

# Efficient analytic approximation for small-scale non-cold relic perturbations

Nanoom Lee,<sup>1,\*</sup> Yacine Ali-Haïmoud,<sup>2,†</sup> and Marc Kamionkowski<sup>1,‡</sup>

<sup>1</sup>*William H. Miller III Department of Physics & Astronomy,  
Johns Hopkins University, Baltimore, Maryland 21218, USA*

<sup>2</sup>*Center for Cosmology and Particle Physics, Department of Physics,  
New York University, New York, New York 10003, USA*

We develop a highly accurate analytic approximation for small-scale non-cold relic perturbations by solving the collisionless Boltzmann equation in the quasi-stationary regime. The approximation is implemented in **CLASSIER** (**CLASS** Integral Equation Revision), a modified version of the Boltzmann solver **CLASS** that replaces the traditional truncated Boltzmann hierarchy of non cold relic multipoles with a small set of integral equations solved iteratively. Applying it to massive neutrinos yields a factor-of-two reduction in total runtime relative to **CLASSIER** without the approximation. Compared to standard **CLASS** runs (with  $\ell_{\text{max}}^{\text{NCDM}} = 40$  and no late-time massive neutrino fluid approximation) under the same precision setting, **CLASSIER** with this approximation is faster by a factor of 3–6. The approximation faithfully reproduces the late-time behavior of massive neutrino perturbations and preserves sub-0.1% accuracy in the matter power spectrum today up to comoving wavenumber  $k = 100 \text{ Mpc}^{-1}$ . With this approximation, massive-neutrino perturbations are no longer the computational bottleneck on small scales for linear-theory predictions. The approach can be readily extendable to non-standard dark-matter models, and offers prospects for further efficiency gains in high-precision cosmological analyses.

## I. INTRODUCTION

Cosmological observations across multiple probes from cosmic microwave background (CMB) anisotropies to the clustering of galaxies are increasingly sensitive to the properties of relic particles beyond standard cold dark matter. Non-cold dark matter (NCDM) or non-cold relics, such as massive neutrinos or other light thermal relics, free-stream on cosmological scales and imprint characteristic signatures on the growth of structure [1–3]. The suppression of power on small scales, the scale-dependent growth of matter perturbations, and the induced anisotropic stress in the early Universe all make NCDM a key target for current and next-generation cosmological surveys [4, 5].

The standard approach to modeling the evolution of NCDM perturbations is to solve the Boltzmann hierarchy: a formally infinite set of coupled differential equations for the multipole moments of the phase-space perturbation [6]. In practice, this hierarchy must be truncated at some finite multipole  $\ell_{\text{max}}$ , with closure relations imposed at the truncation point. While accurate, this procedure can become computationally expensive at high precision or on small scales, since a large number of multipoles must be evolved. In addition, if the hierarchy is truncated too low, small numerical artifacts may appear, which can propagate into predictions for cosmological observables [7]. To mitigate these challenges, model-specific approximations have also been developed [8–11], most notably for massive neutrinos, which are approximated as a viscous fluid once free-streaming suppresses

higher-order multipoles. These approximations substantially reduce computational cost while retaining accuracy on large scales, but they are reliable only for the range of wavenumbers and model parameters that they are calibrated upon, and have a precision no better than the calibration calculations. More recently, improved fluid approximations have been proposed that incorporate a scale-dependent sound speed and anisotropic stress, further extending the regime of validity for massive neutrinos [12]. Nonetheless, high-accuracy calculations of NCDM perturbations still demand evolving a large number of multipoles, which remains a heavy computational expense in modern cosmological analyses. The situation can be even more demanding in certain non-standard dark matter models, where accurately capturing the perturbations requires dense momentum grids in addition to many multipoles, further increasing the computational cost (e.g., see Ref. [13]).

Recently, alternative formulations based on integral equations have been developed as a way to overcome these limitations [14–16]. Instead of evolving the entire Boltzmann hierarchy, the integral approach follows particles along unperturbed geodesics and expresses the perturbations as time-integrals over metric sources convolved with spherical Bessel kernels. This method is similar in spirit to the line-of-sight approach for CMB anisotropies [17], except that it is used not just to compute final observables, but throughout the evolution of perturbations themselves. An identical integral approach has also been used to follow the evolution of massive neutrinos during non-linear structure formation [18, 19]. This representation naturally captures the physics of free-streaming, and its convolutional structure allows for efficient numerical evaluation using fast Fourier transforms (FFTs). The **CLASSIER** code (**CLASS** Integral Equation

\* [nanoom.lee@jhu.edu](mailto:nanoom.lee@jhu.edu)

† [yah2@nyu.edu](mailto:yah2@nyu.edu)

‡ [kamion@jhu.edu](mailto:kamion@jhu.edu)

Revision)<sup>1</sup> [16] implements this method within the public Boltzmann solver CLASS v3.2.2 [7, 11], achieving accurate and efficient evolution of NCDM perturbations.

Despite this progress, evaluating the full integral solutions can remain costly at large wavenumbers, where perturbations oscillate rapidly due to free-streaming. In this regime, analytic approximations are both physically transparent and computationally advantageous: they capture the asymptotic behavior of perturbations while reducing the need for expensive numerical convolutions and preserving the accuracy of the observables of interest.

In this work, we develop such a small-scale approximation for NCDM perturbations, in the form of a perturbative expansion in the ratio of the perturbation wavelength to the free-streaming lengthscale. We emphasize that this approximation is a fully analytic, though approximate, solution to the collisionless Boltzmann equation. It is highly accurate yet computationally efficient, and does not rely on any numerical fits or calibrations. We implement this approximation in CLASSIER, apply it to the case of massive neutrinos, and demonstrate that it improves the total runtime of the code by a factor of two, with even larger gains expected as cosmological analyses push to smaller scales. We also validate the accuracy of the approximation against fully numerical solutions, showing that it faithfully reproduces both the oscillatory features and the late-time asymptotics of the multipole moments. Our results highlight the utility of the small-scale limit as a complementary tool to integral-equation methods, providing further efficiency gains for cosmological modeling in the precision era.

The rest of this paper is organized as follows. In Section II, we derive the small-scale approximation for non-cold relic perturbations. In Section III, we describe its implementation in CLASSIER and assess the accuracy and performance of the approximation when applied to massive neutrino, by comparing with fully numerical solutions. We conclude in Section IV.

