



**HAL**  
open science

# Higher order initial conditions for mixed baryon–CDM simulations

Oliver Hahn, Cornelius Rampf, Cora Uhlemann

► **To cite this version:**

Oliver Hahn, Cornelius Rampf, Cora Uhlemann. Higher order initial conditions for mixed baryon–CDM simulations. *Monthly Notices of the Royal Astronomical Society*, 2021, 503 (1), pp.426-445. 10.1093/mnras/staa3773 . hal-02939986

**HAL Id: hal-02939986**

**<https://hal.science/hal-02939986>**

Submitted on 5 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Higher order initial conditions for mixed baryon–CDM simulations

Oliver Hahn <sup>1,2,3</sup>★ Cornelius Rampf <sup>1</sup> and Cora Uhlemann <sup>4</sup>

<sup>1</sup> *Université Côte d’Azur, Observatoire de la Côte d’Azur, CNRS, Laboratoire Lagrange, Boulevard de l’Observatoire, CS 34229, F-06304 Nice, France*

<sup>2</sup> *Department of Astrophysics, University of Vienna, Türkenschanzstraße 17, A-1180 Vienna, Austria*

<sup>3</sup> *Department of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria*

<sup>4</sup> *School of Mathematics, Statistics and Physics, Herschel Building, Newcastle University, Newcastle upon Tyne NE1 7RU, UK*

Accepted 2020 December 2. Received 2020 November 25; in original form 2020 August 20

## ABSTRACT

We present a novel approach to generate higher order initial conditions (ICs) for cosmological simulations that take into account the distinct evolution of baryons and dark matter. We focus on the numerical implementation and the validation of its performance, based on both collisionless  $N$ -body simulations and full hydrodynamic Eulerian and Lagrangian simulations. We improve in various ways over previous approaches that were limited to first-order Lagrangian perturbation theory (LPT). Specifically, we (1) generalize  $n$ th-order LPT to multifluid systems, allowing 2LPT or 3LPT ICs for two-fluid simulations, (2) employ a novel propagator perturbation theory to set up ICs for Eulerian codes that are fully consistent with 1LPT or 2LPT, (3) demonstrate that our ICs resolve previous problems of two-fluid simulations by using variations in particle masses that eliminate spurious deviations from expected perturbative results, (4) show that the improvements achieved by going to higher order PT are comparable to those seen for single-fluid ICs, and (5) demonstrate the excellent (i.e. few per cent level) agreement between Eulerian and Lagrangian simulations, once high-quality initial conditions are used. The rigorous development of the underlying perturbation theory is presented in a companion paper. All presented algorithms are implemented in the MONOFONIC MUSIC-2 package that we make publicly available.

**Key words:** methods: numerical – (galaxies:) intergalactic medium – (cosmology:) dark matter – (cosmology:) large-scale structure of Universe – cosmology: theory.

## 1 INTRODUCTION

The physics of the cosmic microwave background (e.g. Hu & Dodelson 2002; Durrer 2008) implies that baryons do not trace the distribution of dark matter. Since baryons were tightly coupled to photons prior to recombination, they begin to collapse on subhorizon scales much later than dark matter. Furthermore, sound waves excited prior to recombination decay away over some finite time and the resulting relative streaming motion (Tseliakhovich & Hirata 2010) leaves an imprint on the formation of the very first cosmic objects and their spatial distribution (cf. Dalal, Pen & Seljak 2010; Greif et al. 2011; Yoo, Dalal & Seljak 2011; Fialkov et al. 2012, and many later studies) on scales correlated with the baryon acoustic oscillation (BAO) feature of the power spectrum, which is one of the most sensitive cosmological measures. Such correlations, if they carry over to the galaxy populations in lower redshift surveys, are therefore potentially significant biases in BAO measurements of cosmological parameters (e.g. Slepian & Eisenstein 2015; Ahn 2016; Blazek, McEwen & Hirata 2016; Slepian et al. 2018; Chen, Castorina & White 2019).

While the difference in the clustering of baryons and cold dark matter (CDM) – the baryon bias – is small on large scales today (e.g. Angulo, Hahn & Abel 2013), this difference is more significant at earlier times. These epochs are increasingly within reach of

ever more sensitive observations (e.g. the Square Kilometre Array, SKA, Weltman et al. 2020). At late times and on small scales, the finite temperature of baryons, along with energy injection from supermassive black holes leads to a de-correlation of baryons and CDM in the mature low-redshift Universe (e.g. Chisari et al. 2019), complicating access to cosmological information on those scales. While the cosmological baryon bias is a robust prediction of our cosmological model, the latest generation of cosmological galaxy formation simulations in large-scale structure context (e.g. Dubois et al. 2014; Schaye et al. 2015; Springel et al. 2018; Embersson et al. 2019, for the Horizon-AGN, Eagle, Illustris-TNG, and Borg-Cube simulations) still do not model it due to multiple reasons that we discuss below. At the same time, the precision determination of the matter power spectrum has now reached the level at which baryonic effects should be incorporated in predictions from collisionless ‘total matter’  $N$ -body simulations (cf. Schneider & Teyssier 2015; Huang et al. 2019; Schneider et al. 2019; Aricò et al. 2020).

Accurate numerical studies of the rich dynamics of the two-fluid system of collisionless dark matter and collisional baryons from the cosmological perspective have been limited by a range of problems. This includes the generation of simulation initial conditions (ICs) based on perturbation theory (PT). For single-fluid simulations, the Zel’dovich approximation (ZA; Zel’dovich 1970) has been used since the early days of cosmological  $N$ -body methods to set up ICs (Klypin & Shandarin 1983; Efstathiou et al. 1985), and it is by now standard to employ second-order Lagrangian perturbation theory (2LPT), e.g.

\* E-mail: [oliver.hahn@univie.ac.at](mailto:oliver.hahn@univie.ac.at)

Crocce, Pueblas & Scoccimarro (2006). However, to date, higher order LPT has not been developed for multicomponent systems of baryons and dark matter. Some findings have been obtained in the context of Eulerian PT (Somogyi & Smith 2010; Bernardeau, van de Rijt & Vernizzi 2012), but have had no impact on increasing the accuracy of simulations so far. At the same time, simulations studying the baryon streaming predicted by Tseliakhovich & Hirata (2010) impose an *ad hoc* relative velocity between baryons and dark matter but do not self-consistently account for the non-linear coupling of such a relative velocity in the fluid PT used to set up the simulation ICs. Such relative motion appears as a decaying mode, not sourced by gravity, and is particularly difficult to tackle in the setting of a rigorous PT for simulation ICs.

$N$ -body simulation ICs typically start by considering an initially (statistically) uniform discrete universe of particles, on to which the cosmological perturbations are imprinted, namely by perturbing the particle positions and velocities. The problem of decaying-mode initial conditions is that they are fundamentally inconsistent with this boundary condition (in the sense that for earlier times, one would approach a more inhomogeneous state). In full generality, the dark matter and baryon perturbations computed by Einstein–Boltzmann codes of course contain a multitude of effects that are not captured by standard LPT (due to the absence of physical effects beyond the Newtonian two-fluid model).

For this reason, simulations attempting to take into account more realistic two-fluid perturbations so far employ only a first-order accurate approach, which simply guarantees that density and velocity power spectra imposed on the  $N$ -body particles have amplitudes that are in accordance with the linear Einstein–Boltzmann system (e.g. Yoshida, Sugiyama & Hernquist 2003; Hahn & Abel 2011 for technical details). It has quickly been noted however that simulations of two fluids (i.e. dark matter and baryons) initialized in this way do not accurately reproduce the relative growth between baryons and CDM even on scales where linear PT should apply (O’Leary & McQuinn 2012; Angulo et al. 2013), unless a much larger force softening is applied to the baryon particles than would be typical in usual  $N$ -body simulations. This finding has been confirmed in a later analysis also by Valkenburg & Villaescusa-Navarro (2017). Very recently, Bird et al. (2020) have claimed that the large softening can be circumvented by arranging baryon and dark matter  $N$ -body particles in a more refined way than just on two shifted lattices (on which we will comment later on). In any case, the current state-of-the-art view is that two-fluid  $N$ -body simulations require a more careful suppression of discreteness effects than single fluid ‘total matter’  $N$ -body simulations.

In this paper, we show how to numerically implement high-order ICs for two-fluid cosmological simulations while minimizing discretization errors. Among other things, we discuss how Lagrangian PT can be used to generate growing-mode initial conditions for multiple cold fluids. Such ICs preclude decaying modes, and therefore any relative velocities between the fluids. While we also provide a numerical procedure that includes the linear effects of relative velocities in LPT ICs, it is currently unclear how such decaying modes can be implemented to achieve consistent higher order ICs. We defer this aspect to future work.

The growing-mode approach that we focus on in this paper allows for an initially prescribed scale-dependent baryon density bias which, of course, changes significantly during the non-linear evolution. The essential idea of the growing-mode approach is that, to leading order, the local baryon and CDM fractions are constant in time, and therefore can be absorbed into variations of the masses of Lagrangian fluid elements (these variations are fairly small and constant in

time). This simple trick guarantees that the particle realizations of the baryon and CDM fractions are locally compensated to high precision, meaning that the individual density fractions change without changing the total matter density. This solution thus strongly improves over previous simulations of this kind that were plagued by discreteness errors even on large scales. Furthermore, based on this approach, it is possible to essentially apply standard  $n$ th order LPT results to generate high-order ICs for two-fluid simulations.

Cosmological hydrodynamical simulations usually come in two broad limits: Lagrangian methods, such as smooth particle hydrodynamics or moving mesh techniques – e.g. the widely used GADGET-2/3 (Springel 2005), GASOLINE (Wadsley, Keller & Quinn 2017), AREPO (Springel 2010; Weinberger, Springel & Pakmor 2020), GIZMO (Hopkins 2015), and SWIFT (Schaller et al. 2016) codes – and Eulerian methods which use a spatially fixed mesh that can be dynamically refined – e.g. the widely used ART (Kravtsov, Klypin & Khokhlov 1997), RAMSES (Teyssier 2002), ENZO (Bryan et al. 2014), or NYX (Almgren et al. 2013) codes (but note that a moving mesh can be usually used in both Eulerian or quasi-Lagrangian mode). While the mentioned problems of two-fluid cosmological simulations apply to some degree to all Lagrangian  $N$ -body, smoothed particle hydrodynamics (SPH), or ‘moving mesh’ simulations, the situation for Eulerian ‘fixed mesh’ simulations is arguably even more dire. To achieve comparable accuracy in Eulerian PT, one has to go to significantly higher order than in LPT. Obtaining Eulerian ICs by using LPT-evolved fields in combination with local Lagrangian approximation schemes, as proposed by Hahn & Abel (2011), introduces gravitational non-Gaussianity in the Eulerian density field of the baryons, but is not a consistent PT approach.

To tackle the problem of providing accurate Eulerian ICs for hydrodynamical simulations, we apply in this paper the propagator perturbation theory (PPT; Uhlemann et al. 2019; Rampf, Uhlemann & Hahn 2021) for multiple fluids. This field-level approach, which accurately evaluates LPT-evolved fields at the Eulerian position, has already been used for forward modelling of the matter distribution for Ly  $\alpha$  forest reconstructions by Porqueres et al. (2020), to first order in PPT. Using here second-order PPT to initialize Eulerian hydrodynamical simulations, we are able for the first time to achieve ICs for both (Eulerian) baryons and (Lagrangian) dark matter that are on a similar footing regarding their accuracy in fixed-order PT.

The structure of this paper is as follows. First, in Section 2, we provide a concise summary of the main results from Rampf et al. (2021), as they apply to initial conditions for cosmological simulations of baryons and dark matter. We also quantify the error incurred by neglecting contributions inconsistent with the IC boundary conditions compared to the full cosmological Einstein–Boltzmann solution. In Section 3, we present the numerical simulations we employ in this work, and describe the summary statistics that we use to quantify them. In Section 4, we present results for the non-linear evolution of a collisionless two-fluid  $N$ -body system evolving under self-gravity. We then extend this analysis to full cosmological hydrodynamics plus  $N$ -body simulations in Section 5. We summarize our main results and conclude in Section 6.

