Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **507,** 2614–2631 (2021)
Advance Access publication 2021 August 6

https://doi.org/10.1093/mnras/stab2260

An optimal non-linear method for simulating relic neutrinos

Willem Elbers,[1]⋆ Carlos S. Frenk,[1] Adrian Jenkins,[1] Baojiu Li[1] and Silvia Pascoli[2]

[1] *Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2] *Institute for Particle Physics Phenomenology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

Accepted 2021 July 24. Received 2021 June 15; in original form 2020 October 14

ABSTRACT

Cosmology places the strongest current limits on the sum of neutrino masses. Future observations will further improve the sensitivity and this will require accurate cosmological simulations to quantify possible systematic uncertainties and to make predictions for non-linear scales, where much information resides. However, shot noise arising from neutrino thermal motions limits the accuracy of simulations. In this paper, we introduce a new method for simulating large-scale structure formation with neutrinos that accurately resolves the neutrinos down to small scales and significantly reduces the shot noise. The method works by tracking perturbations to the neutrino phase-space distribution with particles and reduces shot noise in the power spectrum by a factor of $\mathcal{O}\left(10^2\right)$ at $z = 0$ for minimal neutrino masses and significantly more at higher redshifts, without neglecting the back-reaction caused by neutrino clustering. We prove that the method is part of a family of optimal methods that minimize shot noise subject to a maximum deviation from the non-linear solution. Compared to other methods, we find per mille level agreement in the matter power spectrum and per cent level agreement in the large-scale neutrino bias, but large differences in the neutrino component on small scales. A basic version of the method can easily be implemented in existing *N*-body codes and allows neutrino simulations with significantly reduced particle load. Further gains are possible by constructing background models based on perturbation theory. A major advantage of this technique is that it works well for all masses, enabling a consistent exploration of the full neutrino parameter space.

Key words: neutrinos – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

The discovery of neutrino masses (Fukuda et al. 1998; Ahmad et al. 2002; Eguchi et al. 2003) calls for extensions of the standard model of particle physics and provides the only known form of dark matter. Measuring the masses is crucial for understanding their origin and for constraining cosmological parameters. While the neutrino mass squared differences are known to a few per cent, the absolute masses are unknown and there remain two possible mass orderings: normal and inverted. A rich experimental programme is aimed at determining the mass ordering, measuring the mass scale set by the lightest neutrino, and completing the overall picture of neutrino properties. Cosmology plays a vital role in this programme due to its ability to provide an independent and complementary constraint on the sum of neutrino masses, $\sum m_\nu$ (Bond, Efstathiou & Silk 1980; Hu, Eisenstein & Tegmark 1998) with a potential sensitivity below 0.02 eV (Font-Ribera et al. 2014; Chudaykin & Ivanov 2019; Sprenger et al. 2019).

Ongoing and planned neutrino experiments will establish the mass ordering with a discovery expected by the end of the decade. Although oscillation data have shown persistent hints of normal ordering, this preference has decreased to 1.6σ over the past year (Esteban et al. 2020). The mass ordering can be established by exploiting matter effects in long baseline neutrino oscillation experiments, as in the Deep Underground Neutrino Experiment (Acciarri et al. 2015), and in the Earth for atmospheric neutrinos, as in the Oscillation Research with Cosmics in the Abyss (Adrian-Martinez et al. 2016) and Hyper-Kamiokande (Abe et al. 2011) experiments, as well as vacuum oscillations in medium baseline reactor neutrino experiments, specifically the Jiangmen Underground Neutrino Observatory (An et al. 2016). Each approach is challenging, so information from multiple sources is essential. Single β-decay is the experimental strategy of choice for direct mass searches and provides a model-independent determination of neutrino masses, in particular the effective electron–neutrino mass. The Karlsruhe Tritium Neutrino (KATRIN) experiment is ongoing and has put a bound of red $m_\beta < 0.8\,\mathrm{eV}$, assuming quasi-degenerate neutrino masses, with the aim of reaching $m_\beta < 0.2$ eV in the near future (Aker et al. 2021). Project 8 will have the potential to set a limit of $m_\beta < 0.04$ eV (Esfahani et al. 2017). Neutrinoless double β-decay can also provide information on neutrino masses (Bilenky, Pascoli & Petcov 2001; Nunokawa, Teves & Zukanovich Funchal 2002; Pascoli & Petcov 2002), albeit entangled with the value of the Majorana CP-violating phases and affected by uncertainty in the nuclear matrix elements (Vergados, Ejiri & Šimkovic 2016). For a recent review, see e.g. Giuliani et al. (2019).

The complementarity between these different strategies is of great interest. A cosmological measurement of $\sum m_\nu$ would provide a target for direct mass searches (Drexlin et al. 2013; Mertens 2016).

⋆ E-mail: willem.h.elbers@durham.ac.uk

An incompatibility between the two would indicate a non-standard cosmological evolution or new neutrino properties. A cosmological bound of $\sum m_\nu < 100$ meV would suggest a normal mass ordering, which should be confronted with evidence from neutrino experiments. Finally, there is a strong synergy with neutrinoless double β-decay. Knowing the mass ordering and the sum of neutrino masses would narrow down the range of values for the effective Majorana mass parameter, providing a clear target for future experiments.

Measuring the mass scale, and potentially ruling out the inverted mass ordering, is therefore a major target of near-term cosmological surveys, including the Dark Energy Spectroscopic Instrument (Levi et al. 2013), Euclid (Laureijs et al. 2011), and Vera Rubin Observatory (Ivezic et al. 2009). In order to analyse these surveys and to extract a mass measurement, there has been a substantial effort to model precisely the effects of massive neutrinos on structure formation. From the analytical side, a swathe of new techniques such as time RG perturbation theory (Upadhye 2019) and effective field theories (Senatore & Zaldarriaga 2017; Colas et al. 2020) promise to push the validity of perturbation theory into the quasi-linear régime. In the non-linear régime, *N*-body simulations offer the most accurate picture of structure formation. Yet incorporating neutrinos into *N*-body simulations has proved to be a challenge and some doubts remain about the validity of neutrino simulations on small scales.

The main obstacle to simulating neutrinos is that, in contrast to cold dark matter and baryons, neutrinos have a significant velocity dispersion. This effectively turns the 3D problem of structure formation, for which *N*-body simulations are well suited, into a 6D phase-space problem. If no provisions are made, a far greater number of simulation particles are needed to sample properly the phase-space manifold. A further complication arises from the fact that neutrinos are relativistic at high redshifts, such that simulations need to handle both the régime where neutrinos are best described as radiation and the régime where neutrinos are better described as massive particles.

The first 3D cosmological neutrino simulations were carried out by Klypin & Shandarin (1983) and Frenk, White & Davis (1983), when neutrinos were thought to be much more massive and the velocity dispersion not as problematic. Modern simulations with sub-electronvolt neutrinos were pioneered by Brandbyge et al. (2008) and Viel, Haehnelt & Springel (2010). Neutrinos are most commonly included in simulations as particles whose initial velocity is the sum of a peculiar gravitational component and a random component sampled from a Fermi–Dirac distribution (Ma & Bertschinger 1994a, b; Brandbyge et al. 2008; Viel et al. 2010; Bird, Viel & Haehnelt 2012; Villaescusa-Navarro et al. 2014, 2020; Castorina et al. 2015; Inman et al. 2015; Villaescusa-Navarro, Bull & Viel 2015; Adamek, Durrer & Kunz 2017; Emberson et al. 2017; Banerjee et al. 2018). The main difficulty with particle simulations is shot noise caused by the velocity dispersion. This problem is more severe for the smallest neutrino masses, which are observationally most relevant. Because neutrinos are a subdominant component, the error in the total matter distribution is relatively small. However, shot noise obscures the small-scale behaviour of the neutrinos and is clearly undesirable if one is interested in the neutrino component and its effect on structure formation.

To overcome the problems with particle simulations, grid simulations evolve the neutrino distribution using a system of fluid equations, which requires a scheme to close the moment hierarchy at some low order (Brandbyge & Hannestad 2009; Viel et al. 2010; Hannestad, Haugbølle & Schultz 2012; Archidiacono & Hannestad 2016; Banerjee & Dalal 2016; Dakin et al. 2019; Tram et al. 2019; Inman & Yu 2020; Chen, Upadhye & Wong 2021a), or as a linear response to the non-relativistic matter density (Ali-Haïmoud & Bird 2012; Liu et al. 2018; McCarthy et al. 2018; Chen, Upadhye & Wong 2021b). Even more efficiently, but in the same spirit of treating neutrinos perturbatively, the total effect of neutrinos has been included as a post-processing step in the form of a gauge transformation (Partmann et al. 2020). While these approaches do not suffer from shot noise, they are not able to capture the full non-linear evolution of the neutrinos at late times. This problem becomes more severe for more massive neutrinos, but is present even for minimal neutrino masses. A number of hybrid simulations have therefore combined grid and particle methods (Brandbyge & Hannestad 2010; Banerjee & Dalal 2016; Bird et al. 2018), typically transitioning from a fluid method to a particle method at some redshift when the neutrinos become non-linear. Another interesting alternative is to integrate the Poisson–Boltzmann equations directly on the grid (Yoshikawa et al. 2020).

The method proposed in this paper can be considered as a type of hybrid method that integrates neutrino particles but only uses the information contained in the particles to the extent that is necessary. This is accomplished by dynamically transitioning from a smooth background model to a non-linear model at the individual particle level. It relies on the noiseless (but approximate) background model as much as possible, thereby producing the smallest amount of shot noise possible while solving the full non-linear system. The main idea is to decompose the phase-space distribution function $f(x, p, t)$ into a background model $\bar{f}(x, p, t)$ that can be solved without noise and a perturbation that is carried by the simulation particles:

$$f(x, p, t) = \bar{f}(x, p, t) + \delta f(x, p, t).$$

The choice of background model is arbitrary, but the method performs best whenever $\bar{f}(x, p, t)$ is strongly correlated with $f(x, p, t)$, in a way that will be made precise below. If the choice of background model is poor, the method performs no worse than an ordinary *N*-body simulation, except for the small amount of overhead associated with evaluating $\bar{f}(x, p, t)$. Note that the background model is just an approximation of f and can itself be a perturbed Fermi–Dirac distribution.

This type of method has a long history in other fields and is variably known as the method of ‘perturbation particles’ or more commonly as the ‘δf method’, which is the name we shall adopt. Merritt (1987) and Leeuwin, Combes & Binney (1993) discussed the method of perturbation particles in stellar dynamics. Around the same time, the δf method arose in plasma physics (Tajima & Perkins 1983; Dimits & Lee 1993; Parker & Lee 1993; Aydemir 1994). While the method of perturbation particles is not widely known today in astronomy, the δf method is standard fare in plasma physics. A major difficulty in astronomical applications is the absence of a background model that captures enough of the dynamics to be useful. In contrast, plasma physicists are often interested in turbulent phenomena arising in an otherwise stable system, with a natural candidate for a background model $\bar{f}$ at hand. Our work is motivated by the fact that there is also a natural background model for cosmic neutrinos, namely the phase-space density predicted by perturbation theory. There is a major synergy between δf *N*-body simulations proposed here and work on improved perturbation theory methods. A better background model means a smaller dependence on the particles and therefore further reduced shot noise. We will show, however, that even the zeroth-order approximation, which is just a homogeneous redshifted Fermi–Dirac distribution, provides a significant improvement over ordinary *N*-body methods.

The remainder of the paper is structured as follows. In Section 2, we derive the δf method and describe its use as a variance reduction method for N-body simulations. We also show that the method is part of a family of optimal hybrid methods. In Section 3, we illustrate the method with a 1D test problem. In Section 4, we discuss how the method can be embedded in relativistic simulations. Our suite of simulations is then described in Section 5. The method is compared with commonly used alternatives in Section 6. We consider higher order background models based on perturbation theory in Section 7. Finally, we conclude in Section 8.

2 DERIVATION

The phase-space evolution of self-gravitating collisionless particles is described by the Poisson–Boltzmann equations, which in the single-fluid case read

$$Lf \equiv \left[\frac{\partial}{\partial t} + p\cdot\nabla - \nabla\Phi\cdot\nabla_p\right] f = 0,$$

$$\nabla^2\Phi = 4\pi G\rho = 4\pi G\int \mathrm{d}^3p\,\sqrt{m^2+p^2}\,f(x,p,t).$$

Here, Φ is the gravitational potential, ρ is the energy density, and f is the phase-space density. In general, the Liouville operator, L, acts on each fluid separately and the potential should be summed over all fluid components. In relativistic perturbation theory, this system can be written as a hierarchy of moment equations for the neutrinos, which is solved to first order with Boltzmann codes such as CLASS (Lesgourgues 2011) or CAMB (Lewis & Challinor 2011). To extend our predictions to the non-linear régime, we can use N-body codes, which solve the Poisson–Boltzmann system by the method of characteristics. Characteristic curves satisfy

$$\frac{\mathrm{d}x}{\mathrm{d}t} = p \qquad \text{and} \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\nabla\Phi.$$

By construction, one finds that $\mathrm{d}f/\mathrm{d}t = Lf = 0$ along these curves. To infer statistics of the phase-space distribution, we simulate N of these trajectories using marker particles. We can freely choose the phase-space distribution, g, of our simulation particles at the initial time. For instance, assuming an initially homogeneous spatial distribution and momenta from the Fermi–Dirac distribution, we would have $g \propto (\exp\{p/T_\nu\} + 1)^{-1}$. Typically, one chooses $g(x, p, t_0) = f(x, p, t_0)$. Since $Lg = 0$, this equality then holds for all $t \geq t_0$. In general, a phase-space statistic is given by

$$\begin{aligned} A(x,t) = \langle A\rangle_p &= \int \mathrm{d}^3p\; f(x,p,t)A(x,p,t) \\ &\cong \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i,p_i,t)}{g(x_i,p_i,t)}A(x_i,p_i,t). \end{aligned} \tag{1}$$

Following the usual choice of setting $g(x, p, t_0) = f(x, p, t_0)$, the sum reduces to a simple average over marker particles. The error in our estimate of A is then $\sigma_A/\sqrt{N}$. Hence, if the distribution, $f(x, p, t)$, has a large intrinsic scatter, we need a large N to beat down the noise. Alternatively, we might construct an estimator with a smaller error. Let us therefore write the phase-space distribution function, f, as a background model, $\bar{f}$, together with some perturbation, δf:

$$f(x,p,t) = \bar{f}(x,p,t) + \delta f(x,p,t).$$

We can reduce the error by only using the particles to estimate the perturbed distribution, δf. We replace (1) with

$$\begin{aligned} A(x,t) &= \int \mathrm{d}^3p\,\left[\bar{f}(x,p,t)+\delta f(x,p,t)\right]A(x,p,t) \\ &\cong \bar{A}(x,t) + \frac{1}{N}\sum_{i=1}^{N}\frac{\delta f(x_i,p_i,t)}{g(x_i,p_i,t)}A(x_i,p_i,t). \end{aligned}$$

This is useful if

$$\bar{A}(x,t) = \int \mathrm{d}^3p\;\bar{f}(x,p,t)A(x,p,t)$$

can be computed efficiently and if f and $\bar{f}$ are strongly correlated, so that the second term is small. The simplest choice of background model is a homogeneous Fermi–Dirac distribution

$$\bar{f}(x,p,t) = \frac{g_s}{(2\pi)^3}\frac{1}{e^{ap/(k_bT_\nu)}+1}, \tag{2}$$

with g_s internal degrees of freedom. Here, $a = a(t)$ is the scale factor, $T_\nu = 1.95$ K is the present-day neutrino temperature, and ap is the present-day momentum. Since the noise reduction scales with the correlation between $\bar{f}$ and f, we can achieve further gains by adding more information to the background model. The obvious next step is to use perturbation theory to improve on (2). This option is considered in Section 7.

