structure Formation with Relic Neutrinos: Nonlinear Clustering and the Neutrino Mass. *Astrophys. J.*, 904(2):159, 2020.

[186] Marilena LoVerde. Neutrino mass without cosmic variance. *Phys. Rev. D*, 93(10):103526, 2016.

[187] Byeonghee Yu, Robert Z. Knight, Blake D. Sherwin, Simone Ferraro, Lloyd Knox, and Marcel Schmittfull. Toward neutrino mass from cosmology without optical depth information. *Phys. Rev. D*, 107(12):123522, 2023.

[188] Mario Ballardini and Roy Maartens. Constraining the neutrino mass using a multi-tracer combination of two galaxy surveys and cosmic microwave background lensing. *Mon. Not. Roy. Astron. Soc.*, 510(3):4295–4301, 2022.

[189] Marcel Schmittfull and Uros Seljak. Parameter constraints from cross-correlation of CMB lensing with galaxy clustering. *Phys. Rev. D*, 97(12):123540, 2018.

[190] Isabelle Tanseri, Steffen Hagstotz, Sunny Vagnozzi, Elena Giusarma, and Katherine

Freese. Updated neutrino mass constraints from galaxy clustering and CMB lensing-galaxy cross-correlation measurements. *JHEAp*, 36:1–26, 2022.

[191] Frank J. Qu, Blake D. Sherwin, Omar Darwish, Toshiya Namikawa, and Mathew S. Madhavacheril. Probing early structure and model-independent neutrino mass with high-redshift CMB lensing mass maps. 8 2022.

[192] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis. Towards a cosmological neutrino mass detection. *Phys. Rev. D*, 92(12):123535, 2015.

[193] Siddharth Mishra-Sharma, David Alonso, and Joanna Dunkley. Neutrino masses and beyond- ΛCDM cosmology with LSST and future CMB experiments. *Phys. Rev. D*, 97(12):123544, 2018.

[194] Robert H. Brandenberger, Nick Kaiser, and N. Turok. Dissipationless Clustering of Neutrinos Around a Cosmic String Loop. *Phys. Rev. D*, 36:2242, 1987.

[195] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, rev - revised, 2 edition, 2008.

[196] Subrahmanyan Chandrasekhar. Dynamical friction. i. general considerations: the coefficient of dynamical friction. *Astrophysical Journal*, 97:255–262, 1943.

[197] A. J. S. Hamilton. Formulae for growth factors in expanding universes containing matter and a cosmological constant. *Mon. Not. Roy. Astron. Soc.*, 322:419, 2001.

[198] Goran Jelic-Cizmek, Francesca Lepori, Julian Adamek, and Ruth Durrer. The generation of vorticity in cosmological N-body simulations. *JCAP*, 09:006, 2018.

[199] Chung-Pei Ma and James N. Fry. Nonlinear kinetic Sunyaev-Zeldovich effect. *Phys. Rev. Lett.*, 88:211301, 2002.

[200] Wayne Hu. Reionization revisited: secondary cmb anisotropies and polarization. *Astrophys. J.*, 529:12, 2000.

[201] Alexander Mead, Samuel Brieden, Tilman Tröster, and Catherine Heymans. HMcode-2020: Improved modelling of non-linear cosmological power spectra with baryonic feedback. 9 2020.

[202] Shahab Joudaki. Constraints on Neutrino Mass and Light Degrees of Freedom in Extended Cosmological Parameter Spaces. *Phys. Rev. D*, 87:083523, 2013.

[203] Christiane S. Lorenz, Erminia Calabrese, and David Alonso. Distinguishing between Neutrinos and time-varying Dark Energy through Cosmic Time. *Phys. Rev. D*, 96(4):043510, 2017.

[204] Junde Chen, Pengjie Zhang, Yi Zheng, Yu Yu, and Yipeng Jing. Accurate determination of halo velocity bias in simulations and its cosmological implications. *Astrophys. J.*, 861(1):58, 2018.

[205] Lam Hui and Kyle P. Parfrey. The Evolution of Bias: Generalized. *Phys. Rev. D*, 77:043527, 2008.

[206] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Structure of cold dark matter halos. *Astrophys. J.*, 462:563–575, 1996.

[207] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A Universal density profile from hierarchical clustering. *Astrophys. J.*, 490:493–508, 1997.

[208] Aaron A. Dutton and Andrea V. Macciò. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. *Mon. Not. Roy. Astron. Soc.*, 441(4):3359–3374, 2014.

[209] Jeremy L. Tinker, Andrey V. Kravtsov, Anatoly Klypin, Kevork Abazajian, Michael S. Warren, Gustavo Yepes, Stefan Gottlober, and Daniel E. Holz. Toward a halo mass function for precision cosmology: The Limits of universality. *Astrophys. J.*, 688:709–728, 2008.

[210] Jeremy L. Tinker, Brant E. Robertson, Andrey V. Kravtsov, Anatoly Klypin, Michael S. Warren, Gustavo Yepes, and Stefan Gottlober. The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *Astrophys. J.*, 724:878–886, 2010.

[211] Matteo Cataneo, Lucas Lombriser, Catherine Heymans, Alexander Mead, Alexandre Barreira, Sownak Bose, and Baojiu Li. On the road to percent accuracy: non-linear reaction of the matter power spectrum to dark energy and modified gravity. *Mon. Not. Roy. Astron. Soc.*, 488(2):2121–2142, 2019.

[212] Eiichiro Komatsu and David N. Spergel. Acoustic signatures in the primary microwave background bispectrum. *Phys. Rev. D*, 63:063002, 2001.

[213] Olivier Doré et al. Cosmology with the SPHEREX All-Sky Spectral Survey. 12 2014.

[214] David J. Schlegel et al. A Spectroscopic Road Map for Cosmic Frontier: DESI, DESI-II, Stage-5. 9 2022.

[215] Marilena LoVerde. Halo bias in mixed dark matter cosmologies. *Phys. Rev. D*, 90(8):083530, 2014.