Throughout this paper, we adopt cosmological parameters consistent with the PLANCK2018+LSS results (Planck Collaboration VI 2020):  $\Omega_m = 0.3111$ ,  $\Omega_\Lambda = 0.6889$ ,  $\Omega_b = 0.04897$ ,  $\Omega_r = 9.139 \times 10^{-5}$ ,  $h = 0.6766$ ,  $\sigma_8 = 0.8102$ , and  $n_s = 0.9665$ . Einstein–Boltzmann results were computed using the CLASS code<sup>1</sup> (Blas, Lesgourgues & Tram 2011). Our computation of the linear theory

<sup>1</sup>Available from <http://class-code.net/>

growth factor  $D_+ =: D$  always includes the background contribution due to relativistic species; see e.g. Fidler et al. (2017) for details. Note that in this work we use the ‘fixing’ technique of Angulo & Pontzen (2016) – in which the modulus of the white noise Fourier modes is set to unity – in order to suppress the impact of cosmic variance on our results without running large ensembles of simulations, but we consider only single simulations and do not perform the additional ‘pairing’.

## 2 PERTURBATION THEORY IN A NUTSHELL

We consider the evolution of two fluids – specifically CDM and baryons for which we assume a negligible temperature (i.e. a Jeans scale much smaller than the scales of interest) – interacting through gravity in an expanding Universe (parametrized by the cosmic scale factor  $a(t)$ ). We employ co-moving spatial coordinates  $\mathbf{x} = \mathbf{r}/a$  and define peculiar velocities with respect to the co-moving expansion with  $\mathbf{v} = \partial_D \mathbf{x}$ , where  $D$  is the linear growth time in  $\Lambda$ CDM which we also use as the time variable (for simplicity of notation we use  $D := D_+$  synonymously); we suppress temporal dependencies whenever there is no confusion. The governing equations of the system of two fluids for component  $\alpha \in \{b, c\}$  in the zero temperature limit are

$$\partial_D \mathbf{v}_\alpha + \mathbf{v}_\alpha \cdot \nabla \mathbf{v}_\alpha = -\frac{3g}{2D} (\mathbf{v}_\alpha + \nabla \varphi), \quad (1a)$$

$$\partial_D \delta_\alpha + \nabla \cdot [(1 + \delta_\alpha) \mathbf{v}_\alpha] = 0, \quad (1b)$$

$$\nabla^2 \varphi = \frac{1}{D} (f_b \delta_b + f_c \delta_c), \quad (1c)$$

where  $f_b := \Omega_b/\Omega_m$  and  $f_c := 1 - f_b$  are, respectively, the global baryon and CDM mass fractions (neglecting other inhomogeneous contributions), and we have defined

$$g := (D/\partial_t D)^2 a^{-3} = 1 + D^3 \Omega_\Lambda / (11\Omega_m) + \mathcal{O}(D^6). \quad (2)$$

For detailed derivations of the following equations and results, we kindly refer the reader to the companion paper Rampf et al. (2021); for convenience, we provide a brief summary of the key technical steps here as well.

In the following, we first report analytical results in Eulerian coordinates and discuss the validity and limitation of our approach. Results in Lagrangian coordinates as well as for a semiclassical field-based approach are given in Sections 2.3 and 2.4, respectively. Details for the numerical initialization of the involved fields with a linear Einstein–Boltzmann solver are provided in Section 2.5. Our method is fairly distinct from others in the literature, with details given in Section 2.6. Finally, in Section 2.7 we provide ways to effectively include relative velocity effects at linear order.

### 2.1 Analytical findings in Eulerian coordinates

It is convenient (e.g. Schmidt 2016) to rewrite the set of equations (1) in terms of the following ‘sum’ and ‘difference’ variables

$$\delta_m = f_b \delta_b + f_c \delta_c, \quad \theta_m = f_b \theta_b + f_c \theta_c, \quad (3a)$$

$$\delta_{bc} = \delta_b - \delta_c, \quad \theta_{bc} = \theta_b - \theta_c, \quad (3b)$$

where  $\theta_\alpha = \nabla \cdot \mathbf{v}_\alpha$ . Formulated in these new variables, the linearized equations (1) can be combined to

$$\partial_D \theta_m = -\frac{3g}{2D} \left( \theta_m + \frac{\delta_m}{D} \right), \quad \partial_D \delta_m + \theta_m = 0, \quad (4a)$$

$$\partial_D \theta_{bc} = -\frac{3g}{2D} \theta_{bc}, \quad \partial_D \delta_{bc} + \theta_{bc} = 0, \quad (4b)$$

which have the only non-decaying solutions (Rampf et al. 2021)

$$\delta_m = D \nabla^2 \varphi^{\text{ini}}, \quad \theta_m = -\nabla^2 \varphi^{\text{ini}}, \quad (5a)$$

$$\delta_{bc} = \delta_{bc}^{\text{ini}}, \quad \theta_{bc} = 0, \quad (5b)$$

where here and in the following ‘ini’ stands for initial evaluation; see Section 2.5 for details how we generate those initial fields. Using the definitions (3), these solutions imply for the components  $\alpha \in \{b, c\}$  at first order in perturbation theory

$$\delta_\alpha = D \nabla^2 \varphi^{\text{ini}} + \delta_\alpha^{\text{ini}}, \quad \theta_\alpha = -\nabla^2 \varphi^{\text{ini}}, \quad (6)$$

where we have defined  $\delta_b^{\text{ini}} = f_c \delta_{bc}^{\text{ini}}$  and  $\delta_c^{\text{ini}} = -f_b \delta_{bc}^{\text{ini}}$  which, from here on, are sometimes called the compensated constant modes (since  $f_b \delta_b^{\text{ini}} + f_c \delta_c^{\text{ini}} = 0$ ). Here, it is crucial to note that equations (4) remain regular for arbitrarily short times if and only if  $\delta_m \rightarrow 0$  and  $\theta_m \rightarrow -\nabla^2 \varphi^{\text{ini}}$  for  $D \rightarrow 0$ , which by virtue of the definitions (3) implies an initially non-vanishing  $\delta_{bc}^{\text{ini}}$  as well as non-vanishing  $\delta_\alpha^{\text{ini}}$ . Indeed, it can easily be verified that terms such as  $\delta_m/D$  or  $\theta_{bc}/D$  appearing in (4) would otherwise imply quasi-singular behaviour for  $D \rightarrow 0$  (which represents the unperturbed initial state). As mentioned in detail in Section 2.5, being able to initialize the evolution at  $D = 0$  simplifies the boundary analysis tremendously. While only growing modes are naturally selected at  $D = 0$ , the evolution of baryons and CDM can be effectively decoupled from the full multifluid evolution which is governed by the relativistic Einstein–Boltzmann system.

We remark that for initializing single-fluid simulations, the described procedure is standard and here applied to the two-fluid case. We must leave for future work how decaying modes can be consistently incorporated in such schemes (see however Section 2.7), as they are by definition inconsistent with a homogeneous initial state on which the perturbations are imposed.

Before concluding this section, we report for completeness the higher order results in our present model. For this, we begin with the *Ansätze*

$$\delta_m = \sum_{n=1}^{\infty} \delta_m^{(n)}(\mathbf{x}) D^n, \quad \theta_m = -\sum_{n=1}^{\infty} \theta_m^{(n)}(\mathbf{x}) D^{n-1}, \quad (7a)$$

$$\delta_{bc} = \sum_{n=1}^{\infty} \delta_{bc}^{(n)}(\mathbf{x}) D^{n-1}, \quad \theta_{bc} = 0, \quad (7b)$$

where  $\delta_m^{(n)}$  and  $\theta_m^{(n)}$  are coefficients that can easily be determined from the known recursion relations in perturbation theory (see e.g. Bernardeau et al. 2002; Taruya, Nishimichi & Jeong 2018), while the analysis of Rampf et al. (2021) revealed the following recursion relation for the difference density,

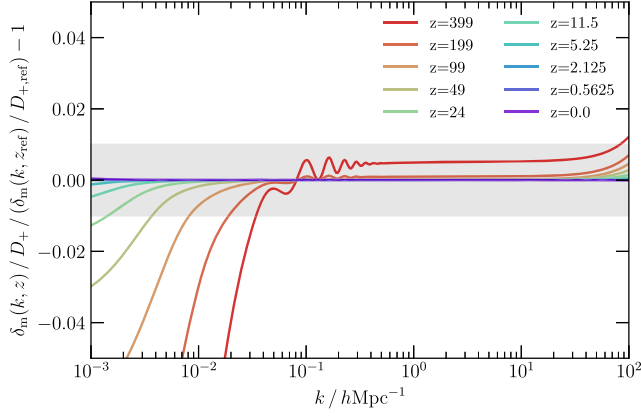
$$\delta_{bc}^{(n)} = \frac{1}{n-1} \sum_{0 < s < n} \nabla \cdot \left[ \delta_{bc}^{(s)} \nabla^{-2} \nabla \theta_m^{(n-s)} \right], \quad (7c)$$

for  $n > 1$ , and  $\delta_{bc}^{(1)} = \delta_{bc}^{\text{ini}}$  for  $n = 1$ . From these recursive relations, it is clear that solutions for  $\delta_b$  and  $\delta_c$  can be easily determined to arbitrarily high orders; for explicit solutions up to third order (see appendix B in the companion paper).

### 2.2 Validation of approximations with CLASS

It is imperative to test the approximations we had to make in order to obtain consistent perturbative results, which effectively ignores decaying modes (see however further below), a finite sound speed of baryons, as well as couplings to relativistic fluid species – except the zeroth-order coupling through the background evolution which we do include. To begin with, assuming the validity of





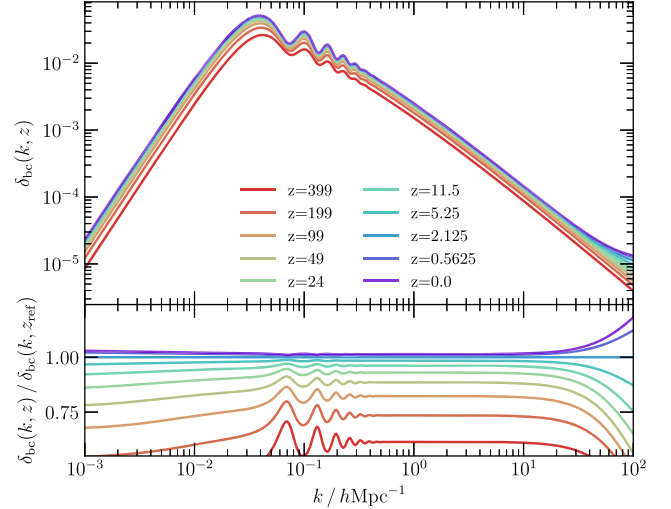
**Figure 1.** Residual scale-dependent evolution of the total matter density  $\delta_m$  from redshift 399 to 0 from CLASS, relative to linear growth with  $D_+$  and the total matter density amplitude at the reference redshift  $z = 2.125$ . After  $z \lesssim 100$ , evolution on small scales (i.e.  $k \gtrsim 0.01 h \text{Mpc}^{-1}$ ) is consistent with a purely growing mode  $\delta \propto D_+(z)$  at much better than one per cent, while residual evolution due to relativistic effects remains on larger (i.e. horizon-scale) scales, and due to finite baryon pressure on small scales  $k \gtrsim 10^2 h \text{Mpc}^{-1}$ . Note that the evolution of  $D_+$  takes a non-zero  $\Omega_r$  into account.

equations (1) implies focusing on a subset of the full set of equations solved in linear Einstein–Boltzmann solvers such as CAMB (Lewis, Challinor & Lasenby 2000) or CLASS (Blas et al. 2011), meaning that any solution based on these equations is already necessarily an approximation. While, clearly, neglecting decaying modes is an important simplification (and one that enables us to carry out the higher order perturbation theory in the first place), it captures the leading effect of the two-fluid system in the late Universe, namely a spatially varying baryon fraction.

To check the validity of the restriction to the growing mode and the compensated constant mode, we show in Fig. 1 the evolution of the total matter overdensity amplitude  $\delta_m$  from redshift 399 to zero, scaled by the linear growth factor  $D_+$  and divided by the amplitude at the reference redshift  $z_{\text{ref}} = 2.125$ .

Clearly, inside the cosmological horizon and for scales  $k \lesssim 10^{-2} h \text{Mpc}^{-1}$  decaying modes have an impact of less than one per cent at all times of interest for ICs for simulations of the late-time Universe. On scales of the horizon and larger, one clearly sees the relativistic effects of horizon growth, as well as at very early times ( $z \gtrsim 199$ ) the effect of radiation drag due to residual ionization on the position of the BAO feature. Horizon scale relativistic effects are well understood and can be easily rectified (see e.g. Brandbyge et al. 2017; Fidler et al. 2017; Zennaro et al. 2017). On the smallest scales one sees the impact of the time-evolution of the baryon temperature-dependent Jeans scale (which in the  $D_+$ -scaled solution is fixed to its co-moving value at the reference redshift).