## II. ANALYTIC SMALL-SCALE APPROXIMATION FOR NCDM PERTURBATIONS

In this Section, we derive analytic approximations to the collisionless Boltzmann equation in the small-scale, free-streaming regime.

### A. Collisionless Boltzmann equation

We denote by  $f(\vec{q}, \vec{k}, \tau) \equiv f_0(q)[1 + \Psi(q, \vec{k}, \mu, \tau)]$  the perturbed NCDM phase-space distribution at confor-

mal time  $\tau$ , with comoving wavenumber  $\vec{k}$  and comoving momentum<sup>2</sup>  $\vec{q}$ . Throughout we denote by  $\epsilon(q, \tau) = \sqrt{q^2 + a^2 m^2}$  the comoving energy, and define  $\mu \equiv \vec{k} \cdot \hat{q}$ . In the synchronous gauge, the collisionless Boltzmann equation governing the evolution of the phase-space perturbation  $\Psi(q, \vec{k}, \mu, \tau)$  is [6]

$$\dot{\Psi} + ik\mu \frac{q}{\epsilon} \Psi + \frac{d \ln f_0}{d \ln q} \left[ \dot{\eta} - \frac{\dot{h} + 6\dot{\eta}}{2} \mu^2 \right] = 0, \quad (1)$$

where overdots denote derivatives with respect to conformal time  $\tau$ . The source terms  $\dot{h}(\vec{k}, \tau)$  and  $\dot{\eta}(\vec{k}, \tau)$  are the derivatives of the synchronous gauge metric perturbations with respect to conformal time.

We may decompose  $\Psi(\mu)$  on the basis of Legendre polynomials as usual through [6]

$$\Psi(q, \vec{k}, \mu, \tau) = \sum_{\ell} (-i)^{\ell} (2\ell + 1) \Psi_{\ell}(\vec{k}, q, \tau) P_{\ell}(\mu). \quad (2)$$

Only the first three multipole moments are directly relevant for the Einstein field equations.

### B. Small-scale quasi-stationary approximation

While the metric perturbations are themselves sourced by NCDM perturbations through the perturbed Einstein field equations, we may still formally treat them as external sources. Our goal is then as follows: given an “exact” solution for  $\Psi$  up to some time  $\tau_*$ , and given metric perturbations at all times, we seek an approximate solution at subsequent times  $\tau \geq \tau_*$ .

Our small-scale approximation relies on two key observations:

- After horizon entry,  $\dot{\eta}$  is a rapidly oscillating function with a quickly decaying amplitude  $\dot{\eta}(k\tau \gg 1) \sim \sin(k\tau/\sqrt{3})/(k\tau)^2$  during radiation domination, and vanishing during matter domination, as we show in Appendix A. At times  $\tau \gg 1/k$  we may thus simply neglect the term  $\dot{\eta}$  in the Boltzmann equation for  $\Psi$ .
- The combination  $\dot{h} + 6\dot{\eta}$  is mostly a smooth function of time, varying on a Hubble timescale, as derived in Appendix A and illustrated in Fig. 1. This function also shows rapid oscillations superimposed on top of its running mean, but their amplitude is much smaller than the overall magnitude of  $\dot{h} + 6\dot{\eta}$ . Note that the relative amplitude of oscillations is slightly smaller for that combination of metric terms than for  $\dot{h}$  alone.

For sufficiently small scales, the source term  $\dot{h} + 6\dot{\eta}$  thus varies on a much longer timescale than the time

<sup>1</sup> Available at <https://github.com/nanoomlee/CLASSIER>.

<sup>2</sup> While we keep equations general in our derivation, when implementing our approximation for massive neutrinos, we follow the CLASS convention and work with comoving momenta and energies rescaled by their temperature today.

it takes to free-stream across a comoving wavelength. We may therefore seek a quasi-stationary (QS) particular solution  $\Psi_{\text{QS}}$  of Eq. (1), for which  $\partial_\tau \Psi_{\text{QS}} \sim aH\Psi_{\text{QS}} \ll (kq/\epsilon)\Psi_{\text{QS}}$ . This quasi-stationary solution is to be summed with a homogeneous solution to smoothly connect with the exact solution at  $\tau_*$ .

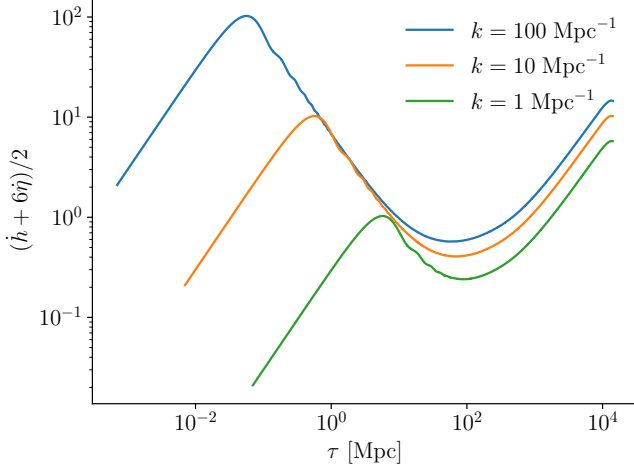


FIG. 1. The source term,  $(\dot{h} + 6\dot{\eta})/2$  in the Boltzmann equation, which determines the smooth late-time behavior of neutrino perturbations. The relatively small rapid oscillations on top of a smooth behavior allows us to find an accurate quasi-stationary analytic approximation for  $\Lambda$ CDM perturbations.