2.1 Implementation

In Appendix E, we outline the practical steps needed to implement the method in cosmological N-body simulations. In essence, whenever we sum over neutrino particles, such as when calculating the gravitational force on a test particle, we replace the particle mass with a statistically weighted mass:

$$m \to mw_i = m\left[\frac{\delta f(x_i,p_i,t)}{g(x_i,p_i,t)}\right] = m\left[\frac{f(x_i,p_i,t)-\bar{f}(x_i,p_i,t)}{g(x_i,p_i,t)}\right].$$

The weights are computed by comparing the true phase-space density with the background model. We know the background model density, because we can evaluate (2) at any time. We also know the true density for each particle, because $Lf = Lg = 0$ along characteristic curves. It is therefore sufficient to record the two numbers f and g at the initial sampled location of each particle in phase space. We note that any sampling distribution g is valid provided that $g \neq 0$ almost everywhere $f \neq 0$. We will continue to use the common choice, $g = f$, where f is the Fermi–Dirac distribution. In general, the optimal choice of g will depend on the phase-space statistic of interest. Choosing a distribution, g, that oversamples slower particles can provide an additional reduction in shot noise.

Given the homogeneous Fermi–Dirac background model (2), the neutrino density becomes

$$\rho_\nu(x,t) = \bar{\rho}_\nu(t) + \textstyle\sum_{i=1}^{N} mw_i\,\delta^{(3)}(x-x_i).$$

Cosmological N-body simulations only compute the perturbed potential, since the background density $\bar{\rho}$ is accounted for in the background equations. The only change affecting the force calculation is therefore the weighting of the particles.

The mean squared weight, $I = \frac{1}{2}\langle w^2\rangle$, is a convenient statistic to quantify the importance of including the neutrino particles. We show the evolution of I for a $\sum m_\nu = 100$ meV simulation with the homogeneous background model (2) in Fig. 1. At early times, particles deviate very little from their initial trajectory and the weights are negligible. We find that $I = 4\times10^{-7}$ at $z = 20$, $I = 3\times10^{-6}$ at $z =$

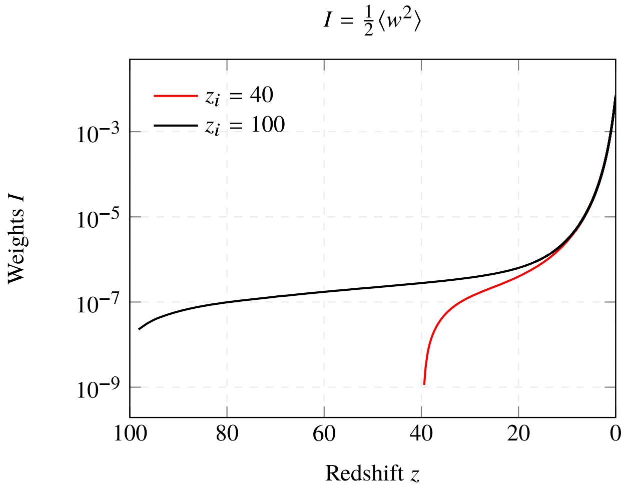

Figure 1. Evolution of particle weights for a $\sum m_\nu = 100$ meV cosmology, starting at different redshifts z_i. The mean squared particle weight $\langle w^2\rangle$ represents the effective reduction in shot noise.

10, and $I = 2\times10^{-5}$ at $z = 5$. This early reduction is important as shot noise at high redshifts inhibits the growth of physical structure and can seed additional fluctuations that grow by gravitational instability. At late times, when non-linear effects become important, the weights increase to $I = 2\times10^{-4}$ at $z = 2$, $I = 1\times10^{-3}$ at $z = 1$, and $I = 6.7\times10^{-3}$ at $z = 0$, independently of the starting redshift of the simulation. This translates to a reduction in shot noise, $\sigma = 2VI/N$, or an effective increase in particle number at $z = 0$ by a factor $(2I)^{-1} = 75$. Finally, we note that one can save computational resources by integrating only a fraction of the neutrino particles as long as I remains small. We do not consider this possibility here.

2.2 Variance reduction

The δf method is an application of the much more general control variates method (Aydemir 1994; Ross 2012). This is a variance reduction technique commonly used in Monte Carlo simulations. See Chartier et al. (2021) for another recent application in cosmology. We briefly review the method here. Let A be a random variable with an unknown expectation $\mathrm{E}[A] = \mathcal{A}$. Given independent random samples A_i, the standard estimator is given by

$$\hat{\mathcal{A}} = \frac{1}{N}\sum_{i=1}^{N} A_i .$$

The error in $\hat{\mathcal{A}}$ is

$$\sigma^2_{\hat{\mathcal{A}}} = \mathrm{E}\left[(A-\hat{\mathcal{A}})^2\right] = \frac{\sigma_A^2}{N}.$$

Let B be another random variable for which the expected value $\mathrm{E}[B] = \mathcal{B}$ is known. By adding and subtracting, we can construct a control variate estimator for A:

$$\hat{\mathcal{A}}_{\rm cv} = \frac{1}{N}\sum_{i=1}^{N}\left[A_i - \alpha B_i\right] + \alpha\mathcal{B},$$

for any constant α. Like $\hat{\mathcal{A}}$, this is an unbiased estimator of $\mathrm{E}[A]$. However, the error in $\hat{\mathcal{A}}_{\rm cv}$ is given by

$$\sigma^2_{\hat{\mathcal{A}}_{\rm cv}} = \tfrac{1}{N}\left(\sigma_A^2 + \alpha^2\sigma_B^2 - 2\alpha\,\mathrm{cov}(A,B)\right).$$

Therefore, the error can be reduced if A and B are correlated. Differentiating, we see that the optimal value of α is given by

$$\alpha^* = \frac{\mathrm{cov}(A,B)}{\sigma_B^2}. \tag{3}$$

For the Fermi–Dirac model considered above, α^* is very close to unity and we simply set $\alpha = 1$ at all times. In general, the value of α^* could be estimated at runtime. This is useful if we add more information about the unknown variable and extend the method to a linear combination of multiple control variates (see Section 7). Furthermore, the method can still be useful when a control variate is not exactly known but can be estimated more efficiently than A.

2.3 Optimality

Let us consider how the δf method compares to other methods. To allow for the broadest possible comparison, we will write down an arbitrary hybrid method that involves some background model, $\bar{f}(x,p,t)$, such as a fluid description or linear response, and a discrete sampling of the distribution with arbitrary particle weights, $w_i(t)$:

$$\begin{aligned} f_{\rm hyb}(x,p,t) &= \alpha(t)\bar{f}(x,p,t) \\ &\quad + \sum_i w_i(t)\delta^{(3)}(x-x_i)\delta^{(3)}(p-p_i), \end{aligned} \tag{4}$$

where $\alpha(t)$ is a weight function for the background. This parametrization captures virtually all existing methods. The ordinary N-body particle method corresponds to $(\alpha, w_i) = (0, 1)$ at all times. Pure grid-based methods have $(\alpha, w_i) = (1, 0)$. Existing hybrid methods switch over from a grid method to a particle method after some time t_s, which corresponds to $(\alpha, w_i) = (1-q, q)$ with $q(t) = I[t\geq t_s]$ a step function. For simplicity, we consider only the case where all particles are switched on at the same time, but the argument extends readily to the more practical case where only some particles are switched on. Given a choice of weight function, $\alpha(t)$, for the background, what choice of particle weights is optimal?

Let $f(x, p, t)$ be the non-linear distribution and $g(x, p, t)$ the sampling distribution of the markers. In the continuous limit, the expected error in the number density is given by

$$\begin{aligned} \langle\epsilon\rangle &= \int \mathrm{d}^3p\; f_{\rm hyb}(x,p,t) - \int \mathrm{d}^3p\; f(x,p,t) \\ &= \int \mathrm{d}^3p \Big(w(x,p,t)g(x,p,t) + \alpha(t)\bar{f}(x,p,t) - f(x,p,t)\Big). \end{aligned}$$

Meanwhile, the shot noise term in the power spectrum grows as the square of the particle weights, so we want to minimize

$$\tfrac{1}{2}\langle w^2\rangle = \int \mathrm{d}^3p\; \tfrac{1}{2}w(x,p,t)^2 g(x,p,t),$$

subject to the constraint $\langle\epsilon\rangle \leq \eta$ for some maximum error η. Assume that the bound is saturated. First, let us look for solutions that extremize the integral constraint. We find the unique solution

$$w = \frac{\delta f}{g} \quad \text{with} \quad \delta f = f - \alpha\bar{f}. \tag{5}$$

This is the δf method introduced above, with optimal α given by (3). Any further solution should extremize the Lagrangian,

$$\begin{aligned} \mathcal{L} = &\ \tfrac{1}{2}w(x,p,t)^2 g(x,p,t) \\ &+ \lambda\Big(w(x,p,t)g(x,p,t) + \alpha(t)\bar{f}(x,p,t) - f(x,p,t)\Big). \end{aligned}$$

Writing down the Euler–Lagrange equations

$$[w+\lambda]\, g\nabla_p w + \tfrac{1}{2}w^2\nabla_p g + \lambda w\nabla_p g = \lambda\nabla_p f - \alpha\lambda\nabla_p\bar{f},$$

one finds a family of quadratic solutions

$$w = -\lambda \pm \sqrt{\lambda^2 + 2\lambda\frac{\delta f}{g}} \quad \text{with} \quad \delta f = f - \alpha \bar{f}.$$

The case $\lambda = 0$ corresponds to the trivial solution $w = 0$. For $\lambda \neq 0$, we obtain the minima

$$w = \frac{\delta f}{g} - \frac{1}{2\lambda}\left(\frac{\delta f}{g}\right)^2 + \mathcal{O}\left(\frac{1}{\lambda^{2/3}}\frac{\delta f}{g}\right)^3. \tag{6}$$

These solutions correspond to small perturbations around the δf method that trade some accuracy for a possible reduction in shot noise. However, since the leading correction is $\propto (\delta f)^2$, this is only possible if the background model is skewed with respect to the non-linear solution. Typically, the skewness and the additional reduction in shot noise are negligible. In fact, since the next-to-leading correction is positive, shot noise increases if the skewness is small.

We have shown that within the broad class of hybrid methods described by equation (4), δf-type methods of the form (6) minimize the amount of shot noise, subject to the constraint that the error in the number density remains below a certain bound. The δf method given by (5), recovered from (6) in the limit $\lambda \to \infty$, is the unique solution for which the expected error $\langle\epsilon\rangle = 0$. The optimal value of α is given by (3), but will be close to 1 if $\bar{f} \approx f$. This is the method we will use exclusively, with the choice $\alpha = 1$.

3 1D EXAMPLE

We now illustrate the method by applying it to a 1D test problem with a known solution. Readers that are satisfied with the mathematical derivation may skip ahead to Section 4.

3.1 The elliptical sine wave

Consider the 1D collisionless Boltzmann equation

$$\frac{\partial f}{\partial t} + p\frac{\partial f}{\partial x} - \frac{\partial \Phi}{\partial x}\frac{\partial f}{\partial p} = 0,$$

where the particles move under a conservative force $F(x) = -\Phi^{'}(x)$. Let us assume a periodic potential given by

$$\Phi(x) = \sin^2(x/2).$$

The steady-state solution can be found to be:

$$f(x, p) = \frac{\bar{\rho}}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{p^2}{2\sigma^2} + \frac{\cos(x)}{2\sigma^2}\right),$$

in terms of the background density $\bar{\rho}$ and velocity dispersion σ. The corresponding density profile $\rho(x)$ is given by

$$\rho(x) = \int_{-\infty}^{\infty} f(x, p)\mathrm{d}p = \bar{\rho}\exp\left(\frac{\cos(x)}{2\sigma^2}\right).$$

To find the general time-dependent solution, we use the method of characteristics. The characteristic equations are

$$\frac{\mathrm{d}x}{\mathrm{d}t} = p, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{1}{2}\sin(x).$$

These equations of motion can be solved in terms of the energy $E = \frac{1}{2}p^2 + \sin^2(x/2)$, which gives

$$\sin(x/2) = \mathrm{sn}\left(\pm\sqrt{E/2}(t - \tau)\right),$$

where τ is an integration constant and sn(x) is the Jacobi elliptic sine function with elliptic modulus $k = 1/\sqrt{E}$ (Weisstein 2002).[1] Assuming a homogeneous Gaussian distribution with mean $\bar{p}$ for the initial momenta p at time $t = 0$,

$$f(x, p, 0) = \frac{\bar{\rho}}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(p - \bar{p})^2}{2\sigma^2}\right), \tag{7}$$

the general solution, $f(x, p, t)$, at later times is a complicated expression involving elliptic sines and arcsines. The details are given in Appendix A. We replicate the problem using N-body methods. A large number of particles are initialized on the interval $x \in [0, 4\pi]$ with momenta drawn from the initial distribution (7). The particles are then integrated using

$$\Delta x = p\Delta t, \qquad \Delta p = -\frac{1}{2}\sin(x)\Delta t.$$

In addition to the ordinary N-body method, we use a δf method, where the background model is given by

$$\bar{f}(x, p, t) = \frac{\bar{\rho}}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(p - \bar{p})^2}{2\sigma^2}\right),$$

and the weights are updated during each step via $w = \delta f/f$. The corresponding density profiles are shown in Fig. 2. The plots were created using $N = 10^6$ particles and the model parameters are $\bar{\rho} = \sigma = 1$ and $\bar{p} = 10$. The results show that both the ordinary N-body simulation and the simulation with a δf step can reproduce the exact solution. However, the ordinary method is very noisy, whereas the δf method reproduces the expected profiles with remarkable accuracy. The reason for this is that while the distribution itself has a large dispersion, resulting in noisy results for the ordinary method, the perturbations from the steady solution are small, which allows the δf method to work. This is exactly analogous to the cosmic neutrino background.

4 RELATIVISTIC EFFECTS

Neutrinos constitute a relativistic fluid at early times, which introduces some subtleties when evolving such a fluid with a Newtonian code. Including relativistic effects is not necessary for the δf method, but we include them in our simulations to allow for a consistent comparison with recent works (Adamek et al. 2017; Tram et al. 2019; Partmann et al. 2020). Furthermore, the higher order δf methods discussed in Section 7 provide a natural setting for including these effects without neglecting the non-linear evolution of the neutrinos. We will work in the Newtonian motion framework of Fidler et al. (2017a) and make modifications to the initial conditions, long-range force calculation, and particle equations of motion as outlined below.