In Fig. 2, we show similar results for the two-fluid case. The top panel shows the evolution of the compensated mode  $\delta_{\text{bc}}$  between redshifts  $z = 399$  and zero, while the bottom panel displays specifically the impact of ignoring decaying modes and baryonic pressure in its evolution (i.e. assuming  $\delta_{\text{bc}}$  to be constant in time). Evidently, ignoring decaying modes in the evolution of  $\delta_{\text{bc}}$  is justified at fairly late times ( $z \lesssim 10$ ) on almost all scales, while on very small scales ( $k \gtrsim 20 h \text{Mpc}^{-1}$ ) the impact of the evolution of the Jeans scale due to the evolution of the baryon temperature becomes visible. The dominant scale-dependent evolution is again due to a shift of the BAO feature at high redshift, as well as a weak horizon-scale evolution. Assuming that  $\delta_{\text{bc}}$  is constant in time thus introduces



**Figure 2.** Scale-dependent evolution of the density difference  $\delta_{\text{bc}}$  from redshift 399 to zero (top panel), and also relative to the density difference at the reference redshift  $z_{\text{ref}} = 2.125$  (bottom panel). The evolution is consistent with a constant mode only at low redshift, roughly  $\lesssim 15$  per cent at  $z \lesssim 50$ . At higher redshifts, there is a significant contribution due to an additional decaying mode and evolution of the horizon and Jeans scale. On the smallest scales,  $k > 20 h \text{Mpc}^{-1}$ , the impact of Jeans damping is visible.

an almost scale-independent error that is increasingly larger at high redshift. We consider it thus important to tune the reference redshift for the IC generation (see Section 2.5) well in order to capture most accurately the time of interest (i.e. for a Ly  $\alpha$  forest simulation, e.g. it is arguably more accurate to use a reference redshift of  $z_{\text{ref}} \sim 2.5$  rather than zero).

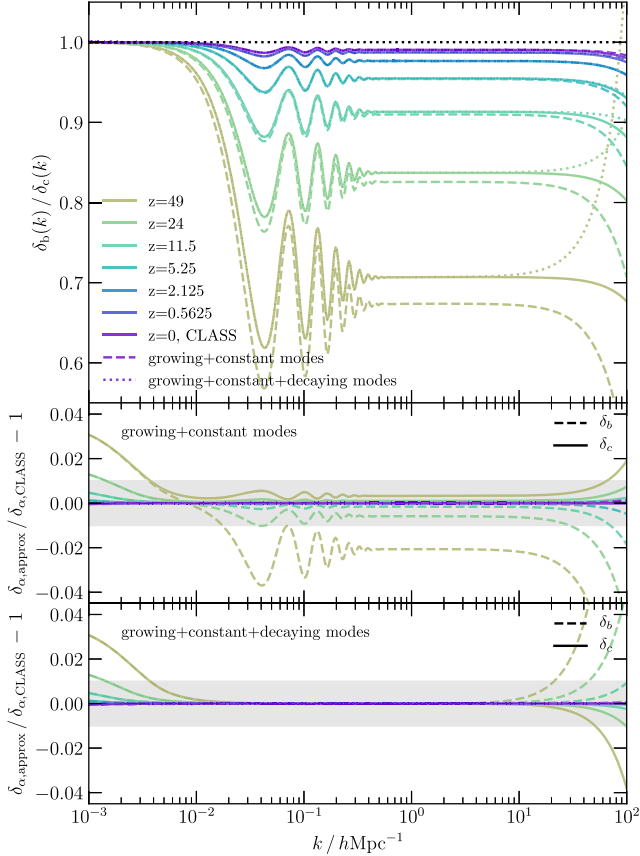
In Fig. 3, we show how well the obtained linear baryon and CDM spectra agree with the full multiphysics evolution. Specifically, we consider two approaches, the first being a two-mode approximation (growing+constant) that forms the basis of our higher order PT, and which obeys equation (6), i.e.

$$\delta_\alpha = D \nabla^2 \varphi^{\text{ini}} + \delta_\alpha^{\text{ini}}, \quad (8)$$

with initial fields obtained as detailed in Section 2.5. The second approach we consider is based on a three-mode approximation (growing+constant+decaying), in which the decaying relative velocity mode is also included (cf. Section 2.7 for generating the additional initial field), i.e. we have then

$$\delta_\alpha = D \nabla^2 \varphi^{\text{ini}} + \delta_\alpha^{\text{ini}} + 2(D^{-1/2} - 1) \theta_\alpha^{\text{ini}}, \quad (9)$$

with  $\delta_b^{\text{ini}} = f_c \delta_{\text{bc}}^{\text{ini}}$ ,  $\theta_b^{\text{ini}} = f_c \delta_{\text{bc}}^{\text{ini}}$ , and  $\delta_c^{\text{ini}} = -f_b \delta_{\text{bc}}^{\text{ini}}$ ,  $\theta_c^{\text{ini}} = -f_b \theta_{\text{bc}}^{\text{ini}}$ . In the top panel, we compare the evolution of  $\delta_b(k)/\delta_c(k)$  from  $z = 49$  to zero as obtained from CLASS (solid lines) including the full linear physics to the two-mode (dashed) and three-mode (dotted) approximations. The relative differences are shown in the bottom panels of the figure. By construction, the solutions coincide at the reference time  $z_{\text{ref}} = 2.125$ . When neglecting the decaying relative velocity mode, the relative error is sub-per-cent for all times  $z \lesssim 24$  on scales  $k \lesssim 10^2 h \text{Mpc}^{-1}$ , but increases rapidly at earlier times. This behaviour is improved when the decaying mode is included, leading to sub-percent agreement over more than four magnitudes in scale at  $z \lesssim 24$ . At late times, the largest error clearly arises from the Jeans scale due to its strong evolution over time, and for simulations aiming at these scales, a more refined discussion, possibly also including temperature fluctuations (e.g. Naoz & Barkana 2005; Naoz, Yoshida & Barkana 2011), might be necessary. Note that the evolution shown



**Figure 3.** Accuracy of the two- and three-mode approximation for the amplitude of baryon and CDM density perturbations. The top panel shows the ratio of the scale-dependent amplitude of baryon to CDM density perturbations as obtained from CLASS (solid lines) and using the growing+constant mode approximation (dashed lines; cf. equation 8), and also including the decaying relative velocity mode (dotted lines; cf. equation 9). The middle (bottom) panel shows the scale-dependent fractional difference in each component between the two-mode (three-mode) approximation and the CLASS amplitude for CDM (solid lines) and baryons (dashed lines). The shaded grey area indicates one per cent deviation. The error in each component is sub-per-cent for redshifts  $z \lesssim 24$  on scales smaller than the horizon. The error on large scales ( $k \lesssim 0.01 h \text{Mpc}^{-1}$ ) is due to neglecting relativistic effects, the error on small scales is due to neglecting finite temperature effects.

here does not include the impact of reionization on the Jeans scale, as this is usually captured in the non-linear simulation directly. Since reionization raises the baryon temperature significantly, the exact evolution of these small scales in the IC backscaling process for simulations that do not explicitly resolve the formation of the reionizing objects is of limited interest.

Clearly, there is some improvement if the decaying mode is included in the linear evolution; however, we do not yet know how to include it self-consistently in a non-linear PT (but see Section 2.7 for a workaround). This means that the impact of streaming velocities (cf. Tseliakhovich & Hirata 2010) cannot be self-consistently included yet. This is arguably the largest drawback and we have to leave the inclusion of a suppression of baryon perturbations on small scales due to the relative motion (which is a decaying mode coupling to the baryon density at second order) for future work. Since this effect is most prominent on the smallest scales (on dimensional grounds it must be close to the baryon Jeans scale), we can safely assume that the approach we present here is accurate for large-scale simulations

that do not include the formation of the very first baryonic objects close to the Jeans scale.

### 2.3 Lagrangian-coordinates approach

Introducing the Lagrangian map for both fluid components  $\alpha \in \{b, c\}$  with  $\mathbf{q} \mapsto \mathbf{x}^\alpha(\mathbf{q}, D) = \mathbf{q} + \boldsymbol{\xi}^\alpha(\mathbf{q}, D)$  with corresponding displacement  $\boldsymbol{\xi}^\alpha$ , equations (1) can be easily transformed into Lagrangian space. For purely growing-mode flows in perturbation theory, the Lagrangian equations of motion take the particularly simple form (Rampf et al. 2021)

$$\nabla_{\mathbf{x}} \cdot \left( \partial_D^2 + \frac{3g}{2D} \partial_D \right) \boldsymbol{\xi}^\alpha(\mathbf{q}, D) = -\frac{3g}{2D} \nabla_{\mathbf{x}}^2 \varphi, \quad (10a)$$

$$\nabla_{\mathbf{x}}^2 \varphi = \delta_m(\mathbf{q}, D)/D, \quad (10b)$$

where the Poisson source is expressed in terms of

$$\delta_m(\mathbf{q}, D) = \frac{1}{\det[\nabla_{\mathbf{q}} \mathbf{x}^m]} - 1, \quad (11a)$$

where

$$\mathbf{x}^m - \mathbf{q} \equiv \boldsymbol{\xi}^m = f_b \boldsymbol{\xi}^b + f_c \boldsymbol{\xi}^c \quad (11b)$$

is the displacement of the combined (or, centre-of-mass) matter fluid. The perturbative solution of the combined matter displacement is well known, which is usually formulated in terms of the following power series

$$\boldsymbol{\xi}^m(\mathbf{q}, D) = \sum_{n=1}^{\infty} \boldsymbol{\xi}^{m(n)}(\mathbf{q}) D^n, \quad (11c)$$

with the first-order coefficient  $\boldsymbol{\xi}^{m(1)} = -\nabla \varphi^{\text{ini}}$  denoting the Zel'dovich approximation, and with the second-order coefficient  $\boldsymbol{\xi}^{m(2)} = -\nabla \varphi_2$  with  $\nabla^2 \varphi_2 = \frac{3}{14} (\varphi_{,ll}^{\text{ini}} \varphi_{,mm}^{\text{ini}} - \varphi_{,lm}^{\text{ini}} \varphi_{,lm}^{\text{ini}})$ . Explicit recursions relations for  $\boldsymbol{\xi}^{m(n)}$  are given by Rampf (2012), Zheligovsky & Frisch (2014), and Matsubara (2015).

Since the combined matter displacement  $\boldsymbol{\xi}^m$  is known to all orders, the Poisson equation (10b) can be easily determined by virtue of (11a). At the same time, having determined the Poisson source allows us to solve the evolution equation (10a) for the component displacement; the solution is particularly simple and reads for the growing modes, for  $\alpha \in \{b, c\}$  (Rampf et al. 2021)

$$\boldsymbol{\xi}^\alpha(\mathbf{q}, D) = \sum_{n=1}^{\infty} \boldsymbol{\xi}^{m(n)}(\mathbf{q}) D^n, \quad (12)$$

which, crucially, must be supplemented with the mass conservation law (cf. with equation 11a)

$$\delta_\alpha = \frac{1 + \delta_\alpha^{\text{ini}}(\mathbf{q})}{\det[\nabla_{\mathbf{q}} \mathbf{x}^\alpha]} - 1. \quad (13)$$

To be specific, although the perturbative solutions for  $\boldsymbol{\xi}^\alpha$  formally agree with the one of  $\boldsymbol{\xi}^m$ , the initial density perturbation  $\delta_\alpha^{\text{ini}}$  appearing in (13) must be taken into account since, as mentioned above, only the inclusion of  $\delta_\alpha^{\text{ini}}$  guarantees the regularity of solutions for arbitrarily short times.

Recently it has been mathematically proven by Zheligovsky & Frisch (2014) and Rampf, Villone & Frisch (2015) that the LPT series for the single fluid converges for realistic random initial conditions in the growing mode, at least for sufficiently short times. Even more recently, numerical convergence studies to very high perturbation orders revealed that LPT converges until the instance of first shell-crossing (and even beyond, although LPT then ceases to be

physically correct; Rampf & Hahn 2021). These findings have direct relevance for the considered two-fluid model with growing mode initial conditions. Indeed, since equations (11c) and (12) coincide, it is clear that the observed convergence behaviour for the single fluid directly carries over to the two-fluid model. This is not surprising as the considered model boils down to transporting initial density perturbations along a shared fluid flow (see the companion paper), which comes with the benefit that for two-fluid growing-mode initial conditions,  $n$ LPT is guaranteed to provide more accurate refinements at successively higher orders (provided shell-crossing has not yet occurred).

We remark that when generating initial conditions for the baryon and CDM fluids,  $\delta_\alpha^{\text{ini}}(\mathbf{q})$  can be explicitly taken into account by varying the particle masses. Further details are provided below (see equation 15).