To derive the solution in a most compact form, for a fixed  $(k, q)$ , we start by defining the dimensionless variable

$$\xi \equiv kq \int_{\tau_*}^{\tau} \frac{d\tau'}{\epsilon(q, \tau')}. \quad (3)$$

Physically, this variable represents the ratio of the comoving distance traveled by a particle of momentum  $q$  between  $\tau_*$  and  $\tau$  to the comoving lengthscale of the perturbation. We then rewrite the Boltzmann equation for  $\Psi$  (neglecting the metric term  $\dot{\eta}$ ) in the form

$$\Psi'(\xi) + i\mu\Psi = S(\xi)\mu^2, \quad (4)$$

where primes denote derivatives with respect to  $\xi$ , and

$$S(\xi) \equiv \frac{\epsilon}{kq} \frac{d \ln f_0}{d \ln q} \frac{\dot{h} + 6\dot{\eta}}{2}. \quad (5)$$

The slow variation of the source term implies that  $S'(\xi)/S(\xi) \sim (aH)\epsilon/kq \equiv s \ll 1$ , and we may seek quasi-stationary solutions as a series expansion in that small parameter. We found that it is sufficient to go to second order in  $s$ , for which the quasi-stationary approximation is

$$\Psi_{\text{QS}}(\xi, \mu) = -iS(\xi)\mu + S'(\xi). \quad (6)$$

A full solution that is continuous at  $\tau_*$  (corresponding to  $\xi_* = 0$ ) is then easily obtained by adding the homoge-

neous solution  $e^{-i\mu\xi}$  with the adequate amplitude:

$$\Psi^{\text{approx}}(\xi, \mu) \equiv \Psi_{\text{QS}}(\xi, \mu) + e^{-i\mu\xi} [\Psi(\xi_*, \mu) - \Psi_{\text{QS}}(\xi_*, \mu)]. \quad (7)$$

This expression can also be obtained starting from the integral solution of the differential equation, integrating by parts twice, and neglecting the remainder proportional to  $S''$ . This full solution is our complete small-scale approximation. Its explicit first three multipole moments are

$$\Psi_0^{\text{approx}}(\xi) = S'(\xi) + \sum_{\ell} (-1)^{\ell} (2\ell + 1) j_{\ell}(\xi) \Delta_{\ell}(\xi_*) \quad (8)$$

$$\Psi_1^{\text{approx}}(\xi) = \frac{1}{3} S(\xi) - \sum_{\ell} (-1)^{\ell} (2\ell + 1) j'_{\ell}(\xi) \Delta_{\ell}(\xi_*) \quad (9)$$

$$\Psi_2^{\text{approx}}(\xi) = \sum_{\ell} (-1)^{\ell} (2\ell + 1) \frac{3j''_{\ell}(\xi) + j_{\ell}(\xi)}{2} \Delta_{\ell}(\xi_*), \quad (10)$$

where

$$\begin{aligned} \Delta_{\ell}(\xi_*) &\equiv \Psi_{\ell}(\xi_*) - \Psi_{\text{QS},\ell}(\xi_*) \\ &= \Psi_{\ell}(\xi_*) - S'(\xi_*)\delta_{\ell 0} - \frac{1}{3}S(\xi_*)\delta_{\ell 1}. \end{aligned} \quad (11)$$

Equations (8)-(11) constitute our main result<sup>3</sup>.

Note that all the terms involving the spherical Bessel functions vanish at  $\xi \rightarrow \infty$ , so that at late times, the approximation reduces to  $\Psi(\xi, \mu) \approx \Psi_{\text{QS}}(\xi, \mu)$  as expected. Given that, at late times,  $\dot{\eta} \rightarrow 0$ , but  $\dot{h}$  grows, the dominant asymptotic terms are

$$\begin{aligned} \Psi_0 &\approx \frac{1}{2} \frac{d \ln f_0}{d \ln q} \frac{\epsilon \dot{h}}{(kq)^2}, \\ \Psi_1 &\approx \frac{1}{6} \frac{d \ln f_0}{d \ln q} \frac{\epsilon \dot{h}}{kq}, \\ \Psi_2 &\rightarrow 0. \end{aligned} \quad (12)$$

Note that the late-time behavior of  $\Psi_0$  above matches the analytic result of Ref. [12] [Eq. (A22)], in the Newtonian gauge (see also Ref. [20] for the asymptotic expression of density perturbations).

The accuracy of our small-scale approximation in Eqs. (8)-(11) is demonstrated in Fig. 2, which compares

<sup>3</sup> The explicit forms of the small-scale approximation [Eqs. (8)-(10)] suggest a potential extension: one could directly rewrite the  $\Lambda$ CDM contributions to the metric source terms in Einstein equations using these expressions, thereby bypassing the explicit evolution of the  $\Lambda$ CDM multipole hierarchy. Such a strategy could yield further efficiency gains while broadening the applicability of the method to non-standard dark matter scenarios with distinctive small-scale signatures. However, we do not consider the implementation of this idea in CLASSIER as it will require solving different differential equations for different momentum bins making the implementation complex, unless the approximation we develop is valid for all momentum bins.

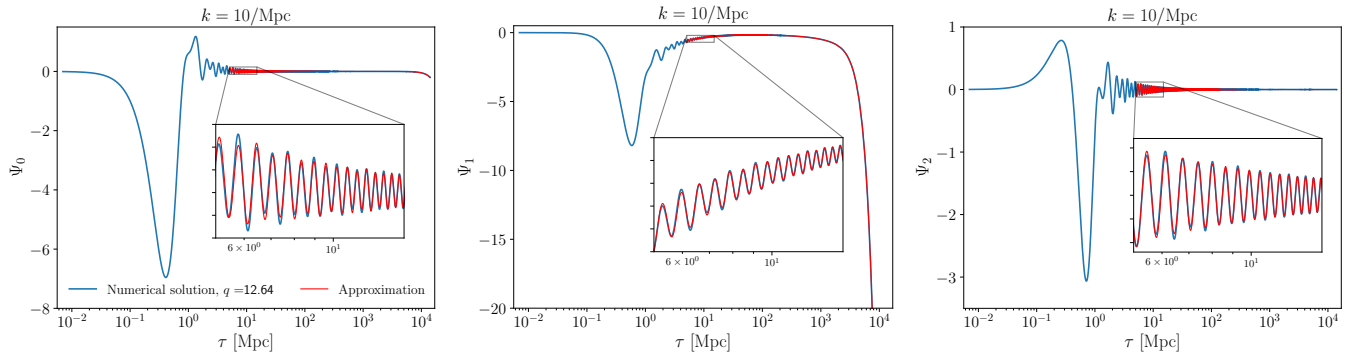


FIG. 2. The first three multipole moments of massive neutrino phase-space distribution function perturbations,  $\Psi_0$ ,  $\Psi_1$ , and  $\Psi_2$ , for comoving wavenumber  $k = 10 \text{ Mpc}^{-1}$  and comoving momentum  $q = 12.64 T_0$ , where  $T_0$  is the neutrino temperature today. Blue: numerical solutions from CLASSIER; Red: small-scale approximations from Eqs. (8)–(10).

it with fully numerical solutions for  $\Psi_0, \Psi_1, \Psi_2$ , shown for comoving wavenumber  $k = 10 \text{ Mpc}^{-1}$  and comoving momentum  $q \simeq 12 T_0$ , where  $T_0$  is the neutrino temperature today. The blue curves are full numerical solutions obtained using CLASSIER, which were shown to agree with high-precision CLASS calculations [16], while the red curves show our small-scale approximations. The figure illustrates that our method accurately reproduces both the late-time behavior of massive neutrino perturbations and their oscillatory features arising from the homogeneous solution encoding early-time contributions. Although the oscillating part near the switch at  $\tau_*$  shows slight differences with the numerical solution, these small differences in rapidly oscillating terms have very little impact on the matter power spectrum, as we show explicitly in the next section.