4.1 Initial conditions

To generate initial conditions for massive neutrinos and to set-up the higher order background models (Section 7), accurate calculation of the linear theory neutrino distribution function is indispensable. This can be done with the Boltzmann codes CAMB (Lewis & Challinor 2011) and CLASS (Lesgourgues 2011). At their default settings, these codes produce accurate total matter and radiation power spectra

---

[1] For $E \to \infty$, we have $\mathrm{sn}\,x \to \sin x$, meaning that $x \propto t$. The particle 'ignores' the potential. For $E = k = 1$, $\mathrm{sn}\,x = \tanh x$, meaning the particle asymptotically approaches a potential peak. For $E < 1$, the particle is bounded and oscillates between peaks.

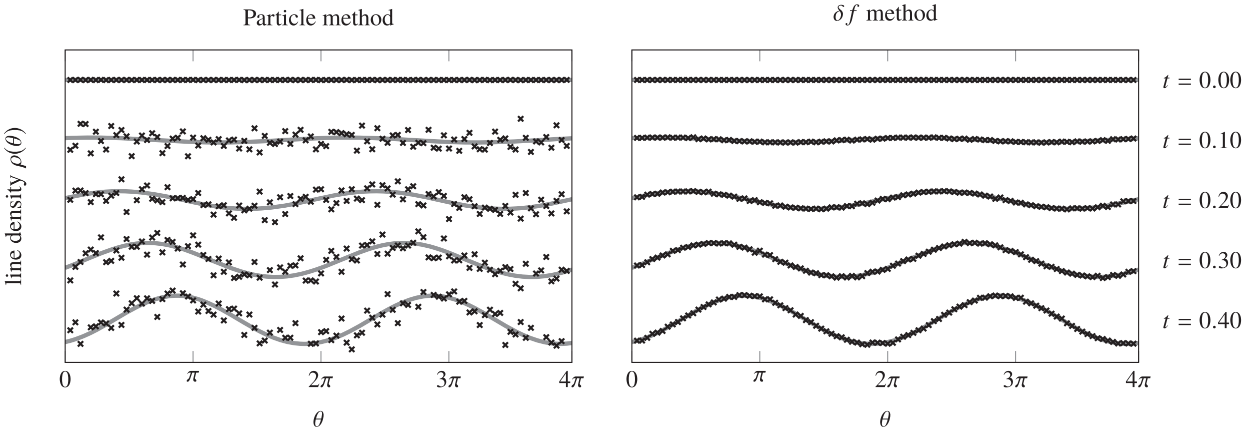

Figure 2. Density profiles for the 1D elliptical sine wave test problem. We counted particles in 100 bins of width $\Delta\theta = 4\pi/100$ to create the empirical density profiles. On the left, an ordinary *N*-body simulation with $N = 10^6$ particles was used. On the right, the *N*-body simulation was extended with a δf step.

(their intended purpose), but the neutrino related transfer functions (e.g. density and velocity) are not converged and can be very inaccurate (Dakin et al. 2019). To obtain converged results, we post-process perturbation vectors from CLASS by integrating source functions up to multipole $\ell_{\max} = 2000$. This prevents the artificial reflection that can happen for low $\ell_{\max}$. See Appendix B for more details.

Initial conditions are then created using the post-processed transfer functions from CLASS in *N*-body gauge at $z = 100$. We do not follow the usual approach of back-scaling the present-day power spectrum, but use the so-called forward Newtonian motion approach (Fidler et al. 2015, 2017a). To our knowledge, forward Newtonian motion initial conditions have always been set-up with the Zel'dovich approximation. However, this approximation is known to be inadequate for precision simulations (Crocce, Pueblas & Scoccimarro 2006). To go beyond Zel'dovich initial conditions, we determine the Lagrangian displacement vectors $\boldsymbol{\psi} = \mathbf{x} - \mathbf{q}$ by solving the Monge–Ampère equation

$$\rho(\mathbf{x}) = \bar{\rho}(1 + \delta(\mathbf{x})) = \det\left[1 + \frac{\partial\psi_i(\mathbf{x})}{\partial q_j}\right]^{-1}.$$

This equation is solved numerically with a fixed-point iterative algorithm that exploits the fact that the density perturbation δ is small. We note that this approach is not equivalent to Lagrangian perturbation theory, but merely provides a more accurate map from the Eulerian initial density field to a Lagrangian displacement field compared to the Zel'dovich approximation. A detailed analysis of this method will be presented elsewhere. Velocities were determined independently using the transfer function for the velocity dispersion $\theta = ik \cdot v$.

We used two methods for setting up the neutrino particles. In the first method, neutrino particles are displaced randomly in phase space according to the perturbed phase-space density function, $f_{\mathrm{PT}}(x, p, t)$, including terms up to $\ell = 5$. This method was used for our $(256\,\mathrm{Mpc})^3$ simulations. In the second method, which accounts also for higher multipoles, neutrino particles are integrated in linear theory from $z = 10^9$ to $z = 10^2$ (Ma & Bertschinger 1994a; Adamek et al. 2017). Both methods produce the same result on small scales, but the latter method was found to be more accurate on large scales, $k < 10^{-2}\,\mathrm{Mpc}^{-1}$. For this reason, we used the second method for our $(1\,\mathrm{Gpc})^3$ simulations.

4.2 Long-range forces

In a relativistic setting, the gravitational evolution is governed by the Einstein–Boltzmann equations. We will approximate this system using a hybrid approach (Brandbyge et al. 2017), in which dark matter and massive neutrinos are evolved using a Newtonian *N*-body code complemented with relativistic corrections to the fluid equations that are pre-solved in linear theory. We will work in *N*-body gauge, which allows the fluid equations for dark matter to be written in a particularly convenient form resembling the Newtonian equations solved by conventional *N*-body codes. The continuity and Euler equations can then be written as (Fidler et al. 2015, 2017b)

$$\dot{\delta} + \nabla \cdot \mathbf{v} = 0,$$
$$\dot{\mathbf{v}} + aH\mathbf{v} = -\nabla\phi + \nabla\gamma^{\mathrm{Nb}},$$

where overdots denote conformal time derivatives, δ is the density contrast, $\mathbf{v}$ the peculiar velocity, and $H = \dot{a}/a^2$. All relativistic corrections are captured by the *N*-body gauge term, $\nabla\gamma^{\mathrm{Nb}}$, which arises from the anisotropic stress of relativistic species. In addition, the scalar potential ϕ receives contributions from all fluid components:

$$\nabla^2\phi = 4\pi G a^2 \sum_i \delta\rho_i,$$

where the sum runs over cold dark matter, baryons, neutrinos, and photons. Density perturbations are actively calculated for all species. In the case of massive neutrinos and the cold dark matter and baryon fluid, this is done with particles in the usual way. For photons and massless neutrinos (and for some runs, the massive neutrinos[2]), this is done by realizing the corresponding transfer functions from CLASS on a grid as part of the long-range force calculation in our *N*-body code SWIFT.

In the absence of relativistic species, the *N*-body gauge term, $\nabla\gamma^{\mathrm{Nb}}$, vanishes and the continuity and Euler equations agree with the Newtonian equations solved in conventional *N*-body codes. This makes *N*-body gauge useful as it allows one to set up initial conditions in *N*-body gauge, evolve them in a Newtonian simulation, and give the results a relativistic interpretation. The relativistic corrections become relevant at the 0.5 per cent level on the largest scales in

---

[2] Specifically, the linear theory runs and the runs with higher order δf methods, as discussed in Sections 6 and 7, respectively.

our Gpc simulations. Ordinarily, these corrections are accounted for in the initial conditions by backscaling the present-day linear power spectrum, ensuring that the linear power spectrum is recovered on large scales at $z = 0$. Here, instead, we actively include these contributions with the aim of recovering the linear power spectrum at earlier times as well.

4.3 Particle content

When simulating light neutrinos from high redshifts, we are evolving relativistic particles in a Newtonian simulation. Such particles can reach superluminal speeds when evolved using the ordinary equations of motion. Following Adamek et al. (2016), we initially addressed this issue by replacing the equations of motion with special relativistic equations that are valid to all orders in u:

$$\dot{\mathbf{u}} = -\frac{2u^2 + a^2}{\sqrt{u^2 + a^2}} \nabla \left(\phi - \gamma^{\mathrm{Nb}}\right), \tag{8}$$

$$\dot{\mathbf{x}} = \frac{\mathbf{u}}{\sqrt{u^2 + a^2}}. \tag{9}$$

Here, a is the scale factor and $\mathbf{u}$ the comoving 3-velocity. However, we encountered two problems with our original approach. The first is that a direct leapfrog implementation of these equations is not symplectic due to the fact that the right-hand side of (8) depends on $\mathbf{u}$. As a consequence, phase-space density is not exactly conserved by the discrete Hamiltonian, as assumed by the δf method. We investigate this issue further in Appendix D, where we offer an alternative solution and conclude that this is not an issue in practice. The second problem is that using these equations, we did not exactly reproduce linear theory on large scales. We found that this was due to a breakdown of (8) in the ultra-relativistic limit. While it is possible to include relativistic corrections to the acceleration in a manner that is consistent with linear theory[3], a simple alternative is to use the special relativistic equation (9) together with the non-relativistic version of (8):

$$\dot{\mathbf{u}} = -a \nabla \left(\phi - \gamma^{\mathrm{Nb}}\right). \tag{10}$$

This choice not only ensures that the neutrinos move the correct subluminal distance and that the integrator is symplectic, but also avoids the problem on large scales. We used this second approach for our Gigaparsec simulations. On small scales, the evolution of the neutrinos at late times is insensitive to their early evolution, making the difference between (8) and (10) immaterial. For this reason, we present the original results for our 256 Mpc simulations, which used the first approach with (8) and (9).

A separate matter from the equations of motion is that the neutrinos have a relativistic energy at early times. Using only the mass–energy, as is done for non-relativistic matter, leads to underestimation of the matter power spectrum on large scales. To rectify this, we replace the weighted mass of the particles with a weighted energy $\epsilon = m\sqrt{a^2 + u^2}$. Here, again the issue of the symplectic integrator plays a role. To ensure that $\dot{\mathbf{u}}$ does not depend on current particle velocities, we used the alternative form $\epsilon = m\sqrt{a^2 + u_0^2}$ for our Gigaparsec simulations. Here, u_0 is the initial particle velocity at $z = 100$. This approximation is extremely accurate, as substantial deviations from u_0 only occur for slow particles at late times in which case the mass term dominates (see Appendix D). For the 256 Mpc simulations, we used the original form $\epsilon = m\sqrt{a^2 + u^2}$.

Table 1. Description of the simulations. The listed particle mass, m_{p}, refers to the combined cold dark matter and baryon particles. The neutrino fraction is listed as $f_\nu = \Omega_\nu/(\Omega_{\mathrm{cb}} + \Omega_\nu)$.

Side Length	N_c	$m_{\mathrm{p}}\ [\mathrm{M}_\odot]$	N_ν	$\sum m_\nu$	f_ν
1024 Mpc	1024^3	3.96×10^{10}	0	0 meV	0
1024 Mpc	1024^3	3.93×10^{10}	1024^3	100 meV	0.0073
1024 Mpc	1024^3	3.81×10^{10}	1024^3	500 meV	0.0376
256 Mpc	512^3	4.95×10^{9}	0	0 meV	0
256 Mpc	512^3	4.92×10^{9}	1024^3	100 meV	0.0073
256 Mpc	512^3	4.77×10^{9}	1024^3	500 meV	0.0376

5 SIMULATIONS

We now describe our neutrino simulations, which were run on the COSMA6 computing facility in Durham. We have implemented the δf method in the cosmological hydrodynamics code SWIFT (Schaller et al. 2016, 2018). SWIFT uses a combination of the Fast Multipole Method for short-range gravitational forces and the Particle Mesh method for long-range forces. It uses a 5th degree polynomial kernel for the force softening with a single time-dependent softening length. The code uses a task-based parallelization paradigm to achieve strong scaling on large clusters and obtain significant speed-ups over competing N-body codes. The main simulations presented in this paper use the basic version of the δf method with a homogeneous Fermi–Dirac distribution as background model. Our choice of cosmological parameters, based on Planck 2018 (Planck Collaboration VI 2020), is $(h, \Omega_{\mathrm{c}} + \Omega_\nu, \Omega_{\mathrm{b}}, A_s, n_s) = (0.6737, 0.265, 0.0492, 2.097 \times 10^{-9}, 0.9652)$. We run two sets of simulations at different resolution to test the large-scale and small-scale behaviour of various methods. The cube sizes and particle numbers are listed in Table 1.

5.1 Choice of neutrino masses

Neutrino oscillations indicate that there are three neutrino mass eigenstates with unknown masses m_i. The mass splittings have been measured with good precision to be

$$\Delta m^2_{21} \equiv m^2_2 - m^2_1 = 7.42^{+0.21}_{-0.20} \times 10^{-5}\,\mathrm{eV}^2,$$

$$\Delta m^2_{3\ell} \equiv m^2_3 - m^2_\ell = \begin{cases} +2.514^{+0.028}_{-0.027} \times 10^{-3}\,\mathrm{eV}^2\ (\mathrm{NO}), \\ -2.497^{+0.028}_{-0.028} \times 10^{-3}\,\mathrm{eV}^2\ (\mathrm{IO}). \end{cases}$$

The sign of Δm^2_{21} is known to be positive, which leaves two possible mass orderings: $m_1 < m_2 < m_3$ with $\ell = 1$ (normal) or $m_3 < m_1 < m_2$ with $\ell = 2$ (inverted). Current oscillation data slightly favour the normal ordering at 1.6σ (Esteban et al. 2020).

The best terrestrial constraint on the absolute mass scale comes from the KATRIN detector, which places a bound of $m_\beta < 0.8\,\mathrm{eV}$ at the 90 per cent C.L. on the effective neutrino mass (Aker et al. 2021). Assuming a quasi-degenerate mass spectrum, this corresponds to a neutrino mass sum of $\sum m_\nu < 2.4\,\mathrm{eV}$. Recent cosmological limits are much stronger and are quoted below at the 95 per cent C.L. Assuming a degenerate mass spectrum, the Planck temperature, polarization, and lensing likelihoods give a constraint of $\sum m_\nu < 0.24\,\mathrm{eV}$ or $\sum m_\nu < 0.26\,\mathrm{eV}$, depending on the details of the high-ℓ polarization analysis (Planck Collaboration VI 2020). Adding BAO data from BOSS DR12, MGS, and 6dFGS, Choudhury & Hannestad (2020) found $\sum m_\nu < 0.12\,\mathrm{eV}$ (degenerate), $\sum m_\nu < 0.15\,\mathrm{eV}$ (normal), and $\sum m_\nu < 0.17\,\mathrm{eV}$ (inverted). An analysis of the shape of the BOSS DR11 redshift-space power spectrum, combined with CMB data and Type 1a supernovae leads to $\sum m_\nu < 0.18\,\mathrm{eV}$

---

[3] This is important at early times and therefore relevant for generating neutrino initial conditions (Elbers, Frenk, Jenkins, Li and Pascoli, in preparation).

(Upadhye 2019). Adding instead the latest SDSS DR14 BOSS and eBOSS Lyman-α forest data to the Planck and BAO data leads to the strongest constraint: $\sum m_\nu < 0.09$ eV (Palanque-Delabrouille et al. 2020).