[216] Sunny Vagnozzi, Thejs Brinckmann, Maria Archidiacono, Katherine Freese, Martina Gerbino, Julien Lesgourgues, and Tim Sprenger. Bias due to neutrinos must not uncorrect'd go. *JCAP*, 09:001, 2018.

[217] Alvise Raccanelli, Licia Verde, and Francisco Villaescusa-Navarro. Biases from neutrino bias: to worry or not to worry? *Mon. Not. Roy. Astron. Soc.*, 483(1):734–743, 2019.

[218] Matteo Costanzi, Francisco Villaescusa-Navarro, Matteo Viel, Jun-Qing Xia, Stefano Borgani, Emanuele Castorina, and Emiliano Sefusatti. Cosmology with massive neutrinos III: the halo mass function and an application to galaxy clusters. *JCAP*, 12:012, 2013.

[219] Emanuele Castorina, Emiliano Sefusatti, Ravi K. Sheth, Francisco Villaescusa-Navarro, and Matteo Viel. Cosmology with massive neutrinos II: on the universality of the halo mass function and bias. *JCAP*, 02:049, 2014.

[220] Uros Seljak. Extracting primordial non-gaussianity without cosmic variance. *Phys. Rev. Lett.*, 102:021302, 2009.

[221] Donghui Jeong and Eiichiro Komatsu. Primordial non-Gaussianity, scale-dependent bias, and the bispectrum of galaxies. *Astrophys. J.*, 703:1230–1248, 2009.

[222] Christian Wagner, Licia Verde, and Raul Jimenez. Effects of the neutrino mass splitting on the non-linear matter power spectrum. *Astrophys. J. Lett.*, 752:L31, 2012.

[223] Maria Archidiacono, Steen Hannestad, and Julien Lesgourgues. What will it take to measure individual neutrino mass states using cosmology? *JCAP*, 09:021, 2020.

[224] Shadab Alam et al. Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory. *Phys. Rev. D*, 103(8):083533, 2021.

[225] B. Abareshi et al. Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument. *Astron. J.*, 164(5):207, 2022.

[226] Scott Dodelson, Katrin Heitmann, Chris Hirata, Klaus Honscheid, Aaron Roodman, Uroš Seljak, Anže Slosar, and Mark Trodden. Cosmic Visions Dark Energy: Science. 4 2016.

[227] I. Sevilla-Noarbe et al. Dark Energy Survey Year 3 Results: Photometric Data Set for Cosmology. *Astrophys. J. Suppl.*, 254(2):24, 2021.

[228] T. M. C. Abbott et al. The Dark Energy Survey Data Release 1. *Astrophys. J. Suppl.*, 239(2):18, 2018.

[229] Rachel Mandelbaum et al. The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document. 9 2018.

[230] Martin Crocce and Roman Scoccimarro. Renormalized cosmological perturbation theory. *Phys. Rev. D*, 73:063519, 2006.

[231] Zachary Slepian et al. The large-scale three-point correlation function of the SDSS BOSS DR12 CMASS galaxies. *Mon. Not. Roy. Astron. Soc.*, 468(1):1070–1083, 2017.

[232] Zachary Slepian and Daniel Eisenstein. On the signature of the baryon–dark matter relative velocity in the two- and three-point galaxy correlation functions. *Mon. Not. Roy. Astron. Soc.*, 448(1):9–26, 2015.

[233] Jaiyul Yoo, Neal Dalal, and Uros Seljak. Supersonic Relative Velocity Effect on the Baryonic Acoustic Oscillation Measurements. *JCAP*, 07:018, 2011.

[234] Kendrick M. Smith, Mathew S. Madhavacheril, Moritz Münchmeyer, Simone Ferraro, Utkarsh Giri, and Matthew C. Johnson. KSZ tomography and the bispectrum. 10 2018.

[235] Nobuyoshi Makino, Misao Sasaki, and Yasushi Suto. Analytic approach to the perturbative expansion of nonlinear gravitational fluctuations in logical density and velocity fields. *Phys. Rev. D*, 46:585–602, 1992.

[236] Bhuvnesh Jain and Edmund Bertschinger. Second order power spectrum and nonlinear evolution at high redshift. *Astrophys. J.*, 431:495, 1994.

[237] M. H. Goroff, Benjamin Grinstein, S. J. Rey, and Mark B. Wise. Coupling of Modes of Cosmological Mass Density Fluctuations. *Astrophys. J.*, 311:6–14, 1986.

[238] Roman Scoccimarro and Joshua Frieman. Loop corrections in nonlinear cosmological perturbation theory. *Astrophys. J. Suppl.*, 105:37, 1996.

[239] Salvatore Bottaro, Emanuele Castorina, Marco Costa, Diego Redigolo, and Ennio Salvioni. Unveiling Dark Forces with Measurements of the Large Scale Structure of the Universe. *Phys. Rev. Lett.*, 132(20):201002, 2024.

[240] Lorenzo Piga, Marco Marinucci, Guido D'Amico, Massimo Pietroni, Filippo Vernizzi, and Bill S. Wright. Constraints on modified gravity from the BOSS galaxy survey. *JCAP*, 04:038, 2023.

[241] Matthew Lewandowski. Violation of the consistency relations for large-scale structure with dark energy. *JCAP*, 08:044, 2020.

[242] Marco Crisostomi, Matthew Lewandowski, and Filippo Vernizzi. Consistency relations for large-scale structure in modified gravity and the matter bispectrum. *Phys. Rev. D*, 101(12):123501, 2020.

[243] Salvatore Bottaro, Emanuele Castorina, Marco Costa, Diego Redigolo, and Ennio Salvioni. From 100 kpc to 10 Gpc: Dark Matter self-interactions before and after DESI. 7 2024.

[244] Jordan Carlson, Martin White, and Nikhil Padmanabhan. A critical look at cosmological perturbation theory techniques. *Phys. Rev. D*, 80:043531, 2009.

[245] Thomas Buchert. Lagrangian theory of gravitational instability of Friedman-Lemaitre cosmologies: Generic third order model for nonlinear clustering. *Mon. Not. Roy. Astron. Soc.*, 267:811–820, 1994.