The small-scale approximation we develop here can easily be applied to massless species (e.g., massless neutrinos). In the massless limit, we have  $\epsilon = q$  and  $\xi = k(\tau - \tau_*)$ . We provide the small-scale approximation in this limit, in Appendix. B

### III. IMPLEMENTATION AND RESULTS

#### A. Implementation in CLASSIER

We implement the small-scale approximation developed in Section II into CLASSIER, a code that replaces the traditional truncated Boltzmann hierarchy for NCDM perturbations with a set of integral equations evaluated iteratively for self-consistency. In this work, we apply the approximation to the standard case of massive neutrinos.

In CLASS, the parameter `ppr->ncdm_fluid_trigger_tau_over_tau_k` (default value 31) sets the conformal time at which the fluid approximation for NCDM begins, i.e.  $\tau = 31/k$ . The same prescription applies in the zeroth iteration of CLASSIER. For improved accuracy, we instead adopt  $\tau = 51/k$  and increase the truncation order of the Boltzmann hierarchy to  $\ell_{\text{max}}^{\text{NCDM}} = 60$  (default 17 in CLASS). These modifications add negligible cost but im-

prove the accuracy of the subsequent integral-equation iterations. In practice, the zeroth iteration evolves the truncated hierarchy up to  $\tau = 51/k$ , after which the calculation switches to the fluid approximation. Later iterations evaluate the integral solutions for neutrino perturbations as convolutions using FINUFFT (Flatiron Institute Nonuniform Fast Fourier Transform) [21, 22]<sup>4</sup>.

As explained earlier, our small-scale approximation is valid as long as the parameter  $s \equiv aH\epsilon/kq$  is small, i.e. as long as the perturbation comoving wavelength is small relative to the comoving distance free-streamed in a Hubble time. While the free-streaming particles are relativistic, this parameter is simply  $s \approx aH/k$ , and decreases monotonically with time. However, once particles become non-relativistic,  $s \approx a^2 H m/(kq)$ , which increases during matter domination. Therefore, it is possible for some momentum bins  $q$  to satisfy  $s \ll 1$  for some period of time, but no longer once the particles become sufficiently non-relativistic.

We could in principle switch off the small-scale approximation at a different time for each momentum  $q$  and wavenumber  $k$ . However, this would significantly complicate our implementation, with little gains in terms of overall computational efficiency. Therefore, we instead choose to use the small-scale approximation for a given  $(q, k)$  pair only if it holds at the present time  $\tau_0$ , i.e. only if  $s \ll 1$  at  $\tau_0$ . If this is indeed the case, we switch on the small-scale approximation at  $\tau_* = 51/k$ , and use it until the present time.

Concretely, we found that the small-scale approximation yields better than 1% accuracy in  $\Psi_0$  today provided that the following criterion is satisfied:

$$s(k, q, \tau_0) \equiv \frac{H_0 \epsilon(q, \tau_0)}{kq} < 0.0025. \quad (13)$$

We show in Fig. 3 the regions of  $(k, q)$  parameter space where the criterion (13) holds. Whenever the criterion

<sup>4</sup> <https://finufft.readthedocs.io/en/latest/>

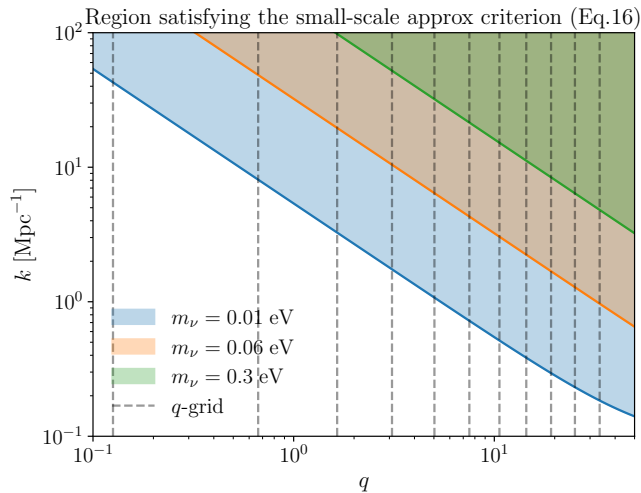


FIG. 3. Regions in  $(k, q)$  space that satisfy the small-scale approximation criterion of Eq. (13). Different colors correspond to different neutrino masses. Vertical dashed lines indicate the  $q$ -grid used. Here  $q$  is the comoving momentum in units of  $T_0$ , the neutrino temperature today.

(13) is satisfied, we replace the full integral evaluation with the small-scale approximation, Eqs. (8)–(10), where the initial multipoles  $\Psi_\ell(\tau_*)$  are evaluated at  $\tau_* \equiv 51/k$  for  $\ell \leq \ell_{\max}^{\text{NCDM}}$ .

The small-scale approximation is straightforward to compute but involves repeated evaluations of spherical Bessel functions  $j_\ell(\xi)$ , of their first and second derivatives. Since these functions only need to be computed once, we precompute them on a grid of 5000 equally spaced values of  $\xi$  over  $0 \leq \xi < 10^3$  and store them in external file (`ncdmfft/j_e11.txt`). When the small-scale approximation is enabled, the code loads these tables and interpolates as needed, eliminating the overhead of repeatedly evaluating the kernel functions during runtime.

To further reduce the computational cost without loss of accuracy, at  $\xi > 1000$ , we switch from the full small-scale approximation [Eqs. (8)–(10)] to these simplified late-time expressions [Eq. (12)].