Given these limits, we consider three values for $\sum m_\nu$, keeping the present-day value $\Omega_{\mathrm{m},0} = \Omega_{\mathrm{cb},0} + \Omega_{\nu,0}$ fixed. Scenario one contains three massless neutrinos, scenario two corresponds to the inverted mass ordering with $\sum m_\nu = 100$ meV,[4] and scenario three to a degenerate spectrum with $\sum m_\nu = 500$ meV. The first two models bracket the most interesting range of values $0 < \sum m_\nu < 100$ meV. The last model has surely been ruled out, but is included for several reasons. First of all, the δf method reduces to the ordinary particle method in the large mass limit at late times. Hence, the $\sum m_\nu = 500$ meV case provides a useful consistency check. Secondly, when simulations are used to emulate statistics for parameter extraction, we should allow for unlikely excursions in MCMC analyses without our simulation methods breaking down (Partmann et al. 2020). Finally, in the extended parameter space around ΛCDM, for example with a non-standard lensing amplitude, A_L, or curvature, or when varying the dark energy equation of state, the possibility of larger neutrino masses remains very relevant (McCarthy et al. 2018; Upadhye 2019; Choudhury & Hannestad 2020; Di Valentino, Melchiorri & Silk 2020).

The two massive scenarios considered in this paper have degenerate neutrino masses (2 × 50 meV and 3 × 167 meV). However, the δf method can easily be extended to account for mass splittings. In that case, particles would be labelled with a given mass state, i, and each state would have its own background model, $\bar{f}_i$. The reduction in shot noise is largest for the smallest neutrino masses, placing different masses on a level footing. This allows for better load balancing between different neutrino masses.

6 RESULTS

We compare our neutrino δf method with three commonly used alternatives. The most common alternative is the ordinary N-body particle method, which is the same in every respect as our method, but with the weighting step disabled. Next, we consider a linear theory method based on Tram et al. (2019) that does not evolve neutrino particles but instead realizes the linear theory neutrino perturbation in N-body gauge on a grid. The neutrinos are then fully accounted for in the long-range forces. Finally, we consider the linear response method of Ali-Haïmoud & Bird (2012) in which the neutrino perturbation is calculated by applying the linear theory transfer function ratio, $\delta_\nu^{\mathrm{lin}}(k)/\delta_{\mathrm{cdm+b}}^{\mathrm{lin}}(k)$, to the simulated cdm+baryon phases.

A visual inspection of the neutrino density plots shown in Figs 3 and 4 reveals the strengths and weaknesses of the four methods. Broadly, we see that the linear theory method does not suffer from shot noise, but fails to reproduce the small-scale behaviour resolved by the particle and δf methods. At the same time, shot noise is clearly visible in the particle simulation with $\sum m_\nu = 100$ meV, despite using $N_\nu = 1024^3$ particles in a 256 Mpc cube. This is evidently cured in the δf plot. We also see that shot noise is much less of a problem for $\sum m_\nu = 500$ meV, but the δf plot is still less grainy than the corresponding particle plot. Finally, Fig. 4 shows that the linear response method greatly improves on the pure linear theory prediction, but still produces neutrino haloes that are too diffuse compared to the particle and δf simulations.

6.1 Neutrino component

We start with an analysis of the probability density function of the neutrino density field, computed on a 1024^3 grid from the 256 Mpc simulations. Refer to the plots in Fig. 5, which bear out the basic picture sketched above. For the $\sum m_\nu = 100$ meV neutrinos, the particle method is plagued by shot noise, but agrees with the δf method in the high-density tail where the particle number is sufficient to obtain a good signal-to-noise ratio. The linear prediction fails in the high and low-density tails. Finally, the linear response method, which applies the linear theory ratio $\delta_\nu(k)/\delta_{\mathrm{cdm+b}}(k)$ to the cdm+baryon phases, is an intermediate case between the linear theory and δf methods. For the more massive scenario, the situation is much the same, except that shot noise is much less of a problem for the particle method on these scales.

Next, we consider two-point statistics and show the neutrino power spectrum at $z = 0$ in Fig. 6, combining the large and small simulations to show a wide range of scales. We use the Gpc simulations for $k < 0.1\,\mathrm{Mpc}^{-1}$ and the 256 Mpc simulations for $k \geq 0.1\,\mathrm{Mpc}^{-1}$. As expected, all methods agree on scales greater than $k = 0.1\,\mathrm{Mpc}^{-1}$ for both neutrino masses. On smaller scales, linear theory significantly underpredicts the amount of neutrino clustering. The linear response method also underpredicts the neutrino power spectrum, but not by as much. The relative difference between the non-linear power spectrum and linear power spectrum is greater for neutrinos than for cdm and baryons. To account for this effect, we fit a non-linear correction to the linear response power spectrum using the measured δf power spectrum up to $k = 1\,\mathrm{Mpc}^{-1}$:

$$P_\nu^{\mathrm{fit}}(k) = P_{\mathrm{cdm+b}}(k)\left[\frac{\delta_\nu^{\mathrm{lin}}(k)}{\delta_{\mathrm{cdm+b}}^{\mathrm{lin}}(k)}\right]^2 e^{\alpha+\beta k}, \tag{11}$$

and find $\alpha = 0.006 \pm 0.004$ and $\beta = 0.90 \pm 0.01$ ($\sum m_\nu = 100$ meV) and $\alpha = -0.06 \pm 0.03$ and $\beta = 0.34 \pm 0.09$ ($\sum m_\nu = 500$ meV). These are shown as the red curves in Fig. 6.

The particle simulations are clearly affected by shot noise, at the level of $V/N = 1/64$, obscuring the neutrino signal on scales smaller than $k = 0.2\,\mathrm{Mpc}^{-1}$ for the lightest scenario and on scales smaller than $k = 1\,\mathrm{Mpc}^{-1}$ for the more massive scenario. Using the δf method, shot noise is significantly reduced in the former case (factor of 87) and slightly reduced in the latter case (factor of 3.5), revealing a signal down to $k = 1$–$2\,\mathrm{Mpc}^{-1}$. Hence, δf simulations can achieve a similar resolution independently of mass without adjusting the particle number.

We also show the cross-spectral coefficient

$$r_{\nu,\mathrm{cb}}(k) = \frac{P_{\nu,\mathrm{cb}}(k)}{\sqrt{P_\nu(k)P_{\mathrm{cb}}(k)}},$$

which captures phase differences between the dark matter and neutrinos. By definition, $r_{\nu,\mathrm{cb}} = 1$ according to the linear response method. However, this does not hold on small scales as can be seen in the bottom panels. Up to the point where shot noise becomes a problem, the particle and δf methods agree, demonstrating that $r_{\nu,\mathrm{cb}} < 1$. This is particularly clear for $\sum m_\nu = 500$ meV.

Next, we consider how well the simulations can resolve the extended neutrino haloes surrounding galaxies and clusters (Brandbyge et al. 2010; Villaescusa-Navarro et al. 2011). In Fig. 7, we show stacked neutrino profiles for haloes with virial mass $M_{\mathrm{cdm+b}}$ in the range $(5, 12) \times 10^{14}\,\mathrm{M}_\odot$. The particle and δf methods agree almost perfectly, once again because of the high signal-to-noise ratio in the largest overdensities. In linear theory, the neutrino haloes are

---

[4] Specifically, two 0.0486 eV neutrinos and one massless neutrino.

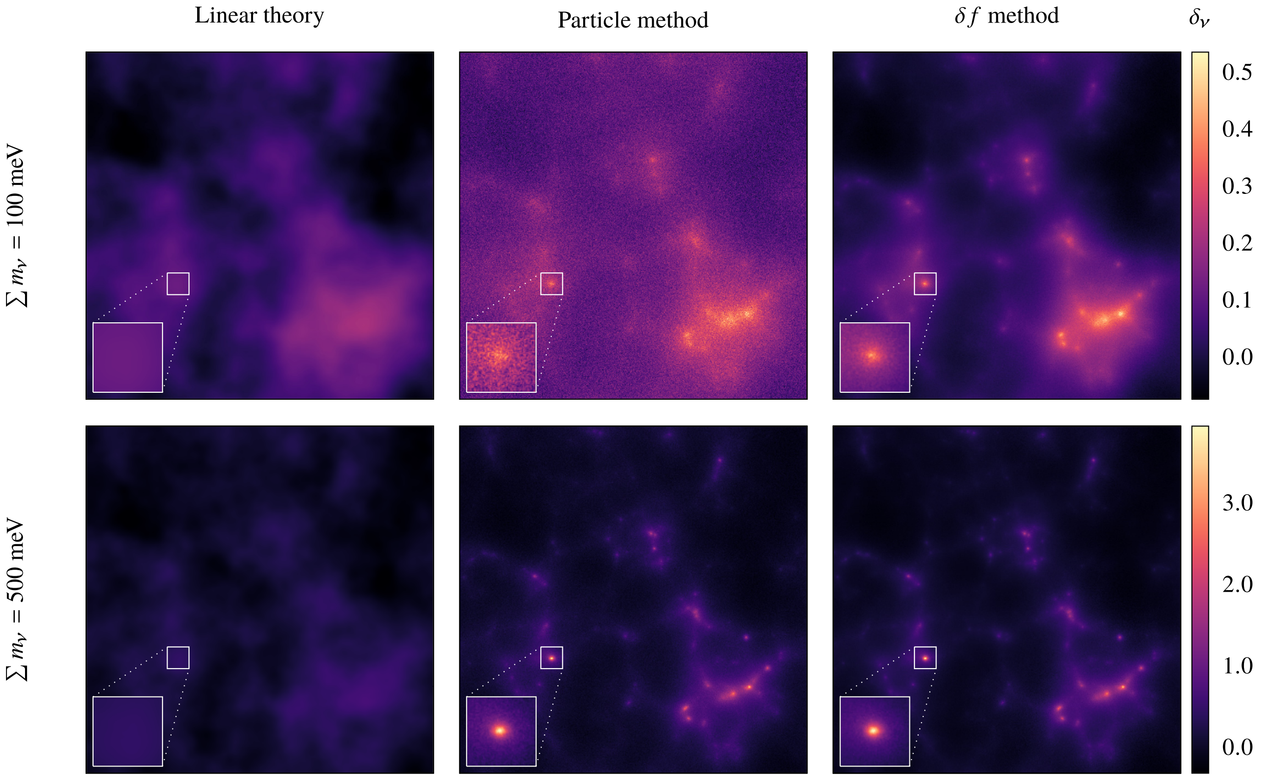

Figure 3. Neutrino density plots of $(256\,\mathrm{Mpc})^3$ cubes at $z = 0$, simulated with two commonly used methods and with the δf method. The particle and δf simulations used $N_\nu = 1024^3$ particles. Shot noise is clearly visible for the particle method, although noticeably less so for $\sum m_\nu = 500\,\mathrm{meV}$. The linear theory model fails to reproduce the small-scale behaviour. The δf method solves both problems. The inset zooms in on a neutrino halo and should also be compared with the linear response prediction in Fig. 4, which uses the same colour scale. The neutrino fraction $f_\nu = \Omega_\nu/(\Omega_{\mathrm{cb}} + \Omega_\nu) = 0.0073$ for the top row and $f_\nu = 0.0376$ for the bottom row.

completely absent as is evident also from the cross-sections in Fig. 3. Finally, the linear response method predicts neutrino haloes that are too diffuse compared to the non-linear simulations, and with too little dispersion from the mean profile. The larger dispersion found in the non-linear simulations is not due to errors in individual profiles, but due to a stronger correlation between $M_{\mathrm{cdm+b}}$ and the local neutrino density.

6.2 Neutrino bias

On larger scales, the neutrino density field can be reconstructed from the density of haloes for a given neutrino mass spectrum (Inman et al. 2015). We therefore construct the halo overdensity field,

$$\delta_{\mathrm{h}}(x) = \frac{n_{\mathrm{h}}(x) - \bar{n}_{\mathrm{h}}}{\bar{n}_{\mathrm{h}}},$$

by calculating the number density, $n_{\mathrm{h}}(x)$, of haloes and the mean density, $\bar{n}_{\mathrm{h}}$, at $z = 0$ in our Gpc simulations identified using the halo finder VELOCIRAPTOR (Elahi et al. 2019). We restrict attention to haloes with virial mass, $M_{\mathrm{cdm+b}} > 10^{12}\,\mathrm{M}_\odot$, and smooth δ_{h} and δ_ν with a tophat filter of comoving radius $R = 30\,h^{-1}$ Mpc. Following Yoshikawa et al. (2020), we study the mean neutrino density at constant halo density $\bar{\delta}_\nu(\delta_{\mathrm{h}})$, defined in terms of the joint probability density function $P(\delta_\nu, \delta_{\mathrm{h}})$ as

$$\bar{\delta}_\nu(\delta_{\mathrm{h}}) = \int \mathrm{d}\delta_\nu \delta_\nu P(\delta_\nu, \delta_{\mathrm{h}}).$$

This relationship is close to linear with slope equal to the neutrino bias, given by

$$b = \frac{\langle \delta_\nu \delta_{\mathrm{h}} \rangle}{\langle \delta_{\mathrm{h}}^2 \rangle}.$$

The degree of non-linearity is captured by

$$\epsilon_{\mathrm{nl}}^2 = \frac{\left\langle \delta_{\mathrm{h}}^2 \right\rangle \left\langle \bar{\delta}_\nu^2 \right\rangle}{\left\langle \bar{\delta}_\nu \delta_{\mathrm{h}} \right\rangle^2} - 1,$$

which satisfies $\epsilon_{\mathrm{nl}} = 0$ if and only if the slope of $\bar{\delta}_\nu(\delta_{\mathrm{h}})$ is independent of δ_{h}. The scatter around the biasing relationship is characterized by the stochasticity,

$$\epsilon_{\mathrm{stoch}}^2 = \frac{\left\langle \delta_{\mathrm{h}}^2 \right\rangle \left\langle (\delta_\nu - \bar{\delta}_\nu)^2 \right\rangle}{\left\langle \bar{\delta}_\nu \delta_{\mathrm{h}} \right\rangle^2}.$$

The non-linearity and stochasticity are related to the correlation coefficient,

$$r_{\nu,\mathrm{h}} = \frac{\langle \delta_\nu \delta_{\mathrm{h}} \rangle}{\sqrt{\langle \delta_\nu^2 \rangle \langle \delta_{\mathrm{h}}^2 \rangle}},$$

via $r_{\nu,\mathrm{h}} \simeq (1 + \epsilon_{\mathrm{nl}}^2 + \epsilon_{\mathrm{stoch}}^2)^{-1/2}$. This model is analogous to the non-linear stochastic galaxy biasing model of Taruya & Suto (2000) and Yoshikawa et al. (2001). We compute the four quantities $(b, \epsilon_{\mathrm{nl}}^2, \epsilon_{\mathrm{stoch}}^2, r_{\nu,\mathrm{h}})$ for each of the methods under consideration. The results are listed in Table 2 and the biasing relationship is shown in Fig. 8. As expected on these large scales, we find good agreement with differences of a few per cent in the bias. The greater the level of