[246] Paolo Catelan. Lagrangian dynamics in nonflat universes and nonlinear gravitational evolution. *Mon. Not. Roy. Astron. Soc.*, 276:115, 1995.

[247] Naonori S. Sugiyama. Using Lagrangian perturbation theory for precision cosmology. *Astrophys. J.*, 788:63, 2014.

[248] A. Pezzotta et al. Euclid preparation. TBD. Galaxy power spectrum modelling in real space. 12 2023.

[249] Anton Chudaykin, Mikhail M. Ivanov, Oliver H. E. Philcox, and Marko Simonović. Nonlinear perturbation theory extension of the Boltzmann code CLASS. *Phys. Rev. D*, 102(6):063533, 2020.

[250] Oliver H. E. Philcox, Mikhail M. Ivanov, Giovanni Cabass, Marko Simonović, Matias Zaldarriaga, and Takahiro Nishimichi. Cosmology with the redshift-space galaxy bispectrum monopole at one-loop order. *Phys. Rev. D*, 106(4):043530, 2022.

[251] Petter Taule and Mathias Garny. The two-loop power spectrum in redshift space. *JCAP*, 11:078, 2023.

[252] Florian Beutler, Chris Blake, Matthew Colless, D. Heath Jones, Lister Staveley-Smith, Lachlan Campbell, Quentin Parker, Will Saunders, and Fred Watson. The 6dF Galaxy Survey: Baryon Acoustic Oscillations and the Local Hubble Constant. *Mon. Not. Roy. Astron. Soc.*, 416:3017–3032, 2011.

[253] Chris Blake et al. The WiggleZ Dark Energy Survey: the growth rate of cosmic structure since redshift z=0.9. *Mon. Not. Roy. Astron. Soc.*, 415:2876, 2011.

[254] Anton Chudaykin and Mikhail M. Ivanov. Cosmological constraints from the power spectrum of eBOSS quasars. *Phys. Rev. D*, 107(4):043518, 2023.

[255] Petter Taule, Marco Marinucci, Giorgia Biselli, Massimo Pietroni, and Filippo Vernizzi. Constraints on dark energy and modified gravity from the BOSS Full-Shape and DESI BAO data. 9 2024.

[256] Alejandro Aviles. Testing gravity with the full-shape galaxy power spectrum: first constraints on scale-dependent modified gravity. 9 2024.

[257] Thomas Colas, Guido D'amico, Leonardo Senatore, Pierre Zhang, and Florian Beutler. Efficient Cosmological Analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure. *JCAP*, 06:001, 2020.

[258] Théo Simon, Pierre Zhang, and Vivian Poulin. Cosmological inference from the EFTofLSS: the eBOSS QSO full-shape analysis. *JCAP*, 07:041, 2023.

[259] A. G. Adame et al. DESI 2024 V: Full-Shape Galaxy Clustering from Galaxies and Quasars. 11 2024.

[260] S. Chen et al. Analysis of DESI×DES using the Lagrangian effective theory of LSS. *Phys. Rev. D*, 110(10):103518, 2024.

[261] Chiara Moretti, Maria Tsedrik, Pedro Carrilho, and Alkistis Pourtsidou. Modified gravity and massive neutrinos: constraints from the full shape analysis of BOSS galaxies and forecasts for Stage IV surveys. *JCAP*, 12:025, 2023.

[262] Martin Crocce and Roman Scoccimarro. Nonlinear Evolution of Baryon Acoustic Oscillations. *Phys. Rev. D*, 77:023533, 2008.

[263] Atsushi Taruya and Takashi Hiramatsu. A Closure Theory for Non-linear Evolution of Cosmological Power Spectra. *Astrophys. J.*, 674:617, 2008.

[264] Francis Bernardeau, Martin Crocce, and Roman Scoccimarro. Multi-Point Propagators in Cosmological Gravitational Instability. *Phys. Rev. D*, 78:103521, 2008.

[265] Atsushi Taruya, Francis Bernardeau, Takahiro Nishimichi, and Sandrine Codis. RegPT: Direct and fast calculation of regularized cosmological power spectrum at two-loop order. *Phys. Rev. D*, 86:103528, 2012.

[266] Mikhail M. Ivanov. *Precision theoretical methods for large-scale structure of the Universe*. PhD thesis, EPFL, 1 2019.

[267] Naonori Sugiyama. Developing a Theoretical Model for the Resummation of Infrared Effects in the Post-Reconstruction Power Spectrum (youtu.be/u1-xx3_4xCg). 2 2024.

[268] Matthew Lewandowski and Leonardo Senatore. An analytic implementation of the IR-resummation for the BAO peak. *JCAP*, 03:018, 2020.

[269] Massimo Pietroni. Flowing with Time: a New Approach to Nonlinear Cosmological Perturbations. *JCAP*, 10:036, 2008.

[270] Shi-Fan Chen, Zvonimir Vlah, and Martin White. The bispectrum in Lagrangian perturbation theory. *JCAP*, 11:012, 2024.

[271] Shi-Fan Chen, Zvonimir Vlah, Emanuele Castorina, and Martin White. Redshift-Space Distortions in Lagrangian Perturbation Theory. *JCAP*, 03:100, 2021.

[272] Takahiko Matsubara. Resumming Cosmological Perturbations via the Lagrangian Picture: One-loop Results in Real Space and in Redshift Space. *Phys. Rev. D*, 77:063530, 2008.

[273] Sean M. Carroll, Stefan Leichenauer, and Jason Pollack. Consistent effective theory of long-wavelength cosmological perturbations. *Phys. Rev. D*, 90(2):023518, 2014.

[274] Massimo Pietroni, Gianpiero Mangano, Ninetta Saviano, and Matteo Viel. Coarse-Grained Cosmological Perturbation Theory. *JCAP*, 01:019, 2012.