We also remark that alternatively to the above LPT method for two fluids, one may also incorporate  $\delta_\alpha^{\text{ini}}(\mathbf{q})$  by perturbing the initial positions (see section 5.4 of Rampf et al. 2021 for details). To second order the resulting displacement takes the form

$$\xi_{\text{pert}}^\alpha = D \xi^{\text{m}(1)} + D^2 \xi^{\text{m}(2)} - \nabla^{-2} \nabla \delta_\alpha^{\text{ini}} + \mathbf{F}^{(2)}(\delta_\alpha^{\text{ini}}), \quad (14)$$

where  $\mathbf{F}^{(2)}$  is a vector-valued function that depends quadratically on  $\delta_\alpha^{\text{ini}}$  and can be read off from equation (56) in Rampf et al. (2021). While the corresponding mass conservation then simplifies to  $1 + \delta_\alpha = 1 / \det[\mathbb{1} + \nabla_q \xi_{\text{pert}}^\alpha]$ , we find that spurious discretization errors are excited (see discussion below and Fig. 6), and thus we do not recommend this avenue for the present context.

### 2.3.1 Pre-initial conditions for numerical implementation

Since these LPT results directly translate to initial conditions for Lagrangian methods, such as  $N$ -body, we can simply (pre-)initialize a set of  $N$  particles with the positions and velocities starting from a discrete set of locations  $\mathbf{q}_{i=1\dots N}$ . In this work, we always place the particles initially on a simple cubic (SC) Bravais lattice, so that initial particle positions coincide with the uniform 3D grid on which we computed the velocity and displacement fields using Fourier methods (cf. Michaux et al. 2020).

For Lagrangian hydrodynamic codes, such as GADGET or AREPO that we discuss below, also the baryon fluid elements need to be set up using LPT. A dilemma arises if one wants to construct a force-free set-up of the initial unperturbed CDM+baryon fluid. In the  $\delta_{\text{bc}} = 0$  case this is possible, two individual SC lattices, each carrying  $N$  particles, shifted by half a cell diagonal, with particle masses  $\bar{m}_\alpha = m_p \Omega_\alpha / \Omega_{\text{m}}$  achieve this ( $m_p$  is  $M_{\text{box}}/N$ ). The result corresponds to the CsCl crystal structure. One can in principle also use any other diatomic crystal structure, such as two shifted face-centred cubic (FCC) lattices corresponding to either a NaCl or a Zincblende crystal, depending on the relative shift vector. To evaluate the perturbation fields at the shifted locations, we use a simple Fourier shift of the field<sup>2</sup> in the inverse direction and evaluate at cell centres.

As discussed above, a non-zero  $\delta_{\text{bc}}$  can be realized in two ways: either by using the total mass LPT displacements  $\xi^{\text{m}}(\mathbf{q}, D)$  for all species and perturbing the individual particle masses

$$m_\alpha(\mathbf{q}) = \bar{m}_\alpha (1 + \delta_\alpha^{\text{ini}}(\mathbf{q})), \quad \bar{m}_\alpha := \Omega_\alpha / \Omega_{\text{m}}, \quad (15)$$

or by absorbing this perturbation into a perturbed displacement, i.e. applying equation (14). Either case leads to discretization errors;

<sup>2</sup>Using that for a field shifted by  $\mathbf{x}_0$ , i.e.  $g := f(\mathbf{x} - \mathbf{x}_0)$ , the Fourier transforms obey  $\hat{g}(\mathbf{k}) = \exp[-i\mathbf{k} \cdot \mathbf{x}_0] \hat{f}(\mathbf{k})$ .

however, we find that only with a perturbed displacement that this error has a spurious growing mode, while for the perturbed masses, the discreteness errors are confined to small scales only. For this reason, we adopt the perturbed mass approach in most parts of this paper. We present an analysis of the impact of mass versus displacement perturbations on the power spectrum in Section 4. Even in ‘forward’ simulations (see Section 4.1 for details), it would seem preferable to use perturbed masses instead of displacements to set up the compensated perturbations.

Note that we do not consider the proposed solution of Bird et al. (2020) in this article, which uses glass pre-initial conditions (cf. White 1996) for baryon particles and an SC lattice for DM particles. While seemingly also solving the spurious growth problem, this approach appears to introduce significant additional noise on small scales compared to a Bravais lattice, so that we see no advantage over our approach.

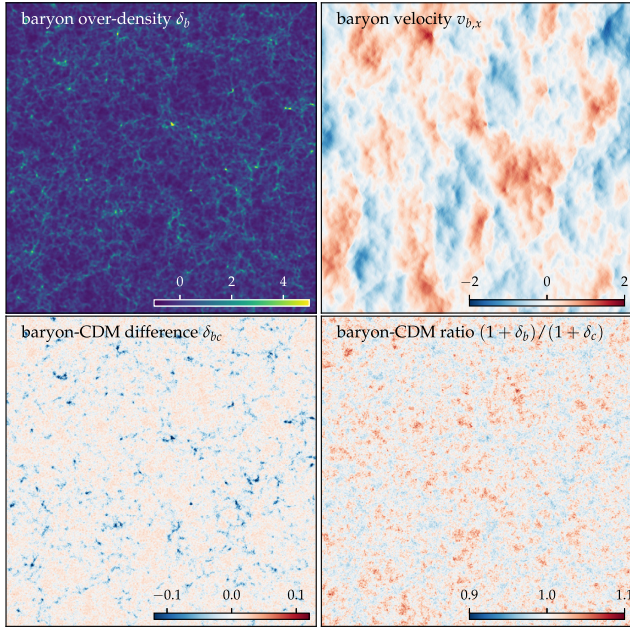
A potential concern in multimass collisionless simulations is the evolution towards mass segregation of  $N$ -body particles in equipartitioned systems (Binney & Tremaine 2008) due to spurious collisional relaxation. We therefore want to emphasize that the mass perturbations introduced by equation (15) are small, independent of the starting redshift (in the fastest growing approximation), and vary on rather large scales. For the set-up we investigate later, i.e. a  $250 h^{-1} \text{Mpc}$  box with  $2 \times 512^3$  particles, the relative fractional variation ( $1\sigma$ ) in particle mass is  $\sim 2.96 \times 10^{-3}$  for the CDM particles and  $\sim 1.69 \times 10^{-2}$  for the baryon fluid elements.<sup>3</sup> In addition, this variation of a few per cent is spatially correlated with a pronounced peak at the BAO scale (cf. Fig. 2), meaning that smaller scale non-linear regions will always have less variation among their particle masses (which can be seen by eye e.g. in the bottom left-hand panel of Fig. 4, where we show the spatial behaviour of  $\delta_{\text{bc}}$ ). Note that furthermore the relative variation in each species is significantly smaller than the difference in particle masses between baryons and CDM in these simulations (which is of order  $\Omega_{\text{b}}/\Omega_{\text{c}} \simeq 1/5.4$ ). With mass differences at the sub-per-cent level for CDM particles, the relaxation time can therefore safely be expected to be much longer than that due to spurious scattering between ‘stars’ and CDM particles (cf. Ludlow et al. 2020).

## 2.4 Propagator perturbation theory

In contrast to Lagrangian methods, cosmological hydrodynamic codes based on Eulerian hydrodynamics, such as the finite-volume codes RAMSES (Teyssier 2002), ENZO (Bryan et al. 2014), or NYX (Almgren et al. 2013), need to start the baryon evolution from the Eulerian density and momentum fields, given at fixed locations discretized in Eulerian space. A possibility to obtain such fields consistent with LPT is by interpolating the fluid elements back to Eulerian grid cells, incurring however the problem of high-quality conservative interpolation. Here, we follow an alternative approach by using propagator perturbation theory (PPT), as proposed by Uhlemann et al. (2019) and extended to two fluids in Rampf et al. (2021), which is able to yield Eulerian density and momentum fields

<sup>3</sup>Note that the mass perturbations have amplitudes of  $\sigma_{\text{m}_b} = \bar{m}_b f_c \sigma_{\text{bc}}$  and  $\sigma_{\text{m}_c} = \bar{m}_c f_b \sigma_{\text{bc}}$  where  $\sigma_{\text{bc}}^2 = (2\pi^2)^{-1} \int_0^{k_{\text{max}}} dk k^2 P_{\text{bc}}(k)$ . If we assume a late-time baryon Jeans scale of order  $k_J \sim 100 h \text{Mpc}^{-1}$  as our  $k_{\text{max}}$ , then  $\sigma_{\text{bc}} \sim 0.026$  for our cosmology. So even at higher resolution, the mass perturbation amounts to at best a few per cent. It would increase of course beyond the (evolving) baryon Jeans scale, but finite temperature effects are beyond the scope of our study here.





**Figure 4.** Eulerian fields at  $z = 8$  obtained with 2PPT as described in Section 2.4: the baryon overdensity  $\delta_b$  (top left),  $x$ -component of the baryon peculiar velocity field  $v_{b,x}$  (top right), the compensated density difference  $\delta_{bc}$  (bottom left), and the ratio of baryon to CDM density fluctuations (bottom right). We show an  $x$ - $y$  slice through the highest density point ( $\delta_{b,\max} \simeq 13$ , north of the centre of the image) for a box of side-length  $250 h^{-1} \text{Mpc}$  computed using a resolution of  $512^3$ .

consistent with LPT without ad hoc interpolation (see also Porqueres et al. 2020).

In the following, we briefly summarize essential equations together with relevant results; further technical details are provided in the companion paper.

#### 2.4.1 Analytical findings in PPT

The central aspect of PPT is to solve for the wavefunction  $\psi_\alpha$  of the fluid components  $\alpha \in \{b, c\}$  whose time evolution is given by the Schrödinger equation

$$i\hbar\partial_D\psi_\alpha = -\frac{\hbar^2}{2}\nabla_x^2\psi_\alpha + V_{\text{eff}}\psi_\alpha, \quad (16)$$

where  $V_{\text{eff}}$  is an ‘effective’ gravitational potential defined in relation to the fluid equations (1). In PPT,  $V_{\text{eff}}$  is treated as an external potential determined by standard perturbation theory. The evolution is expressed through the propagator  $K(\mathbf{q}, \mathbf{x}; D)$  that propagates the initial wavefunction (defined at  $D = 0$ )

$$\psi_\alpha^{\text{ini}}(\mathbf{q}) = \sqrt{1 + \delta_\alpha^{\text{ini}}(\mathbf{q})} \exp\left[\frac{i}{\hbar}\phi_\alpha^{\text{ini}}(\mathbf{q})\right] \quad (17)$$

to the current state at time  $D$  and position  $\mathbf{x}$ , i.e.

$$\psi_\alpha(\mathbf{x}; D) = \int d^3q K(\mathbf{q}, \mathbf{x}; D) \psi_\alpha^{\text{ini}}(\mathbf{q}). \quad (18)$$

At leading order  $V_{\text{eff}} \equiv 0$ , and the solution of the resulting potential-free Schrödinger equation (16) is readily obtained from the ‘free propagator’

$$K_{\text{free}}(\mathbf{q}, \mathbf{x}; D) = (2\pi i\hbar D)^{-3/2} \exp\left[i\frac{(\mathbf{x} - \mathbf{q})^2}{2\hbar D}\right], \quad (19)$$

where the pre-factor guarantees that equation (18) returns  $\psi_\alpha^{\text{ini}}$  for  $D \rightarrow 0$ .

At next-to-leading order, dubbed 2PPT, a time-independent  $V_{\text{eff}}$  becomes relevant and is given by the expression

$$\nabla^2 V_{\text{eff}} = \frac{3}{7} (\varphi_{,ll}^{\text{ini}} \varphi_{,mm}^{\text{ini}} - \varphi_{,lm}^{\text{ini}} \varphi_{,lm}^{\text{ini}}). \quad (20)$$

As shown in the companion paper, the 2PPT propagator reads

$$K(\mathbf{q}, \mathbf{x}; D) = K_{\text{free}}(\mathbf{q}, \mathbf{x}; D) \exp\left[-\frac{iD}{2\hbar} (V_{\text{eff}}(\mathbf{q}) + V_{\text{eff}}(\mathbf{x}))\right]. \quad (21)$$

The semiclassical limits of the free and 2PPT propagators return, respectively, the classical Zel’dovich approximation and the second-order improvement 2LPT. Uhlemann et al. (2019) have shown that the 2PPT results are in fact more accurate than 2LPT since additional symmetries are preserved due to the underlying Hamiltonian structure of (16). Notably, no spurious higher order vorticity is excited.