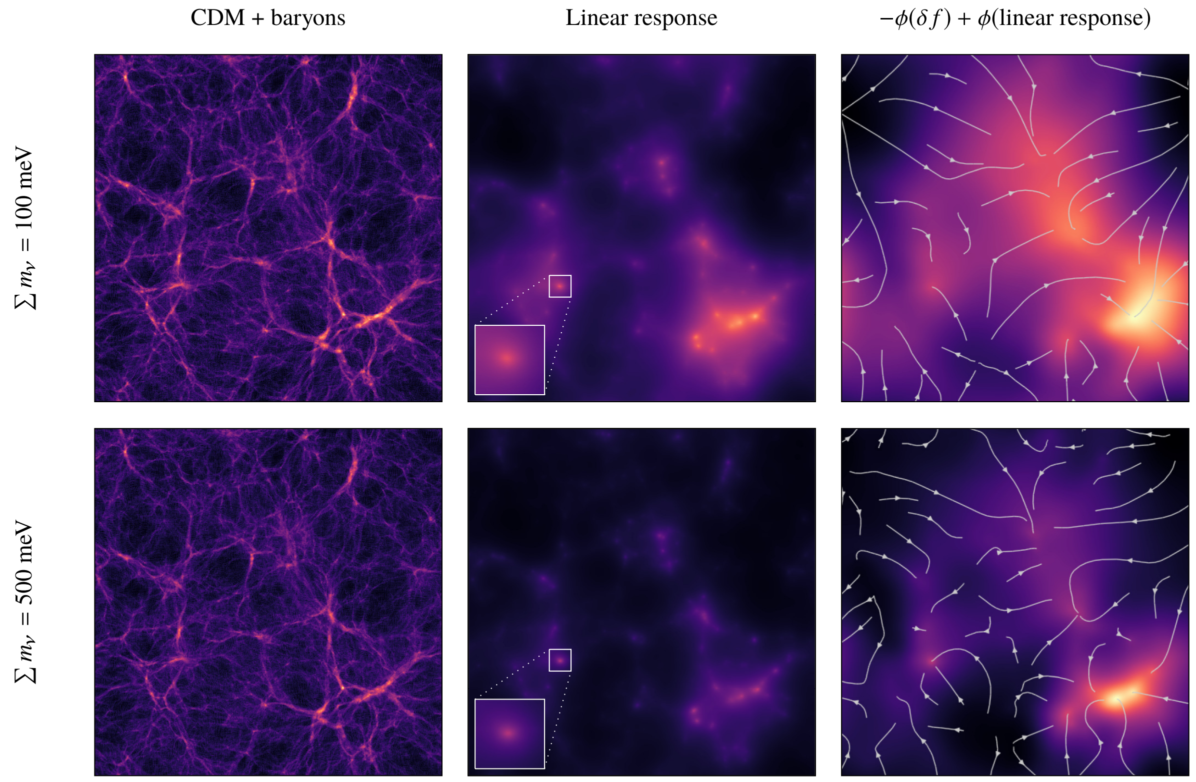

Figure 4. Density plots of $(256\,\mathrm{Mpc})^3$ cubes at $z = 0$. The linear response method applies the linear theory ratio $\delta_\nu^{\mathrm{lin}}(k)/\delta_{\mathrm{cb}}^{\mathrm{lin}}(k)$ to the simulated cdm+baryon phases (Ali-Haïmoud & Bird 2012). Compared to the linear theory prediction, it performs remarkably well, but the neutrino haloes around clusters are significantly more diffuse compared to the particle and δf simulations (compare the zoomed in halo with the δf prediction in Fig. 3). The resulting potential difference is shown in the last column, with flowlines indicating the forces that are not present in the linear response model.

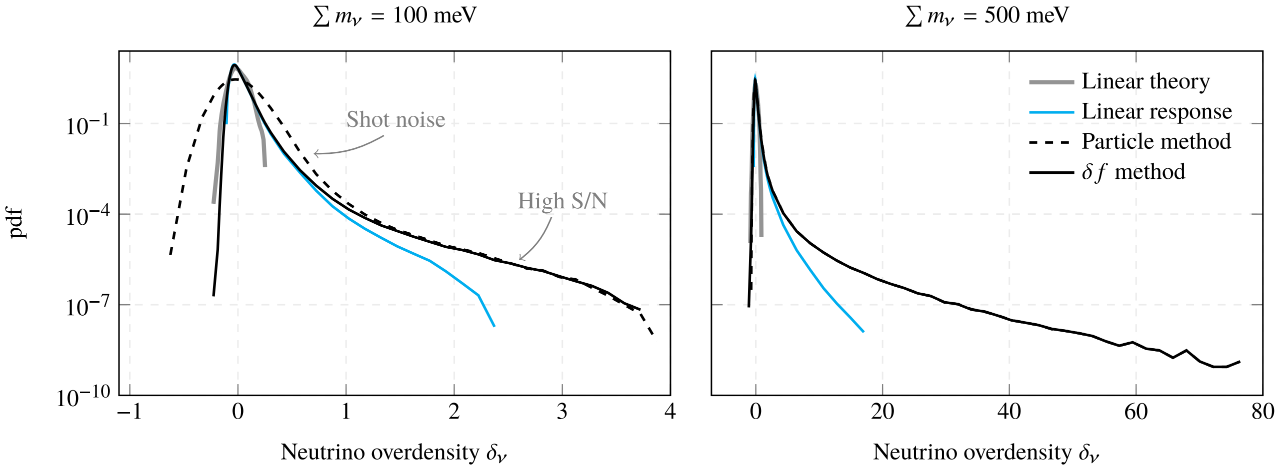

Figure 5. Neutrino probability density functions (pdf) at $z = 0$, computed on a 1024^3 grid from the 256 Mpc simulations, and smoothed with a Gaussian filter with radius $R = 256$ kpc. We compare the δf method with three commonly used alternatives. The particle and δf methods agree in the high density tail, because the largest overdensities have enough particles to achieve a high signal-to-noise ratio. Shot noise plagues the particle method, particularly in underdense regions. The linear methods fail in the high density tail.

neutrino clustering resolved by a given method, the greater the bias b and correlation $r_{\nu,\mathrm{h}}$. The stochasticity follows the opposite pattern. The non-linearity follows no such pattern, but is very small in each case except (amusingly) for the linear theory runs. This is because linear theory does not resolve neutrino haloes, causing the $\bar{\delta}_\nu(\delta_\mathrm{h})$ relation to level off in the high-density tail.

The bias $b = 0.103$ for the 100 meV scenario is in excellent agreement with the bias $b = 0.071$ found by Yoshikawa et al. (2020), when the difference in mass ordering is factored in using the approximately linear relationship between neutrino mass and bias in their results. Yoshikawa et al. (2020) do not consider neutrino masses beyond 400 meV, but our finding of $b = 0.256$ for 500

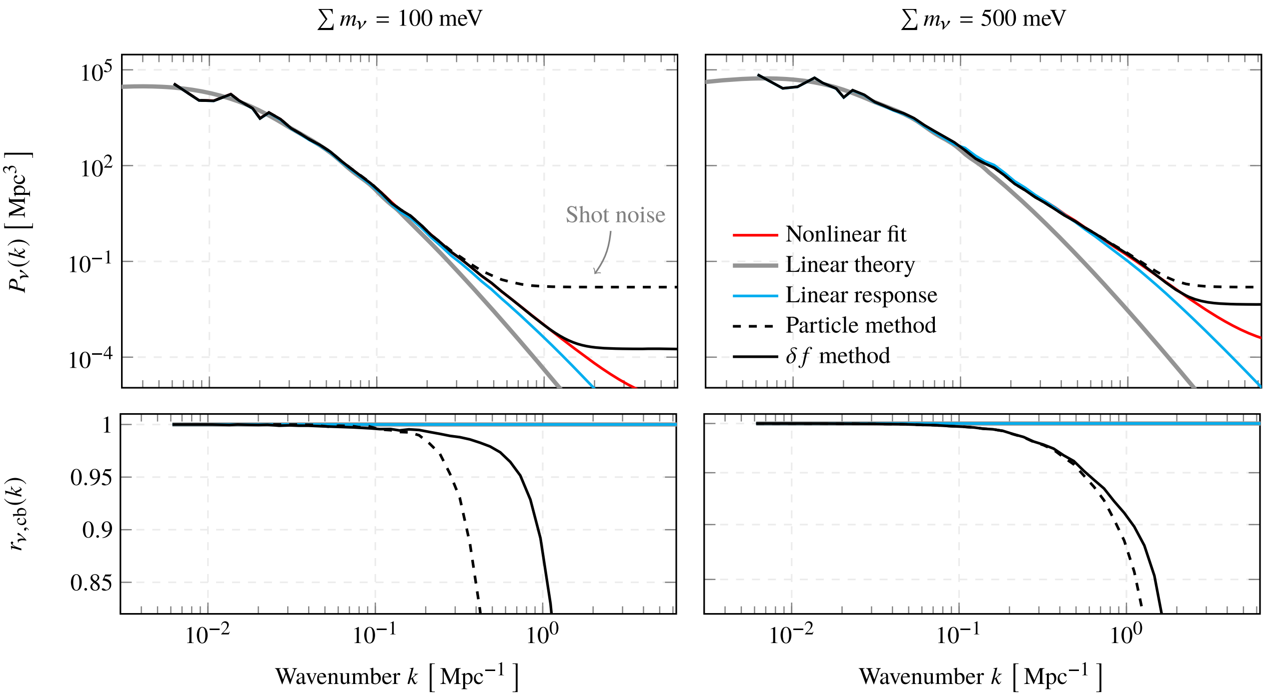

Figure 6. Neutrino power spectra at $z = 0$. We compare the δf method with three commonly used alternatives. Shot noise enters the power spectrum at the constant level $V/N = 1/64$ for the particle method. We also show a fit to the δf power spectrum (red curves), given by equation (11). The bottom panels show the cross-spectral coefficient $r_{\nu,\mathrm{cb}} = P_{\nu,\mathrm{cb}}/\sqrt{P_\nu P_{\mathrm{cb}}}$.

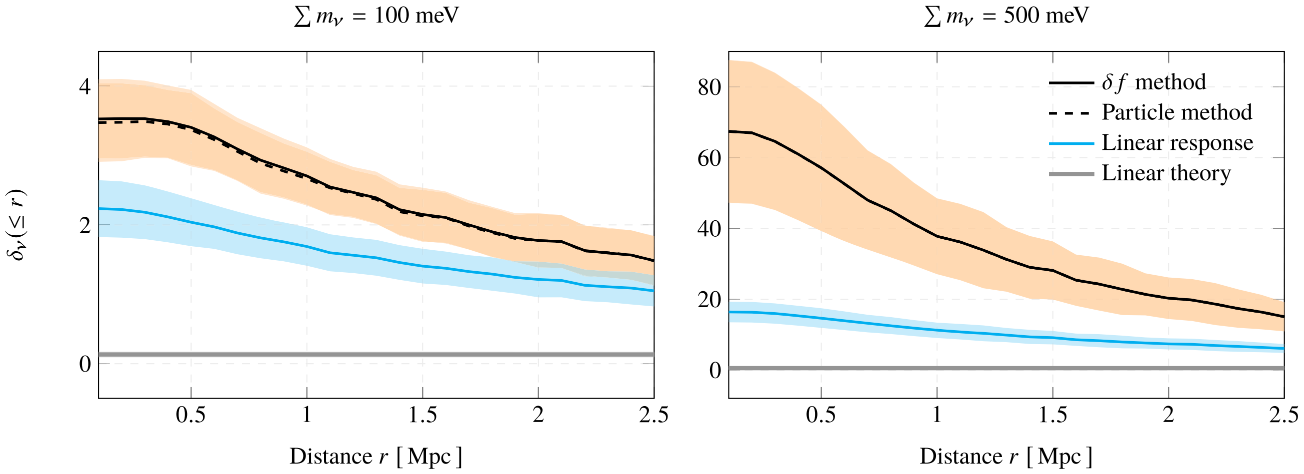

Figure 7. Stacked neutrino density profiles at $z = 0$ for haloes with virial mass $M_{\mathrm{cdm+b}}$ in the range $(5, 14) \times 10^{14}\,\mathrm{M}_\odot$, computed with four different methods from the 256 Mpc simulations. The particle and δf curves overlap almost perfectly. The shaded area indicates the 1σ dispersion around the mean profile.

meV is slightly lower than expected when extrapolating from their results. We also find a larger stochasticity and smaller correlation than might be expected, although the small non-linearities agree. Given the mutual agreement between the different runs in Table 2, these differences are unlikely to be due to our choice of neutrino method. Differences in the N-body code or the identification of haloes could also affect this comparison.

6.3 Matter power spectrum

The suppression of the total matter power spectrum at $z = 0$, relative to a massless neutrino cosmology, is shown in Fig. 9. We see that all methods are in excellent agreement and reproduce the famous spoon-like feature, which has recently been explained in terms of the halo model (Hannestad, Upadhye & Wong 2020). The differences between the methods are most pronounced around $k = 0.6\,\mathrm{Mpc}^{-1}$, where the suppression is largest. The inset graphs zoom in on these scales. For both neutrino masses, the δf method predicts a smaller suppression than the particle and linear methods. This is in line with expectation, as the additional small-scale neutrino clustering, which is obscured by shot noise in the particle method and absent in linear theory, slightly offsets the suppression. Accordingly, the pure linear theory method predicts the least neutrino clustering and the largest suppression. It is interesting to see that the particle and δf methods

Table 2. Neutrino bias relative to dark matter haloes on scales $R = 30\ h^{-1}$ Mpc. Listed are the bias, b; non-linearity, ϵ^2_{nl}; stochasticity, $\epsilon^2_{\mathrm{stoch}}$; and the correlation coefficient, $r_{\nu,\mathrm{h}}$.

	Method	b	ϵ^2_{nl}	$\epsilon^2_{\mathrm{stoch}}$	$r_{\nu,\mathrm{h}}$
100 meV	δf method	0.1032	0.0022	0.4883	0.8195
	Particle method	0.1028	0.0021	0.4955	0.8176
	Linear response	0.1015	0.0022	0.5065	0.8146
	Linear theory	0.0987	0.0206	0.5878	0.7889
500 meV	δf method	0.2556	0.0014	0.1969	0.9137
	Particle method	0.2546	0.0017	0.1927	0.9152
	Linear response	0.2502	0.0019	0.2031	0.9112
	Linear theory	0.2404	0.0257	0.2902	0.8719

do not agree for $\Sigma m_\nu = 500$ meV, despite having similar neutrino power spectra at $z = 0$. This is most likely due to shot noise at high redshift in the particle simulation. Compared to the cold dark matter and baryon fluctuations, the shot noise itself is negligible at $z = 0$, but it has two possible effects on structure formation at earlier times. It could seed non-physical density fluctuations or the random motions of neutrinos could obscure their real contribution to the growth of physical structure. Our results suggest that the latter effect dominates. In either case, these effects highlight the importance of using a hybrid method that eliminates shot noise at high redshift. The differences between the methods are at the per mille level, corresponding to a shift in neutrino mass of several meV. In absolute terms, the differences are larger for $\sum m_\nu = 500$ meV, but less important overall.