[275] John Joseph M. Carrasco, Simon Foreman, Daniel Green, and Leonardo Senatore. The Effective Field Theory of Large Scale Structures at Two Loops. *JCAP*, 07:057, 2014.

[276] Rafael A. Porto, Leonardo Senatore, and Matias Zaldarriaga. The Lagrangian-space Effective Field Theory of Large Scale Structures. *JCAP*, 05:022, 2014.

[277] Zvonimir Vlah, Martin White, and Alejandro Aviles. A Lagrangian effective field theory. *JCAP*, 09:014, 2015.

[278] Matias Zaldarriaga and Mehrdad Mirbabayi. Lagrangian Formulation of the Eulerian-EFT. 11 2015.

[279] Tobias Baldauf, Mehrdad Mirbabayi, Marko Simonović, and Matias Zaldarriaga. LSS constraints with controlled theoretical uncertainties. 2 2016.

[280] Diogo Braganca, Yaniv Donath, Leonardo Senatore, and Henry Zheng. Peeking into the next decade in Large-Scale Structure Cosmology with its Effective Field Theory. 7 2023.

[281] Takahiro Nishimichi, Guido D'Amico, Mikhail M. Ivanov, Leonardo Senatore, Marko Simonović, Masahiro Takada, Matias Zaldarriaga, and Pierre Zhang. Blinded challenge for precision cosmology with large-scale structure: results from effective field theory for the redshift-space galaxy power spectrum. *Phys. Rev. D*, 102(12):123541, 2020.

[282] Thomas Konstandin, Rafael A. Porto, and Henrique Rubira. The effective field theory of large scale structure at three loops. *JCAP*, 11:027, 2019.

[283] Tobias Baldauf, Lorenzo Mercolli, Mehrdad Mirbabayi, and Enrico Pajer. The Bispectrum in the Effective Field Theory of Large Scale Structure. *JCAP*, 05:007, 2015.

[284] Tobias Baldauf, Mathias Garny, Petter Taule, and Theo Steele. Two-loop bispectrum of large-scale structure. *Phys. Rev. D*, 104(12):123551, 2021.

[285] Daniele Bertolini, Katelin Schutz, Mikhail P. Solon, and Kathryn M. Zurek. The Trispectrum in the Effective Field Theory of Large Scale Structure. *JCAP*, 06:052, 2016.

[286] Raul E. Angulo, Matteo Zennaro, Sergio Contreras, Giovanni Aricò, Marcos Pellejero-Ibañez, and Jens Stücker. The BACCO simulation project: exploiting the full power of large-scale structure for cosmology. *Mon. Not. Roy. Astron. Soc.*, 507(4):5869–5881, 2021.

[287] Katrin Heitmann, Earl Lawrence, Juliana Kwan, Salman Habib, and David Higdon. The Coyote Universe Extended: Precision Emulation of the Matter Power Spectrum. *Astrophys. J.*, 780:111, 2014.

[288] Takahiro Nishimichi et al. Dark Quest. I. Fast and Accurate Emulation of Halo Clustering Statistics and Its Application to Galaxy Clustering. *Astrophys. J.*, 884:29, 2019.

[289] Joseph DeRose, Risa H. Wechsler, Jeremy L. Tinker, Matthew R. Becker, Yao-Yuan Mao, Thomas McClintock, Sean McLaughlin, Eduardo Rozo, and Zhongxu Zhai. The Aemulus Project I: Numerical Simulations for Precision Cosmology. *Astrophys. J.*, 875(1):69, 2019.

[290] Benjamin D. Wibking, Andrés N. Salcedo, David H. Weinberg, Lehman H. Garrison, Douglas Ferrer, Jeremy Tinker, Daniel Eisenstein, Marc Metchnik, and Philip Pinto. Emulating galaxy clustering and galaxy–galaxy lensing into the deeply nonlinear regime: methodology, information, and forecasts. *Mon. Not. Roy. Astron. Soc.*, 484(1):989–1006, 2019.

[291] Hans Winther, Santiago Casas, Marco Baldi, Kazuya Koyama, Baojiu Li, Lucas Lombriser, and Gong-Bo Zhao. Emulators for the nonlinear matter power spectrum beyond ΛCDM. *Phys. Rev. D*, 100(12):123540, 2019.

[292] Guilherme Brando, Bartolomeo Fiorini, Kazuya Koyama, and Hans A. Winther. Enabling matter power spectrum emulation in beyond-ΛCDM cosmologies with COLA. *JCAP*, 09:051, 2022.

[293] Renate Mauland, Hans A. Winther, and Cheng-Zong Ruan. Sesame: A power spectrum emulator pipeline for beyond-ΛCDM models. *Astron. Astrophys.*, 685:A156, 2024.

[294] Drew Jamieson, Yin Li, Francisco Villaescusa-Navarro, Shirley Ho, and David N. Spergel. Field-level Emulation of Cosmic Structure Formation with Cosmology and Redshift Dependence. 8 2024.

[295] Drew Jamieson, Yin Li, Renan Alves de Oliveira, Francisco Villaescusa-Navarro, Shirley Ho, and David N. Spergel. Field-level Neural Network Emulator for Cosmological N-body Simulations. *Astrophys. J.*, 952(2):145, 2023.

[296] Mathias Garny, Dominik Laxhuber, and Roman Scoccimarro. Perturbation theory with dispersion and higher cumulants: Framework and linear theory. *Phys. Rev. D*, 107(6):063539, 2023.

[297] Mathias Garny, Dominik Laxhuber, and Roman Scoccimarro. Perturbation theory with dispersion and higher cumulants: Nonlinear regime. *Phys. Rev. D*, 107(6):063540, 2023.

[298] Mathias Garny and Roman Scoccimarro. Vlasov Perturbation Theory and the role of higher cumulants. 2 2025.

[299] Alaric Erschfeld and Stefan Floerchinger. Evolution of dark matter velocity dispersion. *JCAP*, 06:039, 2019.

[300] Alvaro Dominguez. Hydrodynamic approach to the evolution of cosmological structures. 9 2000.