We note that in the original CLASSIER implementation the NUFFT grids were chosen conservatively to resolve the rapid oscillations induced by large momentum bins. With the small-scale approximation handling precisely these bins, such high grid resolution is no longer necessary. Accordingly, when the approximation is active the grid density is adaptively reduced for large- $k$  modes, providing an additional runtime saving on top of the analytic speed-up.<sup>5</sup>

Finally, note that in principle, one could iterate on the procedure we described above, i.e. start with standard CLASS, with the fluid approximation, to compute the metric perturbations, use CLASSIER with the small-scale approximation to compute NCDM perturbations, store the latter, and re-compute metric perturbations sourced by these more accurate NCDM perturbations, etc... In practice, we have found that one single iteration is enough to reach sub-0.1% accuracy on the matter power spectrum, as we describe further in the next section.

## B. Runtime improvements

We now assess the performance of the small-scale approximation developed in this work, both in terms of computational efficiency and accuracy of cosmological observables. Throughout the runtime and accuracy comparisons, we adopt the default cosmological parameters of CLASS v3.2.2 with one massive neutrino of mass 0.06 eV, and use the CLASS:UHP precision setting of Ref. [23], unless stated otherwise.

Figure 4 shows the runtime distribution for three representative  $k$ -modes (100, 10, and 1  $\text{Mpc}^{-1}$ ), separated into contributions from the zeroth-iteration step that solves the truncated Boltzmann hierarchy for NCDM with the late-time fluid approximation (blue), from the integral-equation evaluation (red hatched), and from the rest of the system (orange). The small-scale approximation substantially reduces the time required for the integral-equation step, which otherwise accounts for a noticeable fraction of the runtime at high  $k$ . At lower  $k$ , fewer momentum bins satisfy the criterion of Eq. (13) (see Fig. 3), so the runtime reduction is less pronounced; however, these low- $k$  modes contribute little to the overall runtime budget. Taken together, these results indicate that the small-scale approximation delivers savings close to the maximum possible in the portion of the calculation dominated by NCDM perturbations.

Table I summarizes the runtime performance for different choices of  $k_{\max}$ , for CLASS, the original version of CLASSIER, and with the small-scale approximation described in the present work. For completeness, results under a lower reference precision setting, corresponding to the CLASS default precision with two modifications, `l_max_ncdm=40` ( $\ell_{\max}^{\text{NCDM}} = 40$ ) and `ncdm_fluid_approximation=3` (no fluid approximation), are given in Appendix C.

These results show that the small-scale approximation substantially reduces the cost of evaluating NCDM perturbations in CLASSIER. For example, for  $k_{\max} =$

<sup>5</sup> We also introduce a new parameter `k_ncdmfft_min`, set to  $0.5 \text{ Mpc}^{-1}$  in this work. For  $k$  below this threshold, the code reverts to the standard CLASS treatment with the Runge–Kutta integrator (`evolver=0`). Under the current precision setting, the standard CLASS method can appear faster than CLASSIER, partly

because of the relatively small number of momentum bins and partly due to the present structure of the convolution step in CLASSIER. This latter effect mainly impacts low- $k$  modes, and could potentially be mitigated by restructuring the code in future revisions.



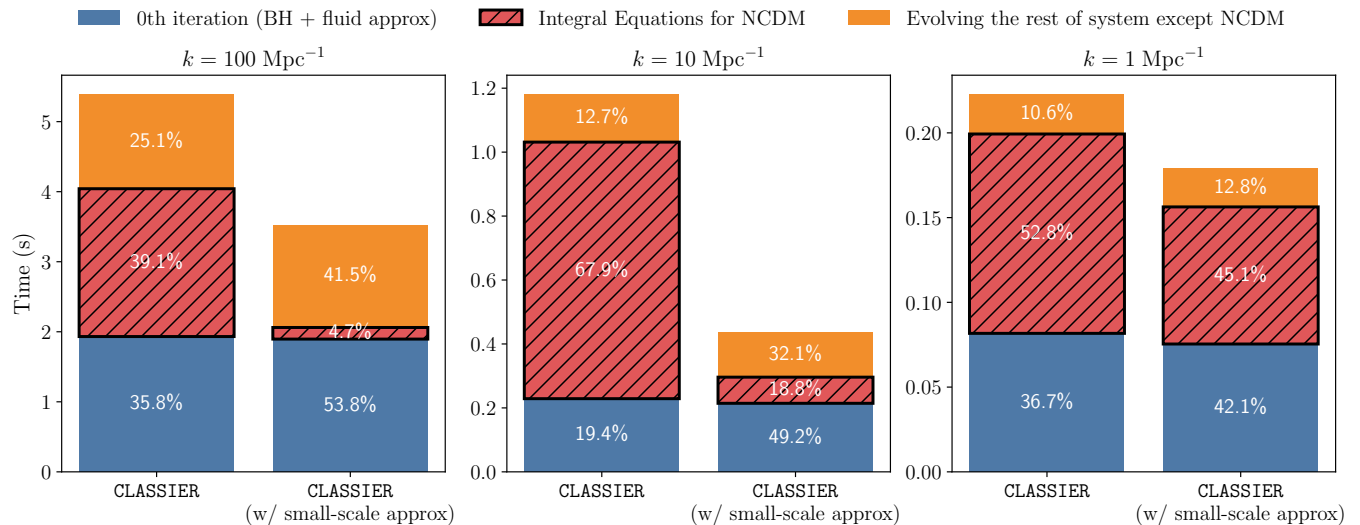


FIG. 4. Runtime distribution for evolving perturbations of three representative  $k$ -modes in CLASSIER (left:  $k = 100 \text{ Mpc}^{-1}$ , middle:  $k = 10 \text{ Mpc}^{-1}$ , right:  $k = 1 \text{ Mpc}^{-1}$ ), shown without (left bar) and with (right bar) the small-scale approximation. Each bar is split into contributions from the zeroth-iteration step that solves the truncated Boltzmann hierarchy (BH) with the late-time fluid approximation (blue), from the integral-equation evaluation of NCDM perturbations (red hatched), and from the rest of the system (orange). The small-scale approximation substantially reduces the time spent on the NCDM integral-equation step, which otherwise accounts for a non-negligible share of the runtime at high  $k$ . At small  $k$ , fewer momenta qualify, but these modes contribute negligibly to the overall runtime. When the approximation is used, we reduce the resolution of  $\tau$ -grid for convolution calculations and this additionally saves the small amount of runtime.