The horizontal line corresponds to the empirical fitting formula, $\Delta P/P = -9.8 f_\nu$ (Brandbyge et al. 2008). Compared to this formula, we find a slightly greater suppression in each case, regardless of the method used to model the neutrinos. For the 100 meV simulations, this can be attributed to our use of the inverted mass ordering. The $\sum m_\nu = 500$ meV case is perhaps more surprising, but seems to be in line with recent works. For example, Partmann et al. (2020) find increasingly larger differences with the fitting formula for increasing masses, although they do not consider models with $\sum m_\nu > 300$ meV.

Globally, the agreement between these very different methods is an encouraging sign and suggests that we have a good handle on the effects of massive neutrinos on the matter power spectrum. The differences, at most a few per mille, may perhaps be relevant when trying to distinguish the effects of individual neutrino masses (Wagner, Verde & Jimenez 2012).

7 HIGHER ORDER δf METHODS

The performance of the δf method scales with the correlation between the non-linear solution $f(x, p, t)$ and the background model $\bar{f}(x, p, t)$, so it is worth investigating other background models. We can go beyond the zeroth-order Fermi–Dirac model by including the linear theory prediction. In that case, the distribution function can be written as

$$\bar{f}(x, p, t) = f_{\mathrm{FD}}(x, p, t)\,[1 + \Psi(x, p, t)]\,, \tag{12}$$

where the perturbation is decomposed into multipole moments (Ma & Bertschinger 1995),

$$\Psi(k, p, t) = \sum_{\ell=0}^{\infty} (-1)^\ell (2\ell + 1) \Psi_\ell(t, k, q) P_\ell(\hat{k} \cdot \hat{n}).$$

Here, $P_\ell(\cdot)$ are Legendre polynomials and the coefficients Ψ_ℓ satisfy an infinite hierarchy of moment equations. The Legendre representation yields simple expressions for the first few fluid moments, but is cumbersome for evaluating the distribution function itself. For our purposes, it is more convenient to use the following monomial representation:

$$\Psi(\mathbf{k}, \hat{n}, q, \tau) = \textstyle\sum_{\ell=0}^{\infty} i^\ell \Phi_\ell(\mathbf{k}, q, \tau)(\mathbf{k} \cdot \hat{n})^\ell,$$

where for a given $\ell_{\max}$, the odd (even) $\Phi_\ell(x, q, t)$ can be expressed in terms of all the odd (even) $\Psi_m(x, q, t)$ with $m \leq \ell$. See Appendix C for details. With this choice of background model, the density integral becomes

$$\rho(x) = \bar{\rho}\,[1 + \delta_\nu(x)] + \textstyle\sum_{i=1}^{N} \sqrt{m^2 + p_i^2}\, w_i\, \delta^{(3)}(x - x_i),$$

with particles weights $w_i = \delta f/f$ and $\bar{f}$ given by (12). Here, $\delta_\nu(x)$ is the linear neutrino overdensity, which is calculated using CLASS. The effect of the δ_ν perturbation should now be included in the long-range force calculation.

As shown in Fig. 10, adding the multipoles Φ_0 and Φ_1 significantly improves the correlation and therefore reduces the shot noise by almost 50 per cent. It is likely that higher order terms could contribute meaningfully too, as the multipole expansion converges only slowly. However, most of the gain is due to the zeroth-order term, which on its own is much easier to implement.

8 DISCUSSION AND CONCLUSIONS

Shot noise in N-body simulations is a major obstacle to modelling the non-linear evolution of light relic neutrinos. In this paper, we demonstrate that the δf method, which decomposes the neutrino distribution into an analytically tractable background component, $\bar{f}$, and a non-linear perturbation, δf, carried by the simulation particles, is an effective variance reduction technique. The reduction in shot noise scales with the dynamic particle weights, parametrized by $I = \frac{1}{2}\left\langle w^2 \right\rangle$. Because the weights are negligible until very late times, the simulation is mostly immunized against the effects of shot noise. Furthermore, shot noise is greatly reduced even at $z = 0$, which makes it possible to resolve neutrino clustering down to much smaller scales than is possible with conventional methods. Using higher order versions of the δf method, which incorporate additional information from perturbation theory, shot noise can be reduced by another factor of $\mathcal{O}(2)$, and possibly more if moments $\ell > 2$ are included. Additional reduction in shot noise is possible by carefully tuning the sampling distribution of the marker particles.

The reduction in shot noise is more significant for smaller neutrino masses, because faster particles deviate less from their initial trajectory, resulting in smaller weights. This is fortunate as shot noise is most problematic for the fastest neutrinos. More generally, particles whose trajectories are not perturbed have negligible weights, whereas particles that are captured by haloes have appreciable weights. This is again fortunate, because particles are needed in the vicinity of haloes where grid methods tend to fail, while the unperturbed particles contain no information and contribute only noise. In between these extremes, particles will have intermediate weights. In this way, the δf method ensures an optimal combination of particles and background.

The method can in principle be combined with any grid or fluid background model to obtain an optimal hybrid method. Any simulation that evolves neutrino particles can be extended with a weighting step to minimize the shot noise as outlined in Section 2.3. It is not necessary, as was done here, to evolve the neutrino particles from the beginning. The I-statistic from a reference simulation can

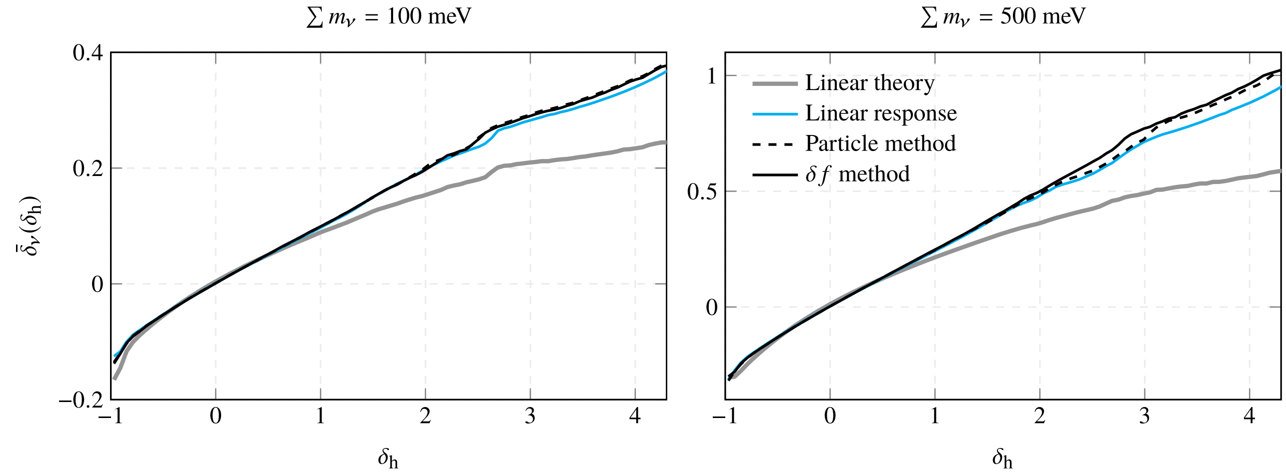

Figure 8. Neutrino bias relative to dark matter haloes with virial mass, $M_\mathrm{cdm+b} > 10^{12}$ M$_\odot$, on scales, $R = 30\,h^{-1}$ Mpc, computed with four different methods from the Gpc simulations at $z = 0$. The $\bar{\delta}_\nu(\delta_\mathrm{h})$ relationship is approximately linear with slope equal to the neutrino bias b.

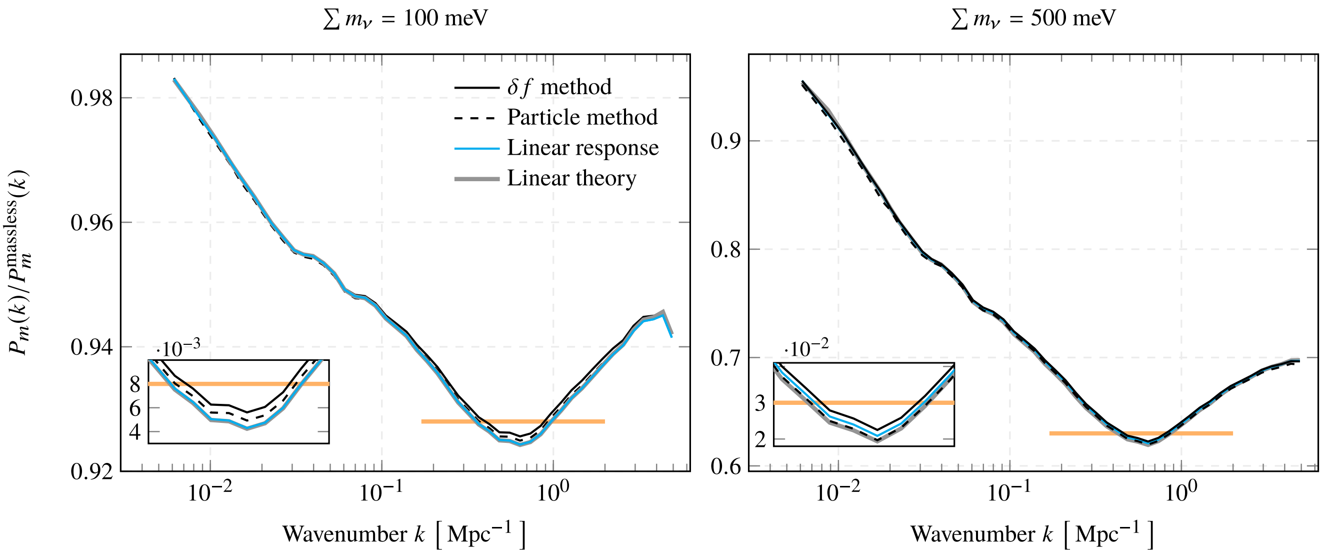

Figure 9. Total matter power spectra at $z = 0$, relative to a massless neutrino cosmology. The plots are based on $(1024\,\mathrm{Mpc})^3$ simulations with $N_\mathrm{cb} = 1024^3$ and (for the particle and δf methods) $N_\nu = 1024^3$ simulation particles. The horizontal line is the empirical fitting formula, $\Delta P/P = -9.8 f_\nu$.

be used to gauge when the neutrinos become non-linear and at what point they can safely be introduced (see Fig. 1).

We know from neutrino oscillations that at least one neutrino has a mass $m_\nu \gtrsim 0.05$ eV. Our results indicate that even for masses close to that bound, neutrinos are not particularly well modelled by linear approximations. For instance, the linear response neutrino power spectrum is off by 10 per cent (60 per cent) at $k = 0.1$ Mpc^{-1} ($k = 1$ Mpc^{-1}) at $z = 0$, and the pure linear theory prediction is off by 14 per cent (96 per cent). Because the neutrinos make up only a small fraction of the total mass, the effect on the matter power spectrum is at most a few per mille. This is the level at which the mass splittings are important (Wagner et al. 2012). Other statistics may be affected at a greater level, particularly if they are more sensitive to neutrino effects. For example, we have shown that the neutrino bias relative to dark matter haloes is affected at the per cent level on 30 h^{-1} Mpc scales. In addition, some novel probes may require accurate modelling of the neutrino dynamics around haloes, such as the neutrino-induced dynamical friction (Okoli et al. 2017) and torque (Yu, Pen & Wang 2019) on haloes. By reducing shot noise without neglecting non-linear terms, the δf method makes it feasible to calculate these effects even for the lightest neutrinos.

ACKNOWLEDGEMENTS

We thank the referee, David Weinberg, for insightful comments that helped improve the manuscript. We also thank Volker Springel and César Hernández-Aguayo for useful discussions on the relativistic aspects of particle simulations. We acknowledge support from the European Research Council through ERC Advanced Investigator grant DMIDAS (GA 786910) to CSF. WE was supported by the ICC PhD Scholarship. SP acknowledges partial support from the European Research Council under ERC Grant NuMass (FP7-IDEAS-

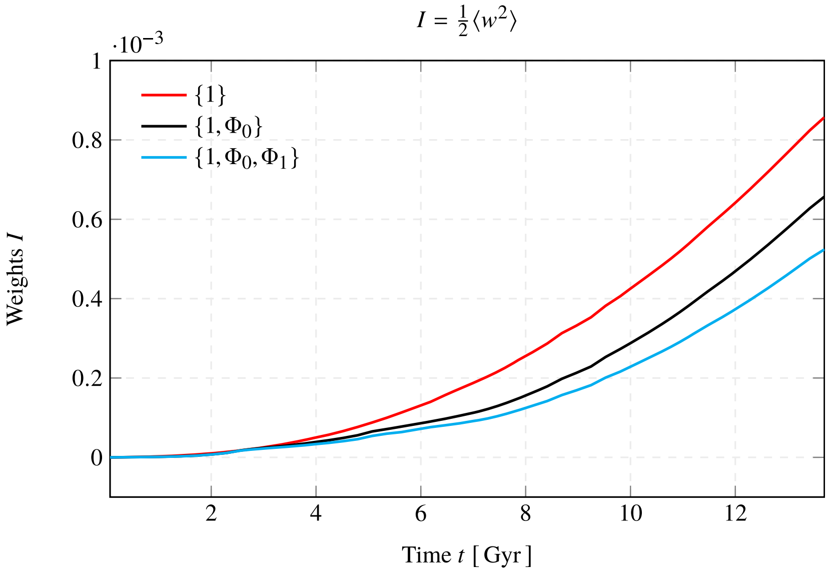

Figure 10. Evolution of the weights, or the effective reduction in shot noise, parametrized by the I-factor $I = \frac{1}{2}\langle w^2 \rangle$, when including higher order perturbations: Φ_0 (density) and Φ_1 (energy flux).

ERC ERC-CG 617143), and the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement numbers 690575 (RISE InvisiblesPlus) and 674896 (ITN Elusives). BL was supported by the European Research Council (ERC) through ERC starting grant number 716532 and STFC consolidated grant (grant numbers ST/I00162X/1 and ST/P000541/1). This work was also supported by STFC consolidated grants for astronomy at Durham (grant numbers ST/P000541/1 and ST/T000244/1). This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, and ST/R002371/1, Durham University, and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

DATA AVAILABILITY

The data analysed in this article can be made available upon reasonable request to the corresponding author. The δf method has been implemented in the open source N-body code SWIFT, which is available at http://swift.dur.ac.uk/.