[301] Patrick McDonald. How to generate a significant effective temperature for cold dark matter, from first principles. *Journal of Cosmology and Astroparticle Physics*, 2011(04):032, 2011.

[302] F. S. Kitaura, F. Sinigaglia, A. Balaguera-Antolínez, and G. Favole. The Cosmic Web from Perturbation Theory. 1 2023.

[303] Sabino Matarrese and Massimo Pietroni. Resumming Cosmic Perturbations. *JCAP*, 06:026, 2007.

[304] Patrick McDonald and Zvonimir Vlah. Large-scale structure perturbation theory without losing stream crossing. *Phys. Rev. D*, 97(2):023508, 2018.

[305] Alejandro Aviles. Dark matter dispersion tensor in perturbation theory. *Phys. Rev. D*, 93:063517, 2016.

[306] Somnath Bharadwaj. Perturbative growth of cosmological clustering. 2: The Two point correlation. *Astrophys. J.*, 460:28–50, 1996.

[307] P. Valageas. Dynamics of gravitational clustering I. building perturbative expansions. *Astron. Astrophys.*, 379:8, 2001.

[308] Svetlin V Tassev. The helmholtz hierarchy: phase space statistics of cold dark matter. *Journal of Cosmology and Astroparticle Physics*, 2011(10):022, 2011.

[309] P. Valageas. A new approach to gravitational clustering: a path-integral formalism and large-n expansions. *Astron. Astrophys.*, 421:23–40, 2004.

[310] Sebastian Pueblas and Roman Scoccimarro. Generation of Vorticity and Velocity Dispersion by Orbit Crossing. *Phys. Rev. D*, 80:043504, 2009.

[311] S. F. Shandarin V. I. Arnold and Ya. B. Zeldovich. The large scale structure of the universe i. general properties. one-and two-dimensional models. *Geophysical & Astrophysical Fluid Dynamics*, 20(1-2):111–130, 1982.

[312] Marcello Musso, Giulia Despali, and Ravi K. Sheth. The energy shear of protohaloes. 5 2024.

[313] Mariana Jaber, Marius Peper, Wojciech A. Hellwing, Miguel Angel Aragon-Calvo, and Octavio Valenzuela. Hierarchical structure of the cosmic web and galaxy properties. 4 2023.

[314] M. A. Aragon-Calvo. Hierarchical Reconstruction of the Cosmic Web, The H-Spine method. 8 2023.

[315] Noam I. Libeskind et al. Tracing the cosmic web. *Mon. Not. Roy. Astron. Soc.*, 473(1):1195–1217, 2018.

[316] Marius Cautun, Rien van de Weygaert, Bernard J. T. Jones, and Carlos S. Frenk. Evolution of the cosmic web. *Mon. Not. Roy. Astron. Soc.*, 441(4):2923–2973, 2014.

[317] Roi Kugel and Rien van de Weygaert. Cosmic Web Dynamics: Forces and Strains. 7 2024.

[318] Michael Buehlmann and Oliver Hahn. Large-Scale Velocity Dispersion and the Cosmic Web. *Mon. Not. Roy. Astron. Soc.*, 487(1):228–245, 2019.

[319] Goran Jelic-Cizmek, Francesca Lepori, Julian Adamek, and Ruth Durrer. The generation of vorticity in cosmological N-body simulations. *JCAP*, 09:006, 2018.

[320] C. Pichon and F. Bernardeau. Vorticity generation in large scale structure caustics. *Astron. Astrophys.*, 343:663, 1999.

[321] Enrico Pajer and Drian van der Woude. Divergence of Perturbation Theory in Large Scale Structures. *JCAP*, 05:039, 2018.

[322] A. J. S. Hamilton. Formulae for growth factors in expanding universes containing matter and a cosmological constant. *Mon. Not. Roy. Astron. Soc.*, 322:419, 2001.

[323] Michael Hartmeier and Mathias Garny. Minimal basis for exact time dependent kernels in cosmological perturbation theory and application to ΛCDM and w $_0$ waCDM. *JCAP*, 12:027, 2023.

[324] Ryuichi Takahashi. Third Order Density Perturbation and One-loop Power Spectrum in a Dark Energy Dominated Universe. *Prog. Theor. Phys.*, 120:549–559, 2008.

[325] Francis Bernardeau. Skewness and Kurtosis in large scale cosmic fields. *Astrophys. J.*, 433:1, 1994.

[326] Matteo Fasiello, Tomohiro Fujita, and Zvonimir Vlah. Perturbation theory of large scale structure in the ΛCDM Universe: Exact time evolution and the two-loop power spectrum. *Phys. Rev. D*, 106(12):123504, 2022.

[327] Matteo Fasiello and Zvonimir Vlah. Nonlinear fields in generalized cosmologies. *Phys. Rev. D*, 94(6):063516, 2016.

[328] Nicholas Choustikov, Zvonimir Vlah, and Anthony Challinor. Optimizing the evolution of perturbations in the ΛCDM universe. *Phys. Rev. D*, 108(2):023529, 2023.

[329] Matthew Lewandowski, Azadeh Maleknejad, and Leonardo Senatore. An effective description of dark matter and dark energy in the mildly non-linear regime. *JCAP*, 05:038, 2017.

[330] Tomohiro Fujita and Zvonimir Vlah. Perturbative description of biased tracers using consistency relations of LSS. *JCAP*, 10:059, 2020.

[331] Yaniv Donath and Leonardo Senatore. Biased Tracers in Redshift Space in the EFTofLSS with exact time dependence. *JCAP*, 10:039, 2020.

[332] Fabian Schmidt. An n-th order Lagrangian Forward Model for Large-Scale Structure. *JCAP*, 04:033, 2021.

[333] Cornelius Rampf, Sonja Ornella Schobesberger, and Oliver Hahn. Analytical growth functions for cosmic structures in a ΛCDM Universe. *Mon. Not. Roy. Astron. Soc.*, 516(2):2840–2850, 2022.

[334] Mathias Garny and Petter Taule. Two-loop power spectrum with full time- and scale-dependence and EFT corrections: impact of massive neutrinos and going beyond EdS. *JCAP*, 09:054, 2022.