Having obtained numerical solutions for the wavefunction (see the following paragraph for details), the desired Eulerian fields, e.g. the density  $\rho_\alpha = 1 + \delta_\alpha$  and the momentum density field  $\pi_\alpha = \rho_\alpha \mathbf{v}_\alpha$  for each species, are

$$\rho_\alpha(\mathbf{x}, a) = \psi_\alpha \overline{\psi_\alpha}, \quad \text{and} \quad (22a)$$

$$\pi_\alpha(\mathbf{x}, a) = \frac{i\hbar}{2} (\psi_\alpha \nabla \overline{\psi_\alpha} - \overline{\psi_\alpha} \nabla \psi_\alpha), \quad (22b)$$

where an overline denotes complex conjugation. In principle, one could also extract an effective temperature from the next higher moment, but we will neglect finite temperature effects here altogether and always assume the cold limit on the PT side.

In Fig. 4, we show the baryon density, velocity  $\mathbf{v}_\alpha = \pi_\alpha / \rho_\alpha$ , and the 2PPT density difference  $\delta_{bc}$  for a  $L = 250 h^{-1} \text{Mpc}$  box with  $512^3$  resolution elements at  $z = 8$  (which is much later than the time we would initialize a simulation and was just chosen for illustrative purposes). For further numerical tests of PPT in the single-fluid case, we refer to Uhlemann et al. (2019).

#### 2.4.2 Numerical implementation of PPT

Numerically, the expression for the free propagator (19) is most conveniently evaluated using a discrete Fourier transform (DFT), since the cyclic convolution with the propagator becomes a simple multiplication in Fourier space. Let us therefore assume without change of notation that all spatial coordinates,  $\mathbf{x}$  and  $\mathbf{q}$ , refer to positions on a discrete regular grid with spacing  $\Delta$ , whenever we refer to the numerical implementation. Then, the equivalent statement of (19) at the operational level can be executed using the ‘drift’ operator  $\hat{D}$ , defined through

$$\begin{aligned} \psi_\alpha(\mathbf{x}, a) &= \hat{D} \psi_\alpha^{\text{ini}} \\ &= \text{DFT}_{\mathbf{k} \rightarrow \mathbf{x}}^{-1} \left\{ \exp\left[-i\hbar D_+(a) \frac{k^2}{2}\right] \text{DFT}_{\mathbf{q} \rightarrow \mathbf{k}} \{ \psi_\alpha^{\text{ini}}(\mathbf{q}) \} \right\}, \end{aligned} \quad (23)$$

where  $\mathbf{k}$  denotes a discrete wave vector and  $k$  its modulus. Similarly, to incorporate the aforementioned 2PPT correction, one introduces the ‘kick’ operator

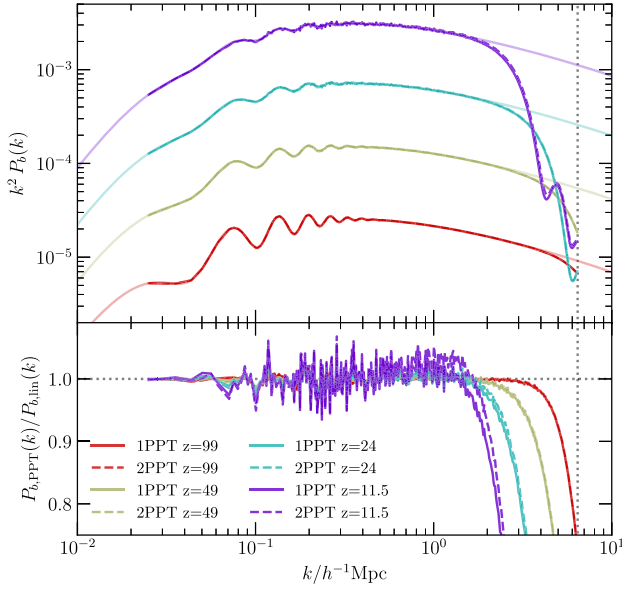
$$\hat{K} := \exp\left[-\frac{i}{\hbar} \frac{D_+(a)}{2} V_{\text{eff}}\right] \quad (24)$$

in real space, which corresponds to a half ‘time step’ in  $D_+$ . The final 2PPT operator evolution equation is given by the single-step leap frog

$$\psi_\alpha(\mathbf{x}; a) = \hat{K} \hat{D} \hat{K} \psi_\alpha^{\text{ini}}. \quad (25)$$

It can be effectively evaluated by performing the drift step in Fourier space and the kick steps in regular space.





**Figure 5.** Evolution of the baryon power spectrum  $P_b(k, z)$  in ‘growing mode’ Eulerian linear theory and in PPT (top panel), and the ratio of the two (bottom panel) for a  $250 h^{-1} \text{Mpc}$  box with  $512^3$  resolution. The finite  $\hbar$ , which is set by numerical resolution (see equation 26), introduces an evolving scale beyond which power is sharply suppressed due to effectively coarse grained dynamics.

#### 2.4.3 The $\hbar$ -parameter

Finally, for numerical implementations of PPT, one chooses a finite  $\hbar$  that is as small as possible in order to be closest to the semiclassical limit. Since we evaluate the propagator using a DFT, the smallest numerically possible  $\hbar$  is determined by the Nyquist–Shannon sampling theorem, which requires that the phase in adjacent sampling points changes by at most  $\pi$ . This implies

$$\hbar \geq \frac{1}{\pi} \max_{\mathbf{q}, d} |\varphi^{\text{ini}}(\mathbf{q}) - \varphi^{\text{ini}}(\mathbf{q} + \Delta \hat{e}_d)|, \quad (26)$$

where  $\hat{e}_d$  is the Cartesian basis vector for the  $d$ th dimension and  $\Delta$  the grid spacing – the expression thus runs over all points and considers the (three) neighbours in three dimensions. We determine  $\hbar$  once we have generated the initial field  $\varphi^{\text{ini}}$  from the input power spectrum, and it depends thus explicitly both on the form of the perturbation spectrum realized in the simulation volume and the grid spacing  $\Delta$ .

The numerically finite value of  $\hbar$  has of course a consequence, namely it acts as an effective coarse-graining scale of the LPT dynamics over phase space cells of size  $\hbar$ . Since  $\hbar$  is determined mostly by the resolution and more weakly by the shape of the perturbation spectrum, the resolution sets the effective temperature of PPT. This manifests itself as a ‘Jeans’-like suppression of power on the smallest scales, similar to what is observed in PT for axion-like particles (cf. e.g. Guth, Hertzberg & Prescod-Weinstein 2015), but note that the scale related to  $\hbar/m$  has a different time dependence in PPT than in the axion-like case (here  $k_{\text{cut}} \propto a^{-1/2} = (1+z)^{1/2}$ ). In Fig. 5, we show the effect on the baryon power spectrum. In the top panel, we show the power spectrum from linear Eulerian PT, restricted to our ‘growing mode’ model, in comparison to the first- and second-order PPT results, measured numerically at different times  $z = 99, 49, 24$ , and  $11.5$ . The effective Jeans smoothing is clearly visible as a sharp power suppression on small scales that increases for later starting times. In the bottom panel, the ratio between Eulerian PT and PPT spectra is shown for a more quantitative comparison. Note that the

power spectrum does however not capture the significant amount of non-Gaussianity that is already present in the fields at the later times. As we show below in Section 5.2, since the suppression affects scales of two to three cells only, and the full non-Gaussian character of LPT is mapped to the Eulerian grid, baryon simulations initialized with PPT evolve quite consistently with those initialized with LPT.

## 2.5 Generating the initial fields – backscaling Einstein–Boltzmann

Traditionally, to generate first-order initial conditions for two-fluid numerical studies for baryons and CDM, one takes the respective fluid variables from a linear Einstein–Boltzmann code at the time when the simulation is to be initialized (see e.g. Yoshida et al. 2003; Hahn & Abel 2011; Angulo et al. 2013; Valkenburg & Villaescusa-Navarro 2017; Bird et al. 2020). By contrast, in simulations for single-matter fields, it is very common to employ the so-called backscaling procedure, which effectively takes the Boltzmann code from very late times, usually around  $z \simeq 0$  (but note that we use  $z = 2.125$  as the pivot redshift in this work), and rescales the respective gravitational potential  $\varphi$  such that the initialized particle configuration at  $z_{\text{ini}}$  has the correct density amplitude. It is important to realize that in a (fictitious) universe with zero radiation content, both approaches reproduce the same initial matter, baryon, and CDM power spectrum. However, the two approaches disagree in a realistic Universe due to the non-trivial evolution of relativistic species that mostly impact the largest scales.

Here, we adopt the backscaling procedure to allow for the initialization of two fluids. This has the advantage that the evolved large-scale power spectra agree, by definition, with the corresponding predictions in general relativity. In addition, also the finite temperature of the baryons is partially included implicitly through the baryon transfer function, determined at the reference time, just not its adiabatic evolution under compression and expansion.

Due to the choice of used boundary conditions (5), which effectively set the decaying modes for the two fluids to zero, only two fields need to be specified initially. One of those fields is the total matter field  $\delta_m$  which relates to the associated gravitational potential according to  $\nabla^2 \tilde{\varphi}(a) = \delta_m(a)/a$ , where  $\tilde{\varphi} = D_+ \varphi/a$ . Since growing-mode initial conditions are obtained from the output of an Einstein–Boltzmann code at sufficiently late times  $a_{\text{ref}}$ , one can write  $\delta_m^{\text{code}}(a_{\text{ref}}) = C_+(x) D_+(a_{\text{ref}})$ . Using these relations, the initial gravitational potential at  $a = 0$  is (cf. Michaux et al. 2020)

$$\varphi^{\text{ini}} = \frac{\nabla^{-2} \delta_m^{\text{code}}(a_{\text{ref}})}{D_+(a_{\text{ref}})} \lim_{a \rightarrow 0} \frac{D_+(a)}{a}. \quad (27)$$

In this paper, we choose  $a_{\text{ref}} = 0.32$  in accordance with our choice of reference redshift  $z_{\text{ref}} = 2.125$ . The other initial field that should be prescribed is the linear difference  $\delta_{\text{bc}}$  which, in the absence of decaying modes, is constant in time. Thus, the amplitude  $\delta_{\text{bc}}$  does not need to be rescaled, and can instead be directly extracted from a Boltzmann code at  $a = a_{\text{ref}}$

$$\delta_{\text{bc}}^{\text{ini}} = \delta_{\text{bc}}^{\text{code}}(a_{\text{ref}}). \quad (28)$$

Of course, having specified both  $\varphi^{\text{ini}}$  and  $\delta_{\text{bc}}^{\text{ini}}$  initially also yields the initial fields for  $\delta_b$  and  $\delta_c$ , as well as  $\theta_b$  and  $\theta_c$ , by virtue of the definitions (3) and boundary conditions (5).

## 2.6 Relation to the forward approach

Previous studies modelling two-fluid dynamics in  $N$ -body simulations (e.g. Yoshida et al. 2003; Angulo et al. 2013; Valkenburg &

Villaescusa-Navarro 2017; Bird et al. 2020) all (to our knowledge) rely on the forward approach, where the Lagrangian displacement and velocity fields are initialized directly with the output of the linear Einstein–Boltzmann code at time  $z_{\text{start}}$  as

$$\begin{aligned} \mathbf{x}_\alpha(\mathbf{q}, z_{\text{start}}) &= \mathbf{q} - \nabla^{-2} \nabla \delta_\alpha^{\text{code}}(z_{\text{start}}), \\ \mathbf{v}_\alpha(\mathbf{q}, z_{\text{start}}) &= \nabla^{-2} \nabla \theta_\alpha^{\text{code}}(z_{\text{start}}). \end{aligned} \quad (29)$$

Within the two-fluid picture, the four code input fields  $\delta_\alpha^{\text{code}}$  and  $\theta_\alpha^{\text{code}}$  may be expressed in terms of the standard growing and decaying modes. Putting aside the decaying modes for the moment (justified at sufficiently late times), one can approximate (29) with

$$\begin{aligned} \mathbf{x}_\alpha(\mathbf{q}, z_{\text{start}}) &\approx \mathbf{q} - D(z_{\text{start}}) \nabla \phi^{\text{ini}} - \nabla^{-2} \nabla \delta_\alpha^{\text{ini}}, \\ \mathbf{v}_\alpha(\mathbf{q}, z_{\text{start}}) &\approx \dot{D}(z_{\text{start}}) \nabla \phi^{\text{ini}}. \end{aligned} \quad (30)$$

where we have used equation (6) to express the growing modes in terms of the input fields in our backscaling approach. Thus, in the forward approach, the initial perturbations  $\delta_\alpha^{\text{ini}}$  are effectively included in equation (30) along with the other modes. Hence, the forward approach is very close in spirit to the ‘displacement perturbation’ approach presented in equation (14), and, therefore, comes with fairly similar numerical challenges as discussed, e.g., in detail by Angulo et al. (2013) and Bird et al. (2020). In Section 4, and in particular in Section 4.1, we present a detailed comparison between ‘displacement perturbed’ and ‘mass perturbed’ ICs.