$k_{\text{max}}$ (Mpc $^{-1}$ )	CLASS $\ell_{\text{max}}^{\text{nCDM}} = 40$	CLASSIER (original)	CLASSIER w/ small-scale approx.
10	16 sec	12 sec	<b>5 sec</b>
50	80 sec	30 sec	<b>15 sec</b>
100	160 sec	49 sec	<b>26 sec</b>

TABLE I. Runtimes (four threads, 16-inch MacBook Pro-M4 chip, Nov 2024 release) under the CLASS:UHP precision setting, for different  $k_{\text{max}}$ . The small-scale approximation leads to a reduction in total runtime relative to the original version of CLASSIER ranging from  $\sim 45\%$  to  $60\%$ . With the small-scale approximation, CLASSIER is faster than CLASS by a factor of 3–6. Single-thread runtimes are roughly four times longer.

$100 \text{ Mpc}^{-1}$  the total runtime decreases by 46% relative to the baseline CLASSIER runs. Similar improvements appear with lower  $k_{\text{max}}$ , with reductions of order 50% across the board. Importantly, these gains are achieved while preserving the accuracy of the solutions (see below). The breakdown in Fig. 4 clarifies that the overall savings are driven primarily by reductions in the NCDM integral-equation step, which constitutes a noticeable fraction of the runtime at high  $k$ . We also have checked that applying the small-scale approximation to massless neutrinos as shown in Appendix B can further reduce the total runtime by up to 20%.<sup>6</sup>

Throughout this work, CLASSIER employs the big- $q$  approximation of Ref. [16],<sup>7</sup> which already lowers the cost of NCDM evaluations. The runtime reductions reported here are therefore an additional improvement on top of that baseline.

### C. Accuracy of the matter power spectrum

To validate accuracy, we compute the  $z = 0$  matter power spectrum and compare it against a high-precision CLASS calculation with  $\ell_{\text{max}}^{\text{nCDM}} = 500$ . Figure 5 shows the absolute fractional differences. For reference, we also show standard CLASS with  $\ell_{\text{max}}^{\text{nCDM}} = 40$ , as well as CLASSIER without the small-scale approximation. This figure demonstrates that our small-scale approximation (orange curve) maintains accuracy at the  $\mathcal{O}(0.1\%)$  level

CLASSIER. The reason is that small numerical artifacts arising during the tight-coupling approximation in CLASS—when both  $\ell_{\text{max}}^{\text{nCDM}}$  and  $\ell_{\text{max}}^{\text{ur}}$  (the truncation scale of the momentum-integrated Boltzmann hierarchy for massless neutrinos) are increased—become amplified within the iterative integral-equation scheme. We plan to include this refinement in a future release once this issue is resolved.

<sup>7</sup> At early times, for sufficiently large comoving momenta  $q$ , the perturbations remain effectively relativistic so their  $q$ -dependence factors out of the source functions. The convolution integrals are then evaluated only for the largest  $q$ -bin and the results for other bins are obtained by rescaling, avoiding redundant calculations.

<sup>6</sup> This feature is not included in the current public release of

across the full  $k$ -range considered, performing essentially indistinguishably from the original integral solution (blue curve). Thus, the small-scale approximation reproduces the correct late-time behavior and the characteristic free-streaming suppression of power, leaving cosmological observables unaffected. We repeat this test with two additional neutrino mass values, 0.01 eV and 0.3 eV, and present with dashed and dot-dashed lines, respectively. This demonstrates that the accuracy of the small-scale approximation remains the same for a wide range of neutrino mass.

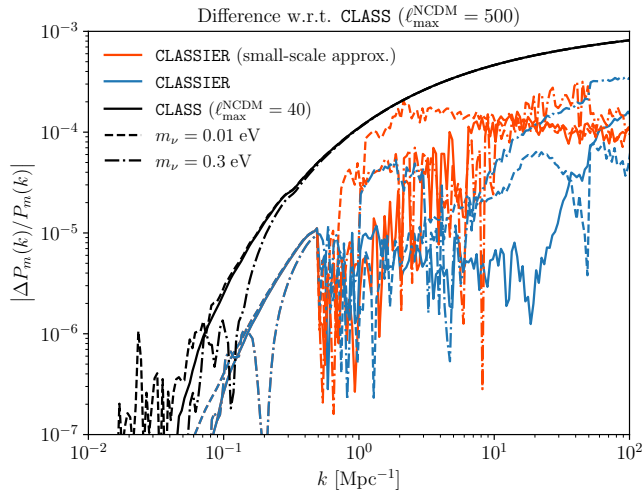


FIG. 5. Absolute fractional differences of the  $z = 0$  matter power spectrum with respect to **CLASS** with  $\ell_{\text{max}}^{\text{NCDM}} = 500$ . The solid lines are with the neutrino mass  $m_\nu = 0.06\text{eV}$ , and the dashed and the dot-dashed lines are with  $m_\nu = 0.01\text{eV}$  and  $0.3\text{eV}$ , respectively.

In summary, the small-scale approximation roughly halves the runtime of **CLASSIER** while preserving the accuracy of the matter power spectrum, making it well-suited for applications that require repeated evaluations of cosmological perturbations, such as parameter inference in upcoming surveys.

#### IV. CONCLUSION

We have derived a highly accurate analytic approximation for small-scale non-cold relic (NCDM) perturbations using quasi-steady-state particular solutions to the collisionless Boltzmann equation and implemented it in **CLASSIER**. Combined with the homogeneous solution which carries the early-time contribution, the approximation replaces costly late-time convolutions with compact analytic expressions. The approximation accurately captures the smooth late-time behavior of NCDM perturbations, in addition to the rapid oscillatory features due to free-streaming.

On the performance side, the approximation reduces the total runtime in **CLASSIER** by roughly a factor of two across a broad range of  $k_{\text{max}}$  (Table I). Compared to

standard **CLASS** runs (with  $\ell_{\text{max}}^{\text{NCDM}} = 40$  and no late-time massive neutrino fluid approximation) under the same precision setting, **CLASSIER** with this approximation is faster by a factor of 3–6. On the accuracy side, comparisons to high-precision **CLASS** calculations show sub-0.1% agreement in the matter power spectrum today (Fig. 5), faithfully capturing the small-scale suppression from free-streaming. The method therefore achieves substantial speed-ups with no degradation in accuracy, making it particularly well suited for parameter inference pipelines in the precision cosmology era.