[335] Greg L. Bryan and Michael L. Norman. Statistical properties of x-ray clusters: Analytic and numerical comparisons. *The Astrophysical Journal*, 495(1):80, mar 1998.

[336] Ravi K. Sheth and Antonaldo Diaferio. Peculiar velocities of galaxies and clusters. *Mon. Not. Roy. Astron. Soc.*, 322:901, 2001.

[337] J. Gasser and H. Leutwyler. Chiral Perturbation Theory to One Loop. *Annals Phys.*, 158:142, 1984.

[338] Enrico Pajer and Matias Zaldarriaga. On the Renormalization of the Effective Field Theory of Large Scale Structures. *JCAP*, 08:037, 2013.

[339] Kenneth G. Wilson. Confinement of Quarks. *Phys. Rev. D*, 10:2445–2459, 1974.

[340] C. McNeile and Christopher Michael. The Decay constant of the first excited pion from lattice QCD. *Phys. Lett. B*, 642:244–247, 2006.

[341] M. Fukugita, N. Ishizuka, H. Mino, M. Okawa, and A. Ukawa. Pion decay constant in full lattice QCD. *Phys. Lett. B*, 301:224–230, 1993.

[342] Ekaterina V. Mastropas and David G. Richards. Decay constants of the pion and its excitations on the lattice. *Phys. Rev. D*, 90(1):014511, 2014.

[343] P.J.E. Peebles. *The Large-scale Structure of the Universe*. Princeton Series in Physics. Princeton University Press, 1980.

[344] P. J. E. Peebles. Phenomenology of the Invisible Universe. *AIP Conf. Proc.*, 1241(1):175–182, 2010.

[345] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, rev - revised, 2 edition, 2008.

[346] Oliver H. E. Philcox and Mikhail M. Ivanov. BOSS DR12 full-shape cosmology: ΛCDM constraints from the large-scale galaxy power spectrum and bispectrum monopole. *Phys. Rev. D*, 105(4):043517, 2022.

[347] Mikhail M. Ivanov, Oliver H. E. Philcox, Giovanni Cabass, Takahiro Nishimichi, Marko Simonović, and Matias Zaldarriaga. Cosmology with the galaxy bispectrum multipoles: Optimal estimation and application to BOSS data. *Phys. Rev. D*, 107(8):083515, 2023.

[348] Emil Brinch Holm, Laura Herold, Théo Simon, Elisa G. M. Ferreira, Steen Hannestad, Vivian Poulin, and Thomas Tram. Bayesian and frequentist investigation of prior effects in EFT of LSS analyses of full-shape BOSS and eBOSS data. *Phys. Rev. D*, 108(12):123514, 2023.

[349] Théo Simon, Pierre Zhang, Vivian Poulin, and Tristan L. Smith. Consistency of effective field theory analyses of the BOSS power spectrum. *Phys. Rev. D*, 107(12):123530, 2023.

[350] M. Maus et al. A comparison of effective field theory models of redshift space galaxy power spectra for DESI 2024 and future surveys. 4 2024.

[351] Pedro Carrilho, Chiara Moretti, and Alkistis Pourtsidou. Cosmology with the EFTofLSS and BOSS: dark energy constraints and a note on priors. *JCAP*, 01:028, 2023.

[352] Jamie Donald-McCann, Rafaela Gsponer, Ruiyang Zhao, Kazuya Koyama, and Florian Beutler. Analysis of unified galaxy power spectrum multipole measurements. *Mon. Not. Roy. Astron. Soc.*, 526(3):3461–3481, 2023.

[353] James E. Gunn and J. Richard Gott, III. On the Infall of Matter into Clusters of Galaxies and Some Effects on Their Evolution. *Astrophys. J.*, 176:1–19, 1972.

[354] William H. Press and Paul Schechter. Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation. *Astrophys. J.*, 187:425–438, 1974.

[355] J. R. Bond, S. Cole, G. Efstathiou, and Nick Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *Astrophys. J.*, 379:440, 1991.

[356] Mikhail M. Ivanov, Alexander A. Kaurov, and Sergey Sibiryakov. Non-perturbative probability distribution function for cosmological counts in cells. *JCAP*, 03:009, 2019.

[357] Andreas A. Berlind and David H. Weinberg. The Halo occupation distribution: Towards an empirical determination of the relation between galaxies and mass. *Astrophys. J.*, 575:587–616, 2002.

[358] Leonardo Senatore and Matias Zaldarriaga. Redshift Space Distortions in the Effective Field Theory of Large Scale Structures. 9 2014.

[359] Mikhail M. Ivanov. *Effective Field Theory for Large-Scale Structure*. 2023.

[360] Simon Foreman, Hideki Perrier, and Leonardo Senatore. Precision Comparison of the Power Spectrum in the EFTofLSS with Simulations. *JCAP*, 05:027, 2016.

[361] Tobias Baldauf, Lorenzo Mercolli, and Matias Zaldarriaga. Effective field theory of large scale structure at two loops: The apparent scale dependence of the speed of sound. *Phys. Rev. D*, 92(12):123007, 2015.

[362] John Joseph M. Carrasco, Simon Foreman, Daniel Green, and Leonardo Senatore. The 2-loop matter power spectrum and the IR-safe integrand. *JCAP*, 07:056, 2014.

[363] Mandar Karandikar, Cristiano Porciani, and Oliver Hahn. Testing the assumptions of the Effective Field Theory of Large-Scale Structure. *JCAP*, 01:051, 2024.

[364] Tobias Baldauf, Emmanuel Schaan, and Matias Zaldarriaga. On the reach of perturbative descriptions for dark matter displacement fields. *JCAP*, 03:017, 2016.

[365] Matthew McQuinn and Martin White. Cosmological perturbation theory in 1+1 dimensions. *JCAP*, 01:043, 2016.

[366] Raul E. Angulo, Simon Foreman, Marcel Schmittfull, and Leonardo Senatore. The One-Loop Matter Bispectrum in the Effective Field Theory of Large Scale Structures. *JCAP*, 10:039, 2015.