## 2.7 Including the decaying relative velocity mode at first order

The two-fluid perturbation theory presented above neglects all decaying modes. This leads to small but noticeable differences between the growing mode PT and the evolution in e.g. CLASS, as shown in Fig. 3. While it is still unclear how to incorporate these decaying modes rigorously in higher order LPT/PPT, we note that it is quite straightforward to include the relative velocity  $v_{\text{bc}}$  between baryons and CDM at linear order in LPT. Given the difference fields  $\delta_{\text{bc}}(\mathbf{q}; a_{\text{ref}})$  and  $\theta_{\text{bc}}(\mathbf{q}; a_{\text{ref}})$  from the Einstein–Boltzmann code at the reference time  $a_{\text{ref}}$ , one can modify the mass perturbation and initial particle velocity by rescaling the decaying relative velocity mode from  $a_{\text{ref}}$  to the starting time  $a_{\text{start}}$  of the simulation as

$$\begin{aligned} m_\alpha(\mathbf{q}; a_{\text{start}}) &\rightarrow m_\alpha(\mathbf{q}; a_{\text{start}}) \\ &\quad + 2\bar{m}_\alpha \left[ \left( \frac{D_+(a_{\text{ref}})}{D_+(a_{\text{start}})} \right)^{1/2} - 1 \right] \theta_\alpha^{\text{ini}}(\mathbf{q}) \\ \mathbf{v}_\alpha(\mathbf{q}; a_{\text{start}}) &\rightarrow \mathbf{v}_\alpha(\mathbf{q}; a_{\text{start}}) \\ &\quad + \left( \frac{D_+(a_{\text{ref}})}{D_+(a_{\text{start}})} \right)^{1/2} \nabla^{-2} \nabla \theta_\alpha^{\text{ini}}(\mathbf{q}). \end{aligned} \quad (31)$$

This allows for a first-order correction to the  $n$ LPT two-fluid ICs that restores the agreement between CLASS and LPT ICs. We note however that it is not part of a rigorous perturbative framework, and it is yet unclear how to incorporate a similar fix in PPT. We demonstrate below in Section 4.3 that the inclusion improves, as expected, the agreement with the linear CLASS calculation at high redshift. It appears therefore that, whenever possible, this mode should be included, even though a self-consistent higher order PT is not known to include its non-linear coupling.

## 3 EMPLOYED SIMULATION SET-UP AND SUMMARY STATISTICS

In this section, we briefly summarize the simulation codes as well as the simulations we use in this work. We also discuss the technicalities of the analyses we perform on the simulations.

### 3.1 Simulation methods

In order to compare the performance of Lagrangian and Eulerian cosmological hydrodynamics codes as well as the impact of the collisional nature of baryons versus the effect of gravity alone, we use a multitude of cosmological simulation codes in this work. Specifically, we use the Tree-SPH code GADGET-2 (Springel 2005) for all gravity-only simulations, in which we do not use the SPH part but evolve both species as collisionless zero-temperature fluids.

For more realistic baryon+CDM simulations that evolve baryons hydrodynamically, we use the finite-volume code RAMSES (Teyssier 2002), as well as the moving mesh code AREPO (Springel 2010; Weinberger et al. 2020), respectively, to evolve our initial conditions. Note that this present choice of simulation codes is fairly arbitrary, and an increasingly larger set of codes is becoming freely available to the community. A more stringent code comparison of the results that we sketch in the following sections, which includes other codes, is certainly desirable at some point in the future. Any details of the two codes we use beyond the Lagrangian–Eulerian distinction of the hydrodynamic scheme are not very important for this paper.

At early times, we are in a regime where the finite temperature of the baryons is negligibly small, and pressure effects become important only after shell-crossing and the related formation of shocks and caustics (e.g. Shandarin & Zeldovich 1989). We decidedly do not include additional physics such as radiative cooling or even astrophysical processes such as star formation or energy injection, and switch off UV and other backgrounds. A Lagrangian method therefore has the trivial advantage of (in principle) solving the cold non-linear advection problem with self-gravity more accurately than a Eulerian method prior to shell crossing.<sup>4</sup> We list all simulations employed in this work in Table 1. Our motivation to consider both a Eulerian and a Lagrangian code was to validate the performance of  $n$ PPT against  $n$ LPT, differences between RAMSES and AREPO are of secondary interest to us here.

#### 3.1.1 Collisionless fluids with GADGET-2

Before considering collisional simulations of the baryons (i.e. in the hydrodynamic limit), we will study the purely gravitational, collisionless evolution using a two component  $N$ -body simulation. For these simulations we use GADGET-2 and treat both baryons and CDM as  $N$ -body particles (i.e. we do not use the SPH part of the code). GADGET-2 uses a tree-PM approach to solve for self-gravity and we employ the code specific parameters listed in Table A2.

#### 3.1.2 Eulerian baryons with RAMSES

RAMSES is based on a second order MUSCL (van Leer 1979) finite-volume scheme to solve the equations of ideal hydrodynamics. For

<sup>4</sup>This statement is strictly speaking not correct since pseudo-spectral Eulerian would have also negligible (possibly even superior) advection errors. Pseudo-spectral methods are however not used in cosmological simulations due to their lack of adaptivity and poor convergence at singularities.

**Table 1.** Simulations marked with superscript ‘1’ are used in Section 4 for the study of the purely gravitational, cold, and collisionless evolution of our two-fluid ICs, those with ‘2’ in Section 5 for full  $N$ -body plus collisional hydrodynamic simulations. For the run marked with ‘\*’, we ran also with ICs with perturbed initial positions (using equation 14 to first order); all others ICs use perturbed masses (see equation 15) which is our preferred method; and for those marked with ‘ $\Delta$ ’, we also ran ICs with the decaying relative velocity mode included at linear order (using equation 31). We also refer to the 1LPT/PPT hydro runs ‘†’ as ‘leading order’ (LO) and the 2LPT/PPT runs ‘‡’ as ‘next-to-leading-order’ (NLO). All simulations represent a cosmological volume of side length  $250 h^{-1}$  Mpc.

PT	$z_{\text{start}}$	$N_{\text{part}}$	code
1LPT	49	$2 \times 512^3$	GADGET-2 <sup>1,*</sup> , $\Delta$
2LPT	49	$2 \times 512^3$	GADGET-2 <sup>1</sup>
2LPT	24	$2 \times 512^3$	GADGET-2 <sup>1,\Delta</sup>
3LPT	24	$2 \times 256^3$	GADGET-2 <sup>1</sup>
3LPT	24	$2 \times 512^3$	GADGET-2 <sup>1</sup>
3LPT	24	$2 \times 1024^3$	GADGET-2 <sup>1</sup>
1LPT + 1PPT	49	$2 \times 512^3$	RAMSES <sup>2,†</sup>
2LPT + 2PPT	24	$2 \times 512^3$	RAMSES <sup>2,‡</sup>
1LPT	49	$2 \times 512^3$	AREPO <sup>2,†</sup>
2LPT	24	$2 \times 512^3$	AREPO <sup>2,‡</sup>

the evolution of the collisionless dark matter, RAMSES employs an adaptive particle mesh scheme. We adopt the usual quasi-Lagrangian refinement strategy in which refinements are triggered by the number of  $N$ -body particles and a gas mass threshold based on the initial average baryon mass per cell. In order to achieve a more accurate large-scale integration, we refine the base grid level already when a cell exceeds 4 times its initial (cosmic average) mass, all higher levels at the default threshold of 8 times. All accuracy-related code specific parameters are listed in Table A1.

### 3.1.3 Lagrangian baryons with AREPO

The moving mesh code AREPO is strictly speaking not a fully Lagrangian method, since the mesh does not strictly follow the Lagrangian tracers, and fluxes between cells are taken into account. Prior to shell crossing we are in the advection dominated regime, and we preferentially probe here the Lagrangian aspect of this approach, while arguably the late time deeply non-linear evolution might be more similar to a Eulerian finite-volume method. We list all accuracy-related code specific parameters in Table A2.

## 3.2 Analysis of simulations – power spectra

In this article, we analyse the density statistics of the baryon-CDM two-fluid system mainly through the isotropic (auto) power spectrum  $P_X(k)$  defined as

$$\langle \delta_X(\mathbf{k}) \delta_X(\mathbf{k}') \rangle = (2\pi)^3 \delta_D^{(3)}(\mathbf{k} + \mathbf{k}') P_X(k), \quad (32)$$

where  $X \in \{\text{b}, \text{c}, \text{bc}, \text{m}\}$ , we have  $k := \|\mathbf{k}\|$ , and  $\delta_D^{(3)}$  is the three-dimensional Dirac delta. Numerically, we compute all power spectra using DFTs based on the mass distribution on a regular mesh.

### 3.2.1 Density fields for particles

If the density field is represented by Lagrangian elements (i.e. particles or moving cells), we employ a ‘cloud-in-cell’ (CIC, cf. Hockney & Eastwood 1981) interpolation to a regular grid. To accurately estimate the power spectrum, we use the interlacing technique

proposed by Sefusatti et al. (2016) along with deconvolution with the CIC assignment kernel. We always employ twice the resolution in the DFT mesh compared to the particle resolution, i.e. for  $N^3$  particles, we compute DFTs of size  $(2N)^3$ , to resolve the particle grid itself. Note that we do not correct for shot noise.

### 3.2.2 Density fields for finite-volume cells

For the finite-volume RAMSES simulations, the baryon density is given as a volume average on the adaptively refined oct-tree mesh. In order to evaluate the density field on a regular grid at the same resolution as the particles, it is necessary to deal with cells that are larger than the grid on which one desires to compute the power spectrum. For those cells, that are at a coarser resolution, we use the slope-limited piecewise linear reconstruction used also during the actual RAMSES simulation to ‘refine’ coarse cells to the target resolution. We found that a deconvolution with the cell volume average is necessary to achieve an estimate of the power spectrum that is relatively independent of the resolution used for its estimation (just as with the interlacing and deconvolution in the case of the particles). The volume average is represented by the convolution with the ‘nearest-grid-point’ (NGP, cf. Hockney & Eastwood 1981) kernel

$$W_{\text{NGP}} = (2\pi)^{3/2} \prod_{i \in \{x,y,z\}} \frac{\sin(\frac{\pi}{2} k_i / k_{\text{Ny}})}{k_i / k_{\text{Ny}}}, \quad (33)$$

where  $k_{\text{Ny}}$  is the grid Nyquist wavenumber. Note also that  $W_{\text{CIC}} = W_{\text{NGP}}^2$  for the kernel used to deconvolve the CIC particle projection.

## 3.3 Analysis of simulations – bispectra

To capture the growth of non-Gaussianity in the baryon-CDM two-fluid system, we also consider the (isotropic) component bispectrum  $B_X(k_1, k_2, k_3)$ , defined by

$$\langle \delta_X(\mathbf{k}_1) \delta_X(\mathbf{k}_2) \delta_X(\mathbf{k}_3) \rangle = (2\pi)^3 \delta_D^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_X(k_1, k_2, k_3), \quad (34)$$

with  $X \in \{\text{b}, \text{c}\}$  (but one could also consider  $\{\text{bc}, \text{m}\}$  as well, of course). To simplify the discussion, we only focus on equilateral bispectra here, i.e. where  $k := k_1 = k_2 = k_3$ . We use the PYTHON package BSKIT (Foreman et al. 2020), to numerically compute the bispectrum from the same three-dimensional component density fields as the power spectra described in the previous subsection (i.e. we perform our own CIC deconvolution for the particle density fields and NGP deconvolution for the finite-volume density field). BSKIT is based on the ‘Scoccimarro estimator’ for the bispectrum (cf. Scoccimarro 2000; Sefusatti et al. 2016; Tomlinson, Jeong & Kim 2019).