The validity of the approximation is governed by a simple scale-dependent condition in Eq. (13) expressed in terms of  $s \equiv aH\epsilon/kq$ , which quantifies the ratio of the perturbation wavelength to the distance free-streamed per Hubble time. Even if a species becomes non-relativistic, the quasi-steady-state solution remains accurate on sufficiently small scales provided  $s$  is small enough throughout the evolution. Within this regime, the late-time behavior of the leading multipoles is controlled by simple  $h$ -derivative terms [Eq. (12)].

Although we have benchmarked the numerics for the canonical case of massive neutrino, the approach applies to exotic non-cold relics [24–29] and with some modifications to models with dark-matter decays to non-cold relics [13, 30, 31]. The improved precision and efficiency in the wavenumber range  $1 \lesssim (k/h \text{ Mpc}^{-1}) \lesssim 100$  may be useful for recent models [32, 33].

We release an implementation in **CLASSIER** that precomputes the required kernels, adapts grids when the approximation is active, and falls back seamlessly to the full integral evaluation when the criterion is not met. We expect the approximation to deliver increasing benefits as upcoming analyses push to smaller scales and as likelihood pipelines demand ever larger numbers of forward model evaluations.

#### ACKNOWLEDGEMENTS

We thank J. Lesgourgues and C. Nascimento for useful discussions. This work was supported at JHU by NSF Grant No. 2412361, NASA ATP Grant No. 80NSSC24K1226, and the Templeton Foundation. NL was supported by the Horizon Fellowship from Johns Hopkins University.

#### Appendix A: Small-scale behavior of synchronous-gauge potentials

In this appendix we derive the small-scale behavior of synchronous-gauge potentials  $\dot{\eta}$  and  $\dot{h}$ . We do so starting from known properties of the Newtonian-gauge potentials  $\phi$ ,  $\psi$  and change gauge to obtain  $\dot{\eta}$ ,  $\dot{h}$ . The gauge-transformation equations for the metric perturbations  $\eta$

and  $\alpha \equiv \dot{h} + 6\dot{\eta}$  are [6]

$$\psi = \frac{1}{2k^2} \frac{1}{a} \frac{d}{d\tau}(a\alpha), \quad \phi = \eta - \frac{aH}{2k^2}\alpha. \quad (\text{A1})$$

• In the radiation era, neglecting anisotropic stress and neutrino contributions, it is well known that

$$\phi(k, \tau) \approx \psi(k, \tau) \approx 3\phi_i(k) \frac{j_1(x)}{x}, \quad x \equiv k\tau/\sqrt{3}. \quad (\text{A2})$$

Using the fact that  $a \propto \tau \propto x$  in the radiation era, we may solve Eqs. (A1) explicitly and obtain

$$\alpha \approx 6\sqrt{3}k\phi_i(k) \frac{1 - j_0(x)}{x}, \quad (\text{A3})$$

$$\eta \approx 3\phi_i(k) \left[ \frac{1 - j_0(x)}{x^2} + \frac{j_1(x)}{x} \right], \quad (\text{A4})$$

implying

$$\dot{\eta} \approx \sqrt{3}k\phi_i(k) \left[ 2 \frac{j_0(x) - 1 - xj_1(x)}{x^3} + \frac{j_0(x)}{x} \right]. \quad (\text{A5})$$

We see that, at late times  $k\tau \gg 1$  (but still in the radiation era),  $\dot{\eta} \propto \sin(k\tau/\sqrt{3})/(k\tau)^2$ . We also see that the oscillatory part of  $\alpha$  is suppressed relative to its running mean by a factor of order  $\sim 1/(k\tau) \ll 1$  at late times.

• During matter domination,  $\phi \approx \psi \approx \text{constant}$  and  $a \propto \tau^2$ , implying  $\alpha \approx (\phi/3)\tau$  and  $\eta \approx \phi/3$ , hence  $\dot{\eta} \approx 0$ .

### Appendix B: Small-scale approximation in massless limit

In the massless limit  $q/\epsilon = 1$  and the perturbation to the phase-space density only depends on momentum through the prefactor  $d \ln f_0/d \ln q$ :

$$\Psi_\ell(\vec{k}, q, \mu, \tau) = \frac{d \ln f_0}{d \ln q} \tilde{\Psi}(\vec{k}, \mu, \tau). \quad (\text{B1})$$

Following Ref. [6], we may integrate out the  $q$ -dependence in the phase-space distribution perturbation:

$$F(\vec{k}, \mu, \tau) \equiv \frac{\int q^2 dq q f_0(q) \Psi(\vec{k}, q, \mu, \tau)}{\int q^2 dq q f_0(q)} = -4\tilde{\Psi}(\vec{k}, \tau), (\text{B2})$$

where in the second equality we inserted Eq. (B1) and integrated by parts, assuming  $f_0(q \rightarrow \infty) = 0$ . The small-scale approximation for the first three multipoles  $F_\ell(\vec{k}, \tau)$  then takes the same form as Eqs. (8)-(11), with  $\xi = k(\tau - \tau_*)$ ,  $\Psi_\ell(\xi_*) \rightarrow F_\ell(\xi_*)$ , and  $S \rightarrow -2(\dot{h} + 6\dot{\eta})/k$ .

### Appendix C: Runtime under lower reference precision setting

For completeness, Table II shows runtime comparisons under a lower reference precision setting, based on the CLASS default precision but with two modifications: `l_max_ncdm=40` ( $\ell_{\text{max}}^{\text{NCDM}} = 40$ ) and `ncdm_fluid_approximation=3` (no fluid approximation).

$k_{\text{max}}$ (Mpc <sup>-1</sup> )	CLASS $\ell_{\text{max}}^{\text{ncdm}} = 40$	CLASSIER (original)	CLASSIER w/ small-scale approx.
10	2.6 sec	1.9 sec	<b>1.2 sec</b>
50	11 sec	3.9 sec	<b>2.2 sec</b>
100	21 sec	6.0 sec	<b>3.4 sec</b>

TABLE II. Runtimes (four threads, 16-inch MacBook Pro-M4 chip, Nov 2024 release) under a lower reference precision setting, corresponding to the CLASS default precision setting except `l_max_ncdm=40` ( $\ell_{\text{max}}^{\text{NCDM}} = 40$ ) and `ncdm_fluid_approximation=3` (no fluid approximation), for different  $k_{\text{max}}$ . The small-scale approximation leads to a reduction in total runtime relative to the original version of CLASSIER  $\sim 40\%$ . With the small-scale approximation, CLASSIER is faster than CLASS by a factor of 2–6. Single-thread runtimes are roughly four times longer.