[367] Simon Foreman and Leonardo Senatore. The EFT of Large Scale Structures at All Redshifts: Analytical Predictions for Lensing. *JCAP*, 04:033, 2016.

[368] Samuel W. Skillman, Michael S. Warren, Matthew J. Turk, Risa H. Wechsler, Daniel E. Holz, and P. M. Sutter. Dark Sky Simulations: Early Data Release. 7 2014.

[369] César Hernández-Aguayo et al. The MillenniumTNG Project: high-precision predictions for matter clustering and halo statistics. *Mon. Not. Roy. Astron. Soc.*, 524(2):2556–2578, 2023.

[370] Ryuichi Takahashi, Masanori Sato, Takahiro Nishimichi, Atsushi Taruya, and Masamune Oguri. Revising the Halofit Model for the Nonlinear Matter Power Spectrum. *Astrophys. J.*, 761:152, 2012.

[371] Ethan R. Siegel and James N. Fry. Effects of inhomogeneities on cosmic expansion. *Astrophys. J. Lett.*, 628:L1–L4, 2005.

[372] Yi-Kuan Chiang, Ryu Makiya, Eiichiro Komatsu, and Brice Ménard. The thermal and gravitational energy densities in the large-scale structure of the Universe. *Astrophys. J.*, 910(1):32, 2021.

[373] A. Jenkins, C. S. Frenk, F. R. Pearce, P. A. Thomas, J. M. Colberg, Simon D. M. White, H. M. P. Couchman, J. A. Peacock, G. Efstathiou, and A. H. Nelson. Evolution of structure in cold dark matter universes. *Astrophys. J.*, 499:20, 1998.

[374] A. Kehagias and A. Riotto. Symmetries and Consistency Relations in the Large Scale Structure of the Universe. *Nucl. Phys. B*, 873:514–529, 2013.

[375] Guido D'Amico, Marco Marinucci, Massimo Pietroni, and Filippo Vernizzi. The large scale structure bootstrap: perturbation theory and bias expansion from symmetries. *JCAP*, 10:069, 2021.

[376] Marco Marinucci, Kevin Pardede, and Massimo Pietroni. Bootstrapping Lagrangian Perturbation Theory for the Large Scale Structure. 5 2024.

[377] Arhum Ansari, Arka Banerjee, Sachin Jain, and Shaunak Padhyegurjar. Time Non-locality in Dark Matter and LSS. 6 2024.

[378] Yaniv Donath, Matthew Lewandowski, and Leonardo Senatore. Direct signatures of the formation time of galaxies. *Phys. Rev. D*, 109(12):123510, 2024.

[379] Christian Wagner, Fabian Schmidt, Chi-Ting Chiang, and Eiichiro Komatsu. Separate Universe Simulations. *Mon. Not. Roy. Astron. Soc.*, 448(1):L11–L15, 2015.

[380] Tobias Baldauf, Uroš Seljak, Leonardo Senatore, and Matias Zaldarriaga. Linear response to long wavelength fluctuations using curvature simulations. *JCAP*, 09:007, 2016.

[381] Yin Li, Wayne Hu, and Masahiro Takada. Super-Sample Covariance in Simulations. *Phys. Rev. D*, 89(8):083519, 2014.

[382] Alex Kehagias, Hideki Perrier, and Antonio Riotto. Equal-time Consistency Relations in the Large-Scale Structure of the Universe. *Mod. Phys. Lett. A*, 29:1450152, 2014.

[383] Ido Ben-Dayan, Thomas Konstandin, Rafael A. Porto, and Laura Sagunski. On Soft Limits of Large-Scale Structure Correlation Functions. *JCAP*, 02:026, 2015.

[384] Blake D Sherwin and Matias Zaldarriaga. Shift of the baryon acoustic oscillation scale: A simple physical picture. *Physical Review D—Particles, Fields, Gravitation, and Cosmology*, 85(10):103523, 2012.

[385] Christian Wagner, Fabian Schmidt, Chi-Ting Chiang, and Eiichiro Komatsu. The angle-averaged squeezed limit of nonlinear matter N-point functions. *JCAP*, 08:042, 2015.

[386] Peter S. Behroozi, Risa H. Wechsler, Hao-Yi Wu, Michael T. Busha, Anatoly A. Klypin, and Joel R. Primack. Gravitationally consistent halo catalogs and merger trees for precision cosmology. *The Astrophysical Journal*, 763(1):18, December 2012.

[387] Andrija Kostić, Nhat-Minh Nguyen, Fabian Schmidt, and Martin Reinecke. Consistency tests of field level inference with the EFT likelihood. *JCAP*, 07:063, 2023.

[388] Beatriz Tucci and Fabian Schmidt. EFTofLSS meets simulation-based inference: σ_8 from biased tracers. *JCAP*, 05:063, 2024.

[389] Marcel Schmittfull, Marko Simonović, Valentin Assassi, and Matias Zaldarriaga. Modeling Biased Tracers at the Field Level. *Phys. Rev. D*, 100(4):043514, 2019.

[390] Nhat-Minh Nguyen, Fabian Schmidt, Beatriz Tucci, Martin Reinecke, and Andrija Kostić. How Much Information Can Be Extracted from Galaxy Clustering at the Field Level? *Phys. Rev. Lett.*, 133(22):221006, 2024.

[391] Shogo Masaki, Takahiro Nishimichi, and Masahiro Takada. Anisotropic separate universe simulations. *Mon. Not. Roy. Astron. Soc.*, 496(1):483–496, 2020.

[392] Jens Stücker, Andreas. S. Schmidt, Simon D. M. White, Fabian Schmidt, and Oliver Hahn. Measuring the tidal response of structure formation: anisotropic separate universe simulations using treepm. *Mon. Not. Roy. Astron. Soc.*, 503(1):1473–1489, 2021.

[393] Yin Li, Marcel Schmittfull, and Uroš Seljak. Galaxy power-spectrum responses and redshift-space super-sample effect. *JCAP*, 02:022, 2018.