## 3.4 Analysis of simulations – cumulants

In addition to the bispectra, to quantify the amount of non-Gaussianity present in the simulation, we also consider directly the third and fourth cumulants (i.e. skewness and kurtosis) of the density field, which we define as the dimensionless quantities

$$C_3^\alpha := \langle \delta_\alpha^3 \rangle_s / \langle \delta_\alpha^2 \rangle_s^{3/2} \quad (35a)$$

$$C_4^\alpha := \langle \delta_\alpha^4 \rangle_s / \langle \delta_\alpha^2 \rangle_s^2 - 3, \quad (35b)$$



where  $\langle \cdot \rangle_s$  is the volume average of the respective field, filtered with a top hat filter of scale  $R_s$ . The skewness is related to the bispectrum through

$$\langle \delta_\alpha^3 \rangle_s = \int d^3k_1 d^3k_2 d^3k_3 B_\alpha(k_1, k_2, k_3) \delta_D^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times W(k_1 R_s) W(k_2 R_s) W(k_3 R_s), \quad (36)$$

where  $W$  is the Fourier kernel of the spherical top-hat smoothing window. Note that we have by definition  $\langle \delta \rangle_s = 0$ . The smoothed density fields are obtained from the same mesh of size  $(2N)^3$  as the power and bispectra, and we also perform the deconvolution with the CIC or NGP filter, as described in Section 3.2, prior to applying the top hat filter.

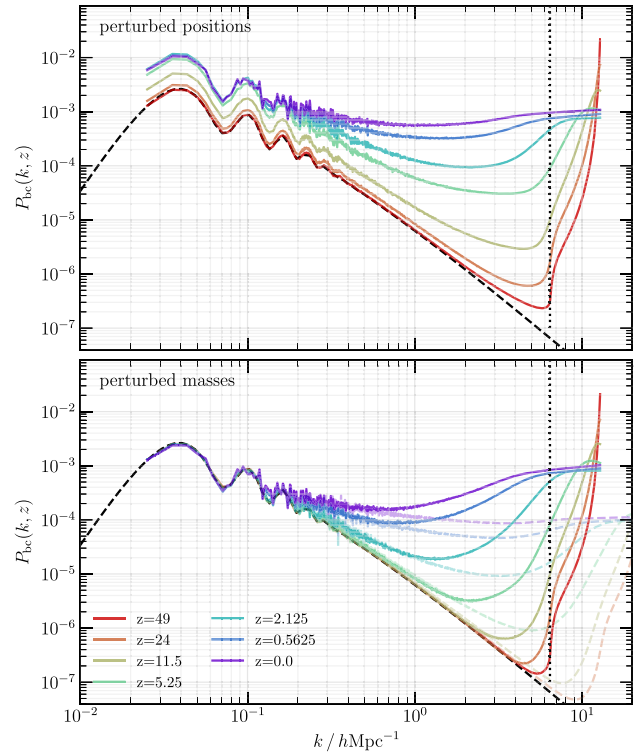
## 4 RESULTS I: PURELY GRAVITATIONAL EVOLUTION IN LAGRANGIAN SIMULATIONS

Before we discuss the performance of our initial conditions evolving the baryons as collisional, it is worth to investigate first the purely gravitational collisionless evolution as a first generalization step of a single fluid cold  $N$ -body simulation to two fluids. We focus exclusively on the evolution of the power spectrum in this section. In these  $2N$ -body simulations, we use identical numbers of baryon and CDM particles with a softening of gravitational forces of 1/20th of the spacing of the initial unperturbed particle lattice, which corresponds to a typical, even slightly conservative, choice in single fluid collisionless  $N$ -body simulations. The simulations that we use for the analysis in this section are listed in Table 1.

### 4.1 Preservation of the compensated mode at linear scales

A first test is the preservation of the compensated mode, which we assume to be constant in time at linear order and thus, it must be preserved in the absence of numerical errors in the linear regime (i.e. early times and/or large scales). It is well known by now that this is not easy to achieve due to discreteness errors. Previous studies have not used our restriction to just two modes, but have instead followed the ‘forward approach’ where the output of the Einstein–Boltzmann code is directly taken at the starting redshift (cf. the discussion in Section 2.6). Without either a gravitational softening of the order of the mean particle separation, or an adaptive softening that arguably achieves this more optimally (O’Leary & McQuinn 2012; Angulo et al. 2013), the relative amplitudes between baryons and CDM do not evolve correctly on any scale, even in the linear regime (cf. also Valkenburg & Villaescusa-Navarro 2017). This is because, effectively,  $\delta_{bc}$  is not compensated at the particle level leading to a slowly growing discrete mode. The low particle-perforce resolution of single-fluid cosmological simulations is usually only possible due to the cold initial conditions, but is known to deviate from the fluid limit on small scales (cf. Joyce et al. 2005; Marcos et al. 2006; Joyce & Marcos 2007). Alternatively, Bird et al. (2020) report that using a mix of grid and glass pre-initial conditions can also suppress the spurious growth of the compensated mode. To our knowledge, there is no theoretical understanding what exactly causes the spurious growing mode, and the exact influence of particle pre-initial conditions on it.

In Fig. 6, we demonstrate the evolution of discreteness effects in the  $\delta_{bc}$  power spectrum for the two ways of setting up ICs. In the top panel, we show  $P_{bc}(k)$  between  $z = 49$  and  $z = 0$  for a simulation with  $2 \times 512^3$  particles initialized with two-fluid 1LPT, where the initial density perturbation  $\delta_\alpha^{(ini)}$  was incorporated into initial particle displacements; as mentioned above and in Section 2.6,



**Figure 6.** Evolution of the baryon–CDM density difference power spectrum  $P_{bc}$  in a gravity-only  $N$ -body simulations, where both baryon and CDM are simulated with  $N$ -body particles with small softening. Simulation results in solid lines (faint dashed lines) with  $2 \times 512^3$  ( $2 \times 1024^3$ ) particles initialized at  $z_{\text{start}} = 49$  ( $z_{\text{start}} = 24$ ) using 1LPT (3LPT) ICs, by suitably backscaling the amplitudes from CLASS at the reference redshift  $z_{\text{ref}} = 2.125$ . Those backscaled power spectra are shown as black dashed lines, before applying the particle realization. The vertical dotted line indicates the particle Nyquist wavenumber. *Top panel:* Simulation results where  $\delta_{bc}^{(ini)}$  is absorbed into particle displacements (using equation 14). We observe the well-known spurious growing mode due to discreteness errors in the particle discretization of the fluids (the discrete  $\delta_{bc}$  is not compensated at the discretization scale, if not sufficiently smoothed by gravitational softening, and therefore grows). *Bottom panel:* Same as above but now  $\delta_{bc}^{(ini)}$  is absorbed into perturbed particle masses (using equation 15). Here, discretization errors are strongly suppressed, and the late-time power asymptotes to a constant noise spectrum beyond  $k_{\text{Ny}}$ . These errors are furthermore vastly reduced when  $2 \times 1024^3$  particles and 3LPT ICs are used.

this approach is implicitly followed in previous two-fluid studies. We see that  $P_{bc}$  evolves on *all* scales due to numerical errors – with the strongest deviations growing at the particle Nyquist wavenumber  $k_{\text{Ny}}$  (indicated by a dotted vertical line) – and linear perturbation theory is *not* recovered even on the largest scales. While we show results only for 1LPT and at starting time  $z_{\text{start}} = 49$ , this result is almost independent of the used LPT order. Since the error is driven by a spurious growing mode, its amplitude is to first order simply determined by the starting time – with earlier starts leading to a larger error. It is also insensitive to the specific choice of softening length, as long as it is appreciably smaller than the mean particle separations. Once the small scales have shell-crossed and collapsed, at  $z \lesssim 2.5$ , the spurious growth is slowed.

In stark contrast, the situation improves dramatically when the initial density perturbation  $\delta_\alpha^{(ini)}$  is incorporated by perturbing the initial particle masses (using equation 15), instead of adding initial displacements. This result is shown in the lower panel of Fig. 6



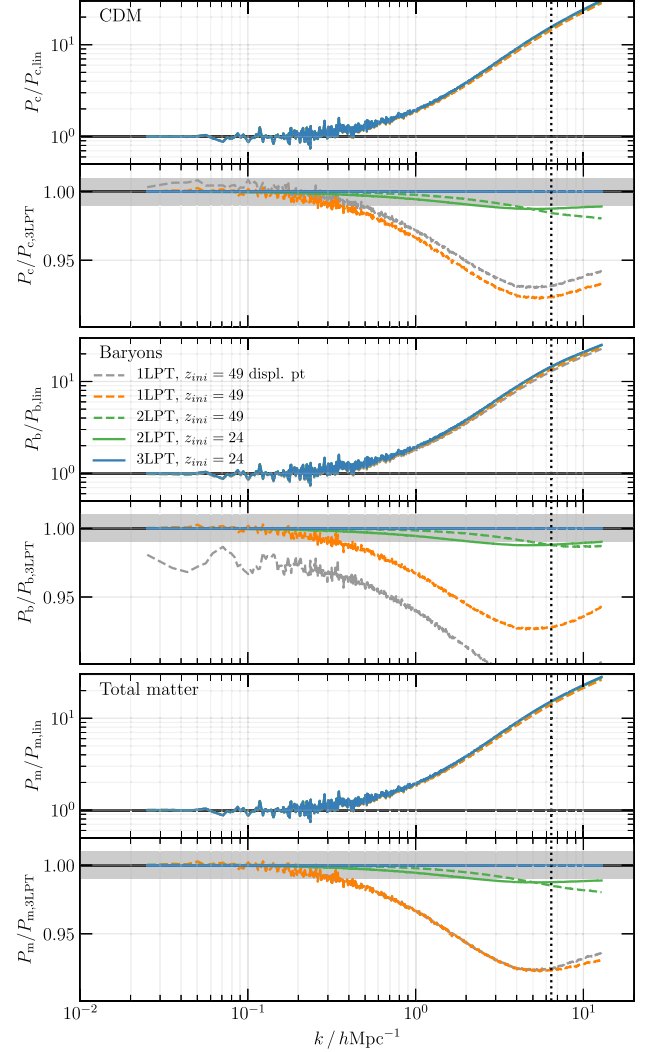
in solid lines with the same particle load and resolution as above, while the results from the higher resolution simulation ( $2 \times 10^{24}$  particles, 3LPT with  $z_{\text{start}} = 24$ ) are shown as faint dashed lines. In this case,  $P_{\text{bc}}$  is exactly constant on large scales and exhibits non-linear growth at intermediate scales, as expected. On the smallest scales, we observe that in all cases the solution asymptotes to a constant power spectrum for  $k \gtrsim k_{\text{Ny}}$  with a scaling inversely proportional to the particle number (as is expected for shot noise in the power spectrum, e.g. Colombi et al. 2009). We expect that this behaviour close to the particle Nyquist wavenumber is thus a combination of residual discreteness errors and shot noise contributions to the measured power spectrum. Note that we have not corrected in any way for shot noise.

Our results compare favourably with those of Bird et al. (2020), who appear to find a stronger discrete evolution on small scales in their mixed glass+SC approach (cf. their fig. 5). Similarly, agreement on large scales can also be achieved using adaptive softening for baryons (e.g. Angulo et al. 2013), which however leads to an artificial suppression of non-linear growth on small scales.

#### 4.2 Impact of the order of perturbation theory

In a next step, we test the impact of the order of  $n$ LPT employed in setting up the ICs on each of the fluid components and on the combined total matter field. To this end, we ran simulations initialized with 1LPT and 2LPT at  $z_{\text{start}} = 49$ , and using 2LPT and 3LPT at  $z_{\text{start}} = 24$  using perturbed masses. In addition, we also use one run with perturbed displacements instead of masses, initialized with 1LPT at  $z_{\text{start}} = 49$ . We show the results at  $z = 2.125$ , a time of relevance for Ly  $\alpha$  studies. At our resolution, this also coincides with the onset of stronger non-linear features in the power spectrum, so that one is still probing also the perturbative regime here. The results of this study are shown in Fig. 7 for the power spectrum  $P_c$  of CDM,  $P_b$  of baryon, and  $P_m$  of the total matter perturbations (top to bottom panels). Each panel is subdivided in two, showing the respective power-spectrum amplitude relative to the linear CLASS prediction, and relative to a reference  $N$ -body simulation. We use the 3LPT,  $z_{\text{start}} = 24$  run as the reference here.