[1] J. Lesgourgues and S. Pastor, *Phys. Rept.* **429**, 307 (2006), [arXiv:astro-ph/0603494](#).  
[2] Y. Y. Y. Wong, *Ann. Rev. Nucl. Part. Sci.* **61**, 69 (2011), [arXiv:1111.1436 \[astro-ph.CO\]](#).  
[3] K. N. Abazajian *et al.* (Topical Conveners: K.N. Abazajian, J.E. Carlstrom, A.T. Lee), *Astropart. Phys.* **63**, 66 (2015), [arXiv:1309.5383 \[astro-ph.CO\]](#).  
[4] P. A. Abell *et al.* (LSST Science, LSST Project), (2009), [arXiv:0912.0201 \[astro-ph.IM\]](#).

[5] K. Abazajian *et al.*, (2019), [arXiv:1907.04473 \[astro-ph.IM\]](#).  
[6] C.-P. Ma and E. Bertschinger, *Astrophys. J.* **455**, 7 (1995), [arXiv:astro-ph/9506072](#).  
[7] D. Blas, J. Lesgourgues, and T. Tram, *JCAP* **07**, 034 (2011), [arXiv:1104.2933 \[astro-ph.CO\]](#).  
[8] W. Hu, *Astrophys. J.* **506**, 485 (1998), [arXiv:astro-ph/9801234](#).  
[9] A. Lewis and A. Challinor, *Phys. Rev. D* **66**, 023531 (2002), [arXiv:astro-ph/0203507](#).



- [10] M. Shoji and E. Komatsu, *Phys. Rev. D* **81**, 123516 (2010), [Erratum: *Phys.Rev.D* **82**, 089901 (2010)], [arXiv:1003.0942 \[astro-ph.CO\]](#).
- [11] J. Lesgourgues and T. Tram, *JCAP* **09**, 032 (2011), [arXiv:1104.2935 \[astro-ph.CO\]](#).
- [12] C. Nascimento, *Phys. Rev. D* **108**, 023505 (2023), [arXiv:2303.09580 \[astro-ph.CO\]](#).
- [13] G. Franco Abellán, R. Murgia, and V. Poulin, *Phys. Rev. D* **104**, 123533 (2021), [arXiv:2102.12498 \[astro-ph.CO\]](#).
- [14] M. Kamionkowski, *Phys. Rev. D* **104**, 063512 (2021), [arXiv:2105.02887 \[astro-ph.CO\]](#).
- [15] L. Ji, M. Kamionkowski, and J. L. Bernal, *Phys. Rev. D* **106**, 103531 (2022), [arXiv:2201.11129 \[astro-ph.CO\]](#).
- [16] N. Lee, J. L. Bernal, S. Günther, L. Ji, and M. Kamionkowski, (2025), [arXiv:2506.01956 \[astro-ph.CO\]](#).
- [17] U. Seljak and M. Zaldarriaga, *Astrophys. J.* **469**, 437 (1996), [arXiv:astro-ph/9603033](#).
- [18] Y. Ali-Haïmoud and S. Bird, *MNRAS* **428**, 3375 (2013), [arXiv:1209.0461 \[astro-ph.CO\]](#).
- [19] S. Bird, Y. Ali-Haïmoud, Y. Feng, and J. Liu, *MNRAS* **481**, 1486 (2018), [arXiv:1803.09854 \[astro-ph.CO\]](#).
- [20] A. Ringwald and Y. Y. Y. Wong, *JCAP* **12**, 005 (2004), [arXiv:hep-ph/0408241](#).
- [21] A. H. Barnett, J. Magland, and L. af Klinteberg, *SIAM J. Sci. Comput.* **41**, C479 (2019), [1808.06736](#).
- [22] A. H. Barnett, *Appl. Comput. Harmon. Anal.* (2020), [10.48550/arXiv.2001.09405](#), in press, [arXiv:2001.09405 \[math.NA\]](#).
- [23] M. Archidiacono *et al.* (Euclid), *Astron. Astrophys.* **693**, A58 (2025), [arXiv:2405.06047 \[astro-ph.CO\]](#).
- [24] F. D’Eramo and A. Lenoci, *JCAP* **10**, 045 (2021), [arXiv:2012.01446 \[hep-ph\]](#).
- [25] A. Banerjee, A. Bhattacharjee, S. Das, A. Maharana, and R. K. Sharma, (2025), [arXiv:2505.22775 \[astro-ph.CO\]](#).
- [26] T. Venumadhav, F.-Y. Cyr-Racine, K. N. Abazajian, and C. M. Hirata, *Phys. Rev. D* **94**, 043515 (2016), [arXiv:1507.06655 \[astro-ph.CO\]](#).
- [27] K. N. Abazajian and A. Kusenko, *Phys. Rev. D* **100**, 103513 (2019), [arXiv:1907.11696 \[hep-ph\]](#).
- [28] S. Bansal, S. Ghosh, M. Low, and Y. Tsai, (2024), [arXiv:2410.19224 \[astro-ph.CO\]](#).
- [29] J. Alvey, M. Escudero, and N. Sabti, *JCAP* **02**, 037 (2022), [arXiv:2111.12726 \[astro-ph.CO\]](#).
- [30] S. Aoyama, T. Sekiguchi, K. Ichiki, and N. Sugiyama, *JCAP* **07**, 021 (2014), [arXiv:1402.2972 \[astro-ph.CO\]](#).
- [31] A. Bencke, N. Lee, and M. Kamionkowski, In preparation.
- [32] A. C. Sobotka, A. L. Erickcek, and T. L. Smith, *Phys. Rev. D* **111**, 123522 (2025), [arXiv:2409.06778 \[astro-ph.CO\]](#).
- [33] R. T. Co, N. Fernandez, A. Ghalsasi, K. Harigaya, and J. Shelton, (2025), [arXiv:2510.01308 \[hep-ph\]](#).