[394] Kazuyuki Akitsu, Yin Li, and Teppei Okumura. Cosmological simulation in tides: power spectra, halo shape responses, and shape assembly bias. *JCAP*, 04:041, 2021.

[395] Mathias Garny, Thomas Konstandin, Rafael A. Porto, and Laura Sagunski. On the Soft Limit of the Large Scale Structure Power Spectrum: UV Dependence. *JCAP*, 11:032, 2015.

[396] Mikhail M. Ivanov, Andrej Obuljen, Carolina Cuesta-Lazaro, and Michael W. Toomey. Full-shape analysis with simulation-based priors: cosmological parameters and the structure growth anomaly. 9 2024.

[397] Hanyu Zhang, Marco Bonici, Guido D'Amico, Simone Paradiso, and Will J. Percival. HOD-informed prior for EFT-based full-shape analyses of LSS. 9 2024.

[398] Mikhail M. Ivanov, Carolina Cuesta-Lazaro, Siddharth Mishra-Sharma, Andrej Obuljen, and Michael W. Toomey. Full-shape analysis with simulation-based priors: Constraints on single field inflation from BOSS. *Phys. Rev. D*, 110(6):063538, 2024.

[399] Job Feldbrugge. Phase-Space Delaunay Tessellation Field Estimator. 2 2024.

[400] Saee Dhawalikar and Aseem Paranjape. Optimization of cosmic filament finders and unbiased recovery of filament phase space profiles using mock filaments. 2 2024.

[401] Alejandro Palomino, Felipe Leonardo Gómez-Cortés, Xiao-Dong Li, and Jaime E. Forero-Romero. Cosmic web classification through stochastic topological ranking, 2024.

[402] Job Feldbrugge and Rien van de Weygaert. What makes a cosmic filament? The dynamical origin and identity of filaments I. fundamentals in 2D. 5 2024.

[403] Suman Sarkar and Biswajit Pandey. Unveiling galaxy pair alignment in cosmic filaments: A 3D exploration using EAGLE simulation. 7 2024.

[404] Mohadese Khoshtinat, Hossein Hatamnia, and Shant Baghram. One-point Statistics in various cosmic environments in the presence of massive neutrinos. 7 2024.

[405] Samo Ilc, Dunja Fabjan, Elena Rasia, Stefano Borgani, and Klaus Dolag. Properties of the diffuse gas component in filaments detected in the Dianoga cosmological simulations. 7 2024.

[406] Henrique Rubira and Fabian Schmidt. Galaxy bias renormalization group. *JCAP*, 01:031, 2024.

[407] Henrique Rubira and Fabian Schmidt. The Renormalization Group for Large-Scale Structure: Origin of Galaxy Stochasticity. 4 2024.

[408] Charalampos Nikolis, Henrique Rubira, and Fabian Schmidt. The Renormalization Group for Large-Scale Structure: Primordial non-Gaussianities. 5 2024.

[409] Valentin Assassi, Daniel Baumann, Enrico Pajer, Yvette Welling, and Drian van der Woude. Effective theory of large-scale structure with primordial non-Gaussianity. *JCAP*, 11:024, 2015.

[410] Jeremy L. Tinker, Andrey V. Kravtsov, Anatoly Klypin, Kevork Abazajian, Michael S. Warren, Gustavo Yepes, Stefan Gottlober, and Daniel E. Holz. Toward a halo mass function for precision cosmology: The Limits of universality. *Astrophys. J.*, 688:709–728, 2008.

[411] Gautham Adamane Pallathadka et al. The Nineteenth Data Release of the Sloan Digital Sky Survey. 7 2025.

[412] M. Abdul Karim et al. Data Release 1 of the Dark Energy Spectroscopic Instrument. 3 2025.

[413] Sigurd Naess et al. The Atacama Cosmology Telescope: DR6 Maps. 3 2025.

[414] M. Archipley et al. Millimeter-wave observations of Euclid Deep Field South using the South Pole Telescope: A data release of temperature maps and catalogs. 5 2025.

[415] M. Tristram et al. Cosmological parameters derived from the final Planck data release (PR4). *Astron. Astrophys.*, 682:A37, 2024.

[416] H Aussel, I Tereno, M Schirmer, G Alguero, B Altieri, E Balbinot, T de Boer, P Casenove, P Corcho-Caballero, H Furusawa, et al. Euclid quick data release (q1)–data release overview. *arXiv preprint arXiv:2503.15302*, 2025.

[417] Angus H. Wright et al. The fifth data release of the Kilo Degree Survey: Multi-epoch optical/NIR imaging covering wide and legacy-calibration fields. *Astron. Astrophys.*, 686:A170, 2024.

[418] K. Bechtol et al. Dark Energy Survey Year 6 Results: Photometric Data Set for Cosmology. 1 2025.

[419] Hiroaki Aihara et al. Third data release of the Hyper Suprime-Cam Subaru Strategic Program. *Publ. Astron. Soc. Jap.*, 74(2):247–272–272, 2022.

[420] Brendan P Crill, Michael Werner, Rachel Akeson, Matthew Ashby, Lindsey Bleem, James J Bock, Sean Bryan, Jill Burnham, Joyce Byunh, Tzu-Ching Chang, et al. Spherex: Nasa's near-infrared spectrophotometric all-sky survey. In *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*, volume 11443, pages 61–77. SPIE, 2020.

[421] Olivier Doré et al. WFIRST: The Essential Cosmology Space Observatory for the Coming Decade. 4 2019.

[422] Yan Gong et al. Introduction to the China Space Station Telescope (CSST). 7 2025.

[423] A. Weltman et al. Fundamental physics with the Square Kilometre Array. *Publ. Astron. Soc. Austral.*, 37:e002, 2020.

[424] Anton Chudaykin, Mikhail M. Ivanov, and Sergey Sibiryakov. Renormalizing one-point probability distribution function for cosmological counts in cells. *JCAP*, 08:079, 2023.