REFERENCES

Abe K. et al., 2011, preprint (arXiv:1109.3262)
Acciarri R. et al., 2015, preprint (arXiv:1512.06148)
Adamek J., Daverio D., Durrer R., Kunz M., 2016, J. Cosmol. Astropart. Phys., 2016, 053
Adamek J., Durrer R., Kunz M., 2017, J. Cosmol. Astropart. Phys., 2017, 004
Adrian-Martinez S. et al., 2016, J. Phys. G: Nucl. Part. Phys., 43, 084001
Ahmad Q. R. et al., 2002, Phys. Rev. Lett., 89, 011301
Aker M. et al., 2021, preprint (arXiv:2105.08533)
Ali-Haïmoud Y., Bird S., 2012, MNRAS, 428, 3375
An F. et al., 2016, J. Phys. G: Nucl. Part. Phys., 43, 030401
Archidiacono M., Hannestad S., 2016, J. Cosmol. Astropart. Phys., 2016, 018
Aydemir A. Y., 1994, Phys. Plasmas, 1, 822
Banerjee A., Dalal N., 2016, J. Cosmol. Astropart. Phys., 2016, 015
Banerjee A., Powell D., Abel T., Villaescusa-Navarro F., 2018, J. Cosmol. Astropart. Phys., 2018, 028
Bilenky S. M., Pascoli S., Petcov S., 2001, Phys. Rev. D, 64, 053010
Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551
Bird S., Ali-Haïmoud Y., Feng Y., Liu J., 2018, MNRAS, 481, 1486
Bond J. R., Efstathiou G., Silk J., 1980, Phys. Rev. Lett., 45, 1980
Brandbyge J., Hannestad S., 2009, J. Cosmol. Astropart. Phys., 2009, 002
Brandbyge J., Hannestad S., 2010, J. Cosmol. Astropart. Phys., 2010, 021
Brandbyge J., Hannestad S., Haugbølle T., Thomsen B., 2008, J. Cosmol. Astropart. Phys., 2008, 020
Brandbyge J., Hannestad S., Haugbølle T., Wong Y. Y., 2010, J. Cosmol. Astropart. Phys., 2010, 014
Brandbyge J., Rampf C., Tram T., Leclercq F., Fidler C., Hannestad S., 2017, MNRAS, 466, L68
Castorina E., Carbone C., Bel J., Sefusatti E., Dolag K., 2015, J. Cosmol. Astropart. Phys., 2015, 043
Chartier N., Wandelt B., Akrami Y., Villaescusa-Navarro F., 2021, MNRAS, 503, 1897
Chen J. Z., Upadhye A., Wong Y. Y., 2021a, J. Cosmol. Astropart. Phys., 2021, 065
Chen J. Z., Upadhye A., Wong Y. Y., 2021b, J. Cosmol. Astropart. Phys., 2021, 078
Choudhury S. R., Hannestad S., 2020, J. Cosmol. Astropart. Phys., 2020, 037
Chudaykin A., Ivanov M. M., 2019, J. Cosmol. Astropart. Phys., 2019, 034
Colas T., d'Amico G., Senatore L., Zhang P., Beutler F., 2020, J. Cosmol. Astropart. Phys., 06, 001
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Dakin J., Brandbyge J., Hannestad S., Haugbølle T., Tram T., 2019, J. Cosmol. Astropart. Phys., 2019, 052
Di Valentino E., Melchiorri A., Silk J., 2020, J. Cosmol. Astropart. Phys., 2020, 013
Dimits A., Lee W. W., 1993, J. Comput. Phys., 107, 309
Drexlin G., Hannen V., Mertens S., Weinheimer C., 2013, Adv. High Energy Phys., 2013, 293986
Eguchi K. et al., 2003, Phys. Rev. Lett., 90, 021802
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019, Publ. Astron. Soc. Aust., 36, e021
Emberson J. et al., 2017, Res. Astron. Astrophys., 17, 085
Esfahani A. A. et al., 2017, J. Phys. G: Nucl. Part. Phys., 44, 054004
Esteban I., Gonzalez-Garcia M., Maltoni M., Schwetz T., Zhou A., 2020, J. High Energy Phys., 2020, 1

Fidler C., Rampf C., Tram T., Crittenden R., Koyama K., Wands D., 2015, Phys. Rev. D, 92, 123517

Fidler C., Tram T., Rampf C., Crittenden R., Koyama K., Wands D., 2017a, J. Cosmol. Astropart. Phys., 2017, 043

Fidler C., Tram T., Rampf C., Crittenden R., Koyama K., Wands D., 2017b, J. Cosmol. Astropart. Phys., 2017, 022

Font-Ribera A., McDonald P., Mostek N., Reid B. A., Seo H.-J., Slosar A., 2014, J. Cosmol. Astropart. Phys., 2014, 023

Frenk C. S., White S. D. M., Davis M., 1983, ApJ, 271, 417

Fukuda Y. et al., 1998, Phys. Rev. Lett., 81, 1562

Giuliani A., Gomez Cadenas J., Pascoli S., Previtali E., Saakyan R., Schäffner K., Schönert S., 2019, preprint (arXiv:1910.04688)

Hannestad S., Haugbølle T., Schultz C., 2012, J. Cosmol. Astropart. Phys., 2012, 045

Hannestad S., Upadhye A., Wong Y. Y., 2020, J. Cosmol. Astropart. Phys., 2020, 062

Hu W., Eisenstein D. J., Tegmark M., 1998, Phys. Rev. Lett., 80, 5255

Inman D., Emberson J., Pen U.-L., Farchi A., Yu H.-R., Harnois-Déraps J., 2015, Phys. Rev. D, 92, 023502

Inman D., Yu H.-R., 2020, ApJS, 250, 21

Ivezic Z. et al., 2009, BAAS, 41, 366

Klypin A., Shandarin S., 1983, MNRAS, 204, 891

Laureijs R. et al., 2011, preprint (arXiv:1110.3193)

Leeuwin F., Combes F., Binney J., 1993, MNRAS, 262, 1013

Lesgourgues J., 2011, preprint (arXiv:1104.2932)

Lesgourgues J., Tram T., 2011, J. Cosmol. Astropart. Phys., 2011, 032

Levi M. et al., 2013, preprint (arXiv:1308.0847)

Lewis A., Challinor A., 2011, Astrophysics Source Code Library, record ascl:1102.026

Liu J., Bird S., Matilla J. M. Z., Hill J. C., Haiman Z., Madhavacheril M. S., Petri A., Spergel D. N., 2018, J. Cosmol. Astropart. Phys., 2018, 049

Ma C.-P., Bertschinger E., 1994a, ApJ, 429, 22

Ma C.-P., Bertschinger E., 1994b, ApJ, 434, L5

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

McCarthy I. G., Bird S., Schaye J., Harnois-Deraps J., Font A. S., Van Waerbeke L., 2018, MNRAS, 476, 2999

Merritt D., 1987, in de Zeeuw T.ed., Symposium-International Astronomical Union, Structure and Dynamics of Elliptical Galaxies. Springer, Dordrecht, Netherlands, p. 315

Mertens S., 2016, J. Phys.: Conf. Ser., 718, 022013

Nunokawa H., Teves W., Zukanovich Funchal R., 2002, Phys. Rev. D, 66, 093010

Okoli C., Scrimgeour M. I., Afshordi N., Hudson M. J., 2017, MNRAS, 468, 2164

Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2020, J. Cosmol. Astropart. Phys., 2020, 038

Parker S., Lee W., 1993, Phys. Fluids B, 5, 77

Partmann C., Fidler C., Rampf C., Hahn O., 2020, J. Cosmol. Astropart. Phys., 2020, 018

Pascoli S., Petcov S., 2002, Phys. Lett. B, 544, 239

Planck Collaboration VI, 2020, A&A, 641, A6

Quinn T., Katz N., Stadel J., Lake G., 1997, ApJ, preprint (arXiv:astro-ph/9710043)

Ross S., 2012, Simulation. Elsevier Science, Amsterdam, Netherlands, p. 147

Schaller M., Gonnet P., Chalk A. B., Draper P. W., 2016, Proceedings of the Platform for Advanced Scientific Computing Conference. Association for Computing Machinery, New York

Schaller M., Gonnet P., Chalk A. B., Draper P. W., 2018, Astrophysics Source Code Library

Senatore L., Zaldarriaga M., 2017, preprint (arXiv:1707.04698)

Sprenger T., Archidiacono M., Brinckmann T., Clesse S., Lesgourgues J., 2019, J. Cosmol. Astropart. Phys., 2019, 047

Tajima T., Perkins F. W., 1983, Sherwood Theory Meeting 1983 Meeting Bulletin, 2P9

Taruya A., Suto Y., 2000, ApJ, 542, 559

Tram T., Brandbyge J., Dakin J., Hannestad S., 2019, J. Cosmol. Astropart. Phys., 2019, 022

Upadhye A., 2019, J. Cosmol. Astropart. Phys., 2019, 041

Vergados J. D., Ejiri H., Šimkovic F., 2016, Int. J. Mod. Phys. E, 25, 1630007

Viel M., Haehnelt M. G., Springel V., 2010, J. Cosmol. Astropart. Phys., 2010, 015

Villaescusa-Navarro F., Bull P., Viel M., 2015, ApJ, 814, 146

Villaescusa-Navarro F., Marulli F., Viel M., Branchini E., Castorina E., Sefusatti E., Saito S., 2014, J. Cosmol. Astropart. Phys., 2014, 011

Villaescusa-Navarro F., Miralda-Escudé J., Peña-Garay C., Quilis V., 2011, J. Cosmol. Astropart. Phys., 2011, 027

Villaescusa-Navarro F. et al., 2020, ApJS, 250, 2

Wagner C., Verde L., Jimenez R., 2012, ApJ, 752, L31

Weisstein E. W., 2002, From MathWorld–A Wolfram Web Resource. Available at: https://mathworld.wolfram.com/JacobiEllipticFunctions.html

Yoshikawa K., Taruya A., Jing Y., Suto Y., 2001, ApJ, 558, 520

Yoshikawa K., Tanaka S., Yoshida N., Saito S., 2020, ApJ, 904, 159

Yu H.-R., Pen U.-L., Wang X., 2019, Phys. Rev. D, 99, 123532

APPENDIX A: ELLIPTICAL SINE WAVE SOLUTION

We define an integral of motion

$$E(x, p, t) = \tfrac{1}{2}p^2 + \sin^2(x/2),$$

which is interpreted as the energy of the particle. Hence,

$$p = \pm\sqrt{2E - 2\sin^2(x/2)}.$$

We have reduced the characteristic equations to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \pm\sqrt{2E - 2\sin^2(x/2)}.$$

This equation is separable,

$$\int \frac{\mathrm{d}x}{\sqrt{2E - 2\sin^2(x/2)}} = \pm\textstyle\int \mathrm{d}t.$$

Let τ be the time when $x(\tau) = 0$. Putting in the integration limits, setting $u = x/2$, and factoring out $2E$, we obtain

$$F(u) \equiv \int_0^u \frac{\mathrm{d}u'}{\sqrt{1 - \sin^2(u')/E}} = \pm\sqrt{E/2}\textstyle\int_\tau^t \mathrm{d}t' \equiv \phi.$$

The elliptic sine function is defined such that $\mathrm{sn}(\phi) = \sin(u)$, where $u = F^{-1}(\phi)$. Hence,

$$\mathrm{sn}\left(\pm\sqrt{E/2}(t-\tau)\right) = \sin u. \tag{A1}$$

There exist the following trigonometric identities (Weisstein 2002):

$$\mathrm{sn}(\phi)^2 + \mathrm{cn}(\phi)^2 = 1 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\phi}\mathrm{sn}(\phi) = \mathrm{cn}(\phi)\mathrm{dn}(\phi),$$

where $\mathrm{cn}(\phi) = \cos(u)$ and $\mathrm{dn}(\phi) = \sqrt{1 - \sin^2(u)/E}$ are the elliptic cosine and delta amplitude functions. Using these identities, one can confirm the solution (A1). To find the phase-space distribution at time t, we use the fact that $f(x, p, t)$ is constant along its characteristic curves. At $t = 0$, let $x_0 = x(0)$ and $p_0 = p(0)$. Using the initial Gaussian distribution (7), we find

$$\begin{aligned} f(x, p, t) &= f(x_0, p_0, 0) \\ &= \frac{\bar{\rho}}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(p_0 - \bar{p})^2}{2\sigma^2}\right). \end{aligned}$$

We need to express p_0 in terms of x, p, and t. First, we use conservation of energy to note that

$$p_0^2 = p^2 + 2\sin^2(x/2) - 2\sin^2(x_0/2).$$

What remains to show is how to express $\sin^2(x_0/2)$ in terms of x, p, and t. But this is simply,

$$\sin^2(x_0/2) = \mathrm{sn}^2\left(\mp\sqrt{\tfrac{1}{4}p^2 + \tfrac{1}{2}\sin^2(x/2)\tau}\right),$$

where the time, τ, is given by

$$\tau = \mp\sqrt{2/E}\,\mathrm{arcsn}\,(\sin(x/2)) + t.$$

Here, we used the inverse of the elliptic sine function, $\mathrm{arcsn}(x) = \phi$, with $x = \mathrm{sn}(\phi)$. It follows that

$$\sin^2(x_0/2) = \mathrm{sn}^2\left(\mathrm{arcsn}\,(\sin(x/2)) \mp \sqrt{\tfrac{1}{4}p^2 + \tfrac{1}{2}\sin^2(x/2)t}\right).$$

Therefore, the distribution function is

$$\begin{aligned} f(x, p, t) &= \frac{\bar{\rho}}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left[h(x, p, t) - 2\bar{p}\sqrt{h(x, p, t)} + \bar{p}^2\right]\right), \\ h(x, p, t) &= k(x, p, 1) - m(x, p, t) \\ k(x, p, t) &= p^2 + 2\sin^2(x/2)t \\ m(x, p, t) &= 2\mathrm{sn}^2\left(\mathrm{arcsn}\,(\sin(x/2)) \mp \sqrt{\tfrac{1}{4}k(x, p, t)}\right). \end{aligned}$$

To find the density profile, $\rho(x, t)$, we integrate

$$\rho(x, t) = \textstyle\int_{-\infty}^{\infty} f(x, p, t)\mathrm{d}p,$$

which can be done numerically. This gives the solution curves in Fig. 2.