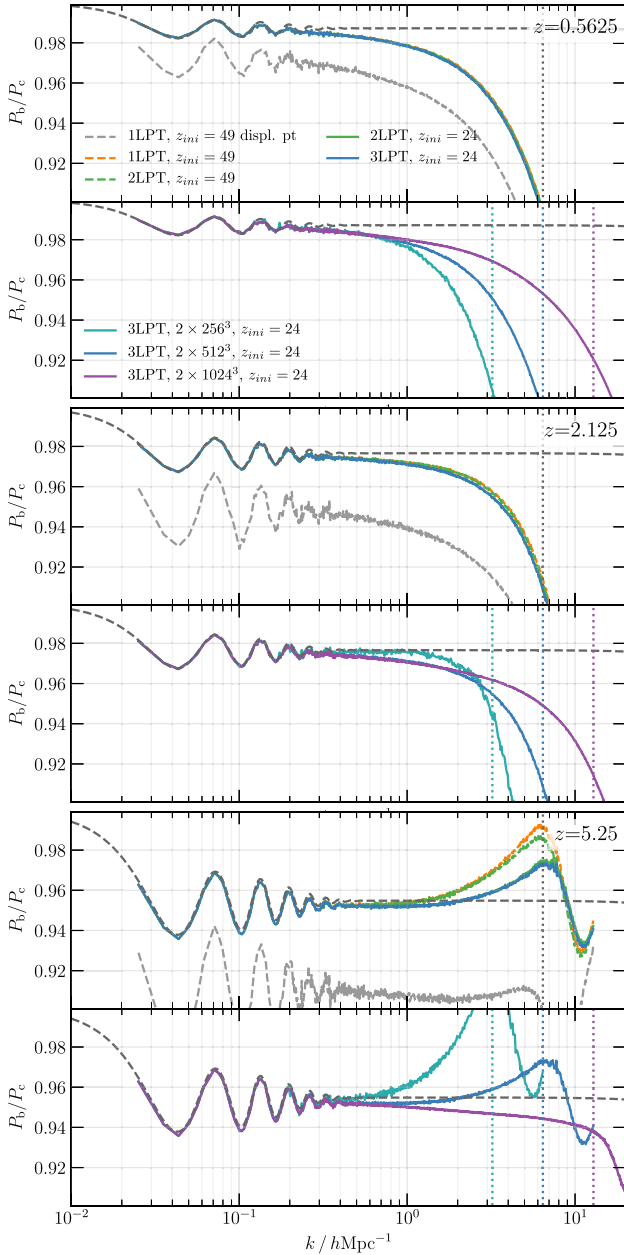
As already discussed in the previous section, without large softening, the run that uses perturbed displacements to absorb the compensated density perturbation (grey lines) shows the wrong growth in each fluid component. The deviation is larger in the baryons than in CDM (arguably due to the particle mass difference); however, there is no error in the total matter spectrum. In the case of perturbed masses, for each component, essentially the same discreteness errors arise as in the single fluid case. The results we find for the runs with perturbed masses are consistent with those of Michaux et al. (2020) for the single fluid case. Essentially, starting later with higher order LPT is preferable, but for our set-up all runs that use at least 2LPT show errors at less than two per cent level at all wave numbers up to the particle Nyquist wavenumber. In contrast, 1LPT is, as already shown 15 yr ago by e.g. Crocce et al. (2006), not particularly accurate. As Michaux et al. (2020) have argued, in principle lower order LPT can be rectified by earlier starts, but then quickly discreteness errors become dominant over LPT-truncation errors, so that this is practically not an option when the goal is to be economic, i.e. to push to the highest wave numbers with the lowest possible number of particles. These discreteness errors can, of course, be suppressed by resorting to a larger force softening, which however also comes at the price of suppressing power on small scales since the force there is no longer Newtonian.



**Figure 7.** CDM, baryon, and total matter power spectra at  $z = 2.125$  for different orders of LPT and starting redshifts in gravity-only simulations. The panels show the CDM power spectrum relative to the linear CLASS solution (top) and relative to a reference run using 3LPT and  $z_{\text{start}} = 24$  (second from top), as well as the same for the baryon power spectrum (third and fourth from top), and for the total matter power spectrum (bottom two panels). Line styles represent different combinations of LPT and starting time: 1LPT (orange), 2LPT (green), 3LPT (blue),  $z_{\text{start}} = 24$  (solid), and  $z_{\text{start}} = 49$  (dashed). All simulations use perturbed masses. For comparison we also show the result of a simulation using displacement perturbations in grey dashed. The vertical dotted line indicates the particle Nyquist wavenumber, and the shaded area indicates a one per cent deviation.

#### 4.3 Convergence of the baryon-CDM ratio and impact of the decaying relative velocity mode

Finally, we also study the ratio of CDM to baryon power spectra,  $P_b/P_c$ , as a function of the order of LPT, starting time, and numerical resolution. We present the results in Fig. 8 at redshifts  $z = 0.5625$ ,  $z = 2.125$ , and  $z = 5.25$  (top to bottom panels). Each panel is again sub-divided in two, the upper one showing the effect of variations of the order of LPT and starting time  $z_{\text{start}}$  of the simulation, the lower showing the effect of varying the number of particles used in the simulation. Starting our discussion with the earliest time  $z = 5.25$ , one notices a pronounced peak at the particle Nyquist wavenumber whose amplitude strongly depends on the starting time



**Figure 8.** Purely gravitational evolution of the relative baryon-CDM power  $P_b(k)/P_c(k)$  from  $z = 5.25$  (bottom two panels) to  $z = 2.125$  (middle panels) to  $z = 0.5625$  (top panels). For each redshift, we show the dependence on the order of LPT and starting redshift in the upper panels (colour and line styles have same meaning as in Fig. 7), and the dependence on the particle resolution in the lower panels (light blue corresponding to  $2 \times 256^3$ , blue to  $2 \times 512^3$ , purple to  $2 \times 1024^3$ ). The dark grey dashed line indicates the CLASS prediction. The dotted vertical lines indicate the particle Nyquist wavenumber.

of the simulation, and the resolution, but also to a much lesser degree on the order of LPT used. It is clear that this peak arises as a consequence of discreteness errors during the early phases of the simulation. Notably, starting the simulation at  $z = 24$  instead of  $z = 49$  reduces this discrete error by almost a factor of two at fixed order of LPT (2LPT in this case). The improvement of 3LPT over 2LPT is minor compared to the discreteness error, but it would be expected that one could push for a later start of the simulation with 3LPT, thus further reducing this specific error (cf. also Michaux et al. 2020).

At intermediate redshift,  $z = 2.125$ , and the resolution we consider here, the pronounced peak at the particle Nyquist wavenumber has disappeared and a physical suppression of  $P_b/P_c$  due to non-linearities becomes visible at small scales. This non-linear suppression is slightly, but visibly, stronger for higher order LPT, for which one expects more accurate non-linear growth. The difference between the different resolutions at fixed order of LPT is much less pronounced than at higher redshift, but it is clear that the low resolution run does not capture the non-linear suppression yet at this redshift. Note that this suppression has been predicted in perturbation theory by Somogyi & Smith (2010) and has also previously been measured in exactly this type of collisionless two-fluid simulations by Angulo et al. (2013) (who however used the ‘forward’ approach to set up their ICs). The suppression essentially means that fluctuations in the baryon fraction become locked once a given scale collapses.

Finally, at the lowest redshift we consider,  $z = 0.5625$ , the situation is similar as at  $z \sim 2$ . The differences due to the order of LPT have further decreased, as was already reported by Michaux et al. (2020) – essentially non-linearities transport power from larger to smaller scales, and differences in LPT are always smaller at larger scales. The resolution of the simulation still plays an important role in setting the suppression of the power spectrum ratio on small scales. It appears as if at this late time, the suppression is converged at scales  $\lesssim k_{\text{Ny}}/3$ .

Last but not least, we also investigate the impact of including the linear decaying relative-velocity mode, as discussed in Section 2.7. The impact on the ratio of baryon to CDM power spectra between redshifts 11.5 and 2.125 is shown in Fig. 9. We show the ratio of baryon to CDM power spectra with and without the inclusion of the relative velocity mode at linear order. It is obvious that the inclusion of the mode clearly improves the agreement at high redshift ( $z \gtrsim 5$ ). At late times ( $z \lesssim 5$ ), the effect is however sub-per-cent on the scales we investigate here. Naturally, the relative velocity mode will have a much stronger effect if scales smaller than those probed here are investigated, where the baryon streaming can have a significant effect on the growth of structures. Since the inclusion of the decaying mode is only carried along at first order, one expects however a strong dependence on the starting redshift since non-linear effects due to the relative velocity mode are not captured in the ICs. In order to improve agreement with the linear evolution, it should however be included when possible (i.e. for particle simulations, since it is not clear yet how to include this mode in the PPT framework).

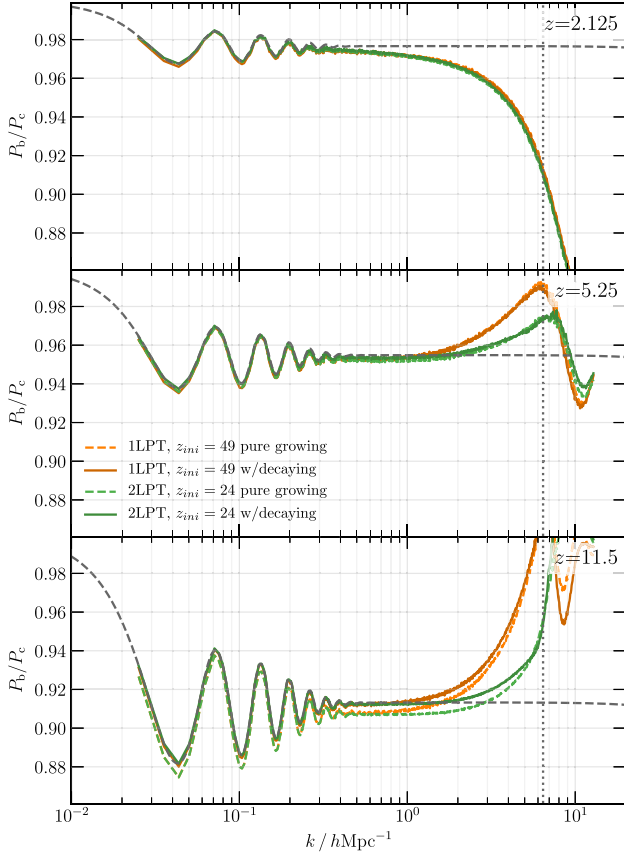
## 5 RESULTS II: MIXED CDM+BARYON SIMULATIONS

After having considered the collisionless, purely gravitational evolution in the previous section, we now turn to fully hydrodynamic simulations using the Eulerian code RAMSES and the moving mesh AREPO codes. The simulations analysed in this section are listed in Table 1. We first present the evolution of density power spectra and compare the results between the two codes and with the purely gravitational evolution. We then analyse in detail the improvements brought about by higher order PT, which is known to be more prominent in higher order correlations (e.g. Munshi, Sahni & Starobinsky 1994).

### 5.1 Power spectrum evolution

#### 5.1.1 Evolution of component spectra

In Fig. 10, we show the evolution of various density power spectra obtained with the RAMSES (left-hand panels) and AREPO (right-hand panels) codes for initial conditions for the same  $250 h^{-1} \text{Mpc}$  box as



**Figure 9.** Effect of the decaying relative velocity mode on the purely gravitational evolution of the relative baryon–CDM power  $P_b(k)/P_c(k)$  from  $z = 11.5$  (bottom panel) to  $z = 5.25$  (middle panel) to  $z = 2.125$  (top panel). For each redshift, we show lines obtained with 1LPT and  $z_{\text{start}} = 49$  (orange) and with 2LPT and  $z_{\text{start}} = 24$  (green) for ICs that include growing modes only (dashed, brighter hues) and that also include the relative velocity decaying mode (solid, darker hues). The grey dashed line indicates the CLASS prediction. The dotted vertical lines indicate the particle Nyquist wavenumber. Including the relative velocity mode improves the agreement at high redshift but has only sub-per-cent effect on the power spectrum on these scales at low  $z$ .

in the previous section. Both simulations use the same initial number of resolution elements. The top-most panels show the evolution of the component, i.e. baryon and CDM, power spectra  $P_\alpha(k)$  between the initial time  $z = 24$  and the final output we considered at  $z \simeq 0.56$ , followed by the evolution of the ratio  $P_b/P_c$  in the second panel from top. We observe that at the initial time  $z = 24$ , the 2PPT initial conditions used for the baryons with RAMSES are slightly smoother than the Lagrangian field used for AREPO, with a suppression of baryon power on scales of about two to three top grid cells (see also Fig. 5 and the related discussion, where we showed how this depends on starting time and resolution). This suppression persists during the quasi-linear stages of the evolution, but we find that once non-linear structure has formed at  $z \lesssim 2.5$ , the non-linear spectra obtained with RAMSES and AREPO agree very well. In the power ratio  $P_b/P_c$ , we observe about 1 per cent deviation of the RAMSES results from linear theory on large scales, while AREPO follows the linear theory perfectly on large scales. This is arguably due to advection errors in RAMSES causing a slight diffusion that affects even large scales, but note that this error is at the per cent level only.

### 5.1.2 Evolution of the difference spectrum