APPENDIX B: ACCURATE CALCULATION OF NEUTRINO MOMENTS

For completeness, we outline the linear theory calculation of the neutrino distribution function following Ma & Bertschinger (1995). The neutrino phase-space distribution function can be written as

$$f(\mathbf{x}, \mathbf{q}, \tau) = \bar{f}(q)\left[1 + \Psi(\mathbf{x}, q, \hat{n}, \tau)\right], \tag{B1}$$

where $\bar{f}(q)$ is the homogeneous Fermi–Dirac distribution and the momentum $\mathbf{q}$ has been decomposed into a magnitude q and a unit vector $\hat{n}$. We will express the momentum q and energy $\epsilon = \sqrt{q^2 + a^2m^2}$ in units of k_bT. In synchronous gauge, the perturbation Ψ evolves as

$$\dot{\Psi} + i\frac{qc}{\epsilon}(\mathbf{k}\cdot\hat{n})\Psi + \frac{1}{c^2}\frac{\mathrm{d}\ln\bar{f}}{\mathrm{d}\ln q}\left[\dot{\eta} - \frac{\dot{h} + 6\dot{\eta}}{2}(\hat{k}\cdot\hat{n})^2\right] = 0,$$

where we switched to momentum space and where overdots denote conformal time derivatives. Here, h and η are the scalar metric perturbations in synchronous gauge. To solve this equation, Ψ is decomposed into a Legendre series

$$\Psi(\mathbf{k}, \hat{n}, q, \tau) = \sum_{\ell=0}^{\infty}(-i)^{\ell}(2\ell + 1)\Psi_{\ell}(\mathbf{k}, q, \tau)P_{\ell}(\hat{k}\cdot\hat{n}). \tag{B2}$$

The Boltzmann equation then becomes an infinite tower of equations:

$$\dot{\Psi}_0 = -\frac{qkc}{\epsilon}\Psi_1 + \frac{\dot{h}}{6}\frac{1}{c^2}\frac{\mathrm{d}\ln\bar{f}}{\mathrm{d}\ln q}, \tag{B3}$$

$$\dot{\Psi}_1 = \frac{qkc}{3\epsilon}(\Psi_0 - 2\Psi_2), \tag{B4}$$

$$\dot{\Psi}_2 = \frac{qkc}{5\epsilon}(2\Psi_1 - 3\Psi_3) - \left(\frac{\dot{h}}{15} + \frac{2\dot{\eta}}{5}\right)\frac{1}{c^2}\frac{\mathrm{d}\ln\bar{f}}{\mathrm{d}\ln q}, \tag{B5}$$

$$\dot{\Psi}_{\ell} = \frac{qkc}{(2\ell + 1)\epsilon}\left[\ell\Psi_{\ell-1} - (\ell + 1)\Psi_{\ell+1}\right], \qquad \text{for} \quad \ell \geq 3. \tag{B6}$$

The hierarchy can be truncated at $\ell = \ell_{\max} + 1$ using the ansatz

$$\Psi_{\ell_{\max}+1} = \frac{(2\ell_{\max} + 1)\epsilon}{qkc\tau}\Psi_{\ell_{\max}} - \Psi_{\ell_{\max}-1}.$$

The precision of this calculation is set by two parameters: the maximum multipole $\ell_{\max}$ and the number of momentum bins N_q. By default, CLASS partly relies on a set of fluid equations and partly on integrating the hierarchy, using $\ell_{\max} = 50$ and $N_q = 28$ at the pre-set high precision settings (Lesgourgues & Tram 2011). The differences in the CMB anisotropies are at the per mille level. However, the neutrino transfer functions have still not converged. To obtain converged results, Dakin et al. (2019) ran calculations with $N_q = 2000$ bins and $\ell_{\max} = 2000$, which each required hundreds of CPU hours. This is to be contrasted with a default CLASS run, which completes in seconds. To circumvent this computational cost, we use a different approach, which involves a post-processing step of CLASS tables.

To quickly integrate the Boltzmann hierarchy for high N_q and $\ell_{\max}$, we note that the source terms in the evolution equations depend on the matter content only through the scalar potential derivatives $\dot{h}$ and $\dot{\eta}$, which can be calculated accurately with much lower settings.[5] Therefore, we make the assumption that we can decouple the potential terms from most of the neutrino moments Ψ_{ℓ}. We first evolve all source functions in CLASS at a reasonable precision setting. This gives the metric perturbations $\dot{h}(k, \tau)$ and $\dot{\eta}(k, \tau)$, which we then take as given and use to integrate the multipole moments Ψ_{ℓ} at high precision where they are needed.

APPENDIX C: MONOMIAL BASIS FOR THE DISTRIBUTION FUNCTION

Boltzmann codes can solve for the functions $\Psi_{\ell}(k, q, \tau)$. But evaluating the distribution function, $f(\mathbf{x}, \mathbf{q}, \tau)$, requires substituting these back into the definitions (B1) and (B2). This presents a challenge as the Ψ_{ℓ} are large discretely sampled arrays of amplitudes that need to be convolved with the random phases. It would be prohibitively expensive to do this repeatedly for each term in the Legendre expansion. We therefore adopt the following scheme. First, we use the following representation of the ℓth Legendre polynomial,

$$P_{\ell}(x) = 2^{\ell}\sum_{n=0}^{\ell} x^n \binom{\ell}{n}\binom{\frac{n+\ell-1}{2}}{\ell},$$

where the last factor is a generalized binomial coefficient. This allows us to expand Ψ and collect monomial terms in $\hat{k}\cdot\hat{n}$. We write

$$\Psi(\mathbf{k}, \hat{n}, q, \tau) = \textstyle\sum_{\ell=0}^{\infty} i^{\ell}\Phi_{\ell}(\mathbf{k}, q, \tau)(\mathbf{k}\cdot\hat{n})^{\ell},$$

where the functions Φ_{ℓ} are defined by

$$\Phi_{\ell}(\mathbf{k}, q, \tau) = \frac{1}{k^{\ell}}\sum_{n=0}^{\infty}(-2)^n\binom{n}{\ell}\binom{\frac{n+\ell-1}{2}}{n}(2n + 1)\Psi_n(\mathbf{k}, q, \tau).$$

Note that we factored out the magnitude of $\mathbf{k} = k\hat{k}$ and write the expansion in terms of $(\mathbf{k}\cdot\hat{n})^{\ell}$ and not $(\hat{k}\cdot\hat{n})^{\ell}$. This is to facilitate taking derivatives, as shown below. The Fourier transform of Ψ is

$$\Psi(\mathbf{x}, \hat{n}, q, \tau) = \int\frac{\mathrm{d}^3k}{(2\pi)^3}\Psi(\mathbf{k}, \hat{n}, q, \tau)e^{i\mathbf{x}\cdot\mathbf{k}},$$

and similarly for the Ψ_{ℓ} and Φ_{ℓ}. We write the directional derivative along the unit vector $\hat{n}$ as $\mathcal{D}_{\hat{n}} = n^i\partial_{x_i}$. In other words,

$$\mathcal{F}\left\{\mathcal{D}_{\hat{n}}\Phi_{\ell}(\mathbf{x}, \hat{n}, q, \tau)\right\} \Longleftrightarrow i(\mathbf{k}\cdot\hat{n})\Phi_{\ell}(\mathbf{k}, q, \tau).$$

---

[5] In the reference model with $N_q = 28$ and $\ell_{\max} = 50$, relative errors in $\dot{h}$ and $(\dot{h}/3 + 2\dot{\eta})$ are of the order of 10^{-4}. Although $\dot{\eta}$ still fluctuates at the several per cent level, this term is much smaller than $\dot{h}$.

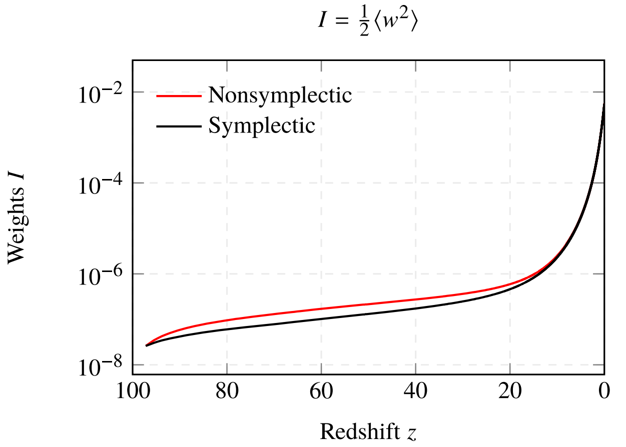

Figure D1. Evolution of particle weights for a $\sum m_\nu = 150$ meV cosmology, using special relativistic equations of motion, evolved with symplectic and non-symplectic leapfrog integrators.

Hence, we obtain

$$i^\ell \Phi_\ell(\mathbf{k}, q, \tau)(\mathbf{k} \cdot \hat{n})^\ell \iff \mathcal{F}\left\{\mathcal{D}_{\hat{n}}^\ell \Phi_\ell(\mathbf{x}, q, \tau)\right\}.$$

And so, the overall perturbation, Ψ, is

$$\Psi(\mathbf{x}, \hat{n}, q, \tau) = \sum_{\ell=0}^{\infty} \mathcal{D}_{\hat{n}}^\ell \Phi_\ell(\mathbf{x}, q, \tau).$$

A convenient numerical scheme is to store the Fourier transformed grids $\Phi_\ell(\mathbf{x}, q, \tau)$, in which case we can evaluate the distribution function efficiently by taking finite differences.

APPENDIX D: SYMPLECTIC INTEGRATOR

A point of concern is that the special relativistic equation of motion (8) may not be suited for the usual leapfrog integration in cosmological *N*-body simulations (Quinn et al. 1997). A straightforward substitution for the non-relativistic equation (10) produces a leapfrog integrator that is not symplectic and may therefore not explicitly conserve phase-space volume. The problem is that the equation for $\dot{\mathbf{u}}$ depends on $\mathbf{u}$. Here, we assess the impact of this error and provide a workaround. Fortunately, it is easy to construct a symplectic integrator that closely approximates the relativistic form. We simply replace equation (8) with

$$\dot{\mathbf{u}} = -\frac{2u_0^2 + a^2}{\sqrt{u_0^2 + a^2}} \nabla\left(\phi - \gamma^{\mathrm{Nb}}\right), \tag{D1}$$

where u_0 is the magnitude of $\mathbf{u}$ at the starting redshift of the simulation. Moreover, we use $\epsilon = m\sqrt{u_0^2 + a^2}$ when computing the energy density. As confirmed below, this is a good approximation due to the fact that $u_0 \ll a$ whenever u deviates much from u_0, i.e. for slow particles at late times. Equation (9) is unchanged. A leapfrog discretization of these equations is

$$\begin{aligned}
\mathbf{x}_{k+1/2} &= \mathbf{x}_k + \frac{1}{2}\Delta t \frac{\mathbf{u}_k}{\sqrt{u_k^2 + a^2}}, \\
\mathbf{u}_{k+1} &= \mathbf{u}_k - \Delta t \frac{2u_0^2 + a^2}{\sqrt{u_0^2 + a^2}} \nabla\left(\phi(\mathbf{x}_{k+1/2}) - \gamma^{\mathrm{Nb}}(\mathbf{x}_{k+1/2})\right), \\
\mathbf{x}_{k+1} &= \mathbf{x}_{k+1/2} + \frac{1}{2}\Delta t \frac{\mathbf{u}_{k+1}}{\sqrt{u_{k+1}^2 + a^2}}.
\end{aligned}$$

To determine whether this is symplectic, one considers the Jacobian, J, of the transformation $\psi : \mathbf{z}_k \mapsto \mathbf{z}_{k+1}$, where $\mathbf{z}_k = (\mathbf{x}_k, \mathbf{u}_k)$. One can confirm that

$$J^{\mathrm{T}}\Omega J = \Omega \equiv \begin{pmatrix} 0 & I_3 \\ -I_3 & 0 \end{pmatrix}$$

to show that ψ is a symplectomorphism. It follows that $\det(J) = 1$, which ensures that the leapfrog integrator is volume-preserving.

Since the validity of the δf algorithm depends on conservation of phase-space density along particle trajectories, we need to determine to what extent this is violated when using a non-symplectic discretization of (8). Deviations will be of the order of $\mathcal{O}(u^2)$, which is small for any non-trivial neutrino orbit. Therefore, the difference should be negligible when the weights are large. To test this assertion, we evolve 256^3 neutrino and dark matter particles in a volume with sidelength 256 Mpc and assuming $\sum m_\nu = 3 \times 50$ meV, using both the special relativistic equations (8) and the alternative equation (D1). First, we confirm that (D1) is a good approximation of (8), by checking that the ratios

$$r = \sqrt{\frac{u_0^2 + a^2}{u^2 + a^2}}, \qquad \text{and} \qquad s = \frac{1}{r}\frac{2u_0^2 + a^2}{2u^2 + a^2}$$

are close to unity. We find that $|r - 1| < 0.02$ per cent and $|s - 1| < 0.03$ per cent for 99 per cent of particles at all times.

The evolution of the weights is shown in Fig. D1. At very high redshifts, when $I = \frac{1}{2}\langle w^2 \rangle < 10^{-6}$, the mean squared weight is about 60 per cent larger for the non-symplectic integrator. We interpret this as being due to small perturbations to neutrino trajectories that are absent in the symplectic case. The results converge after $z = 20$, when density perturbations approach non-linearity, driving up the weights. The difference in I is 7 per cent at $z = 10$, decreasing to 0.2 per cent at $z = 2$, and 0.03 per cent at $z = 0$. The difference in weights $|w_{\mathrm{ns}} - w_{\mathrm{s}}| < 0.025$ for 99 per cent of particles at all times. The difference will be even smaller for larger neutrino masses.

As a result, we find that the use of the non-symplectic integrator has a negligible effect on observables at late times. In particular, there is < 0.1 per cent difference in the total matter and neutrino power spectra at $z = 0$. The difference in $P_\nu(k)$ grows to 0.3 per cent at $z = 2$, and 2 per cent at $z = 10$. For the total matter power spectrum, the difference is always below 0.1 per cent. For future simulations, we recommend using a leapfrog discretization based on the modified expression (D1) or using the non-relativistic version of (8) together with (9), as discussed in Section 4.3.

APPENDIX E: PRACTICAL IMPLEMENTATION

We briefly outline how to implement the δf method in a typical *N*-body code. First, a choice needs to be made for the background model. The simplest choice is the homogeneous Fermi–Dirac distribution, which we repeat here without pre-factor:

$$\bar{f}(x, p, t) = \bar{f}(p, a) = \left[\exp\left(\frac{ap}{k_b T_\nu}\right) + 1\right]^{-1}, \tag{E1}$$

where $T_\nu = 1.95$ K is the present-day neutrino temperature, a is the scale factor, and p is the 3-momentum. The method can be implemented as follows:

(i) Implement a function $\bar{f}(x, p, t)$ that returns the phase-space density at a particle's location according to the background model.

(ii) Generate initial conditions with neutrino particles having a random initial momentum sampled from the background model. Perturbations can be applied afterwards.

(iii) For each particle, record the value of $\bar{f}(x, p, t) = f_0$ at the initial sampled position in phase space.[6] The numerical value of f_0 does not change if perturbations are applied to the initial conditions.

(iv) During subsequent time-steps, for each particle:

  (a) Compute $\bar{f}(x, p, t)$ using the new position and momentum.

  (b) Compute the weight $w = (f_0 - \bar{f})/f_0$.

  (c) For the purposes of the gravity force calculation, use the weighted mass mw.

The weights should be used when calculating statistics of the neutrino ensemble, such as the neutrino density, $\rho_\nu(x)$, that enters into the gravity force calculation. We emphasize that the weights should not be used in relationships such as $\mathbf{p} = m\mathbf{u}$.

For tree codes that perform a multipole expansion around the centre of mass, like SWIFT, there is a final point of consideration. Since neutrino particles can have negative weights, it is possible in rare circumstances that groups of particles have a nearly vanishing total mass. In that case, the centre of mass can be far removed from the particles and the multipole expansion breaks down. There is a simple solution in such cases, which is to expand around any other point such as the geometric centre of the particles or the centre of absolute mass:

$$\mathbf{x}_{\mathrm{CoAM}} = \frac{\sum_i |m_i w_i| \mathbf{x}_i}{\sum_i |m_i w_i|},$$

which has the advantage that it is very close to the ordinary centre of mass in most cases. When the background model agrees with the non-linear solution, the weights are exactly zero. In that case, they can be set to a small value $w \ll 1$. This ensures that the centre of absolute mass is always well defined.

---

[6] Because the background model was used for sampling the initial momenta, we initially have $\bar{f} = f = g$. Conservation of phase-space density then ensures that $f = g = f_0$ for all particles at all times.

This paper has been typeset from a TEX/LATEX file prepared by the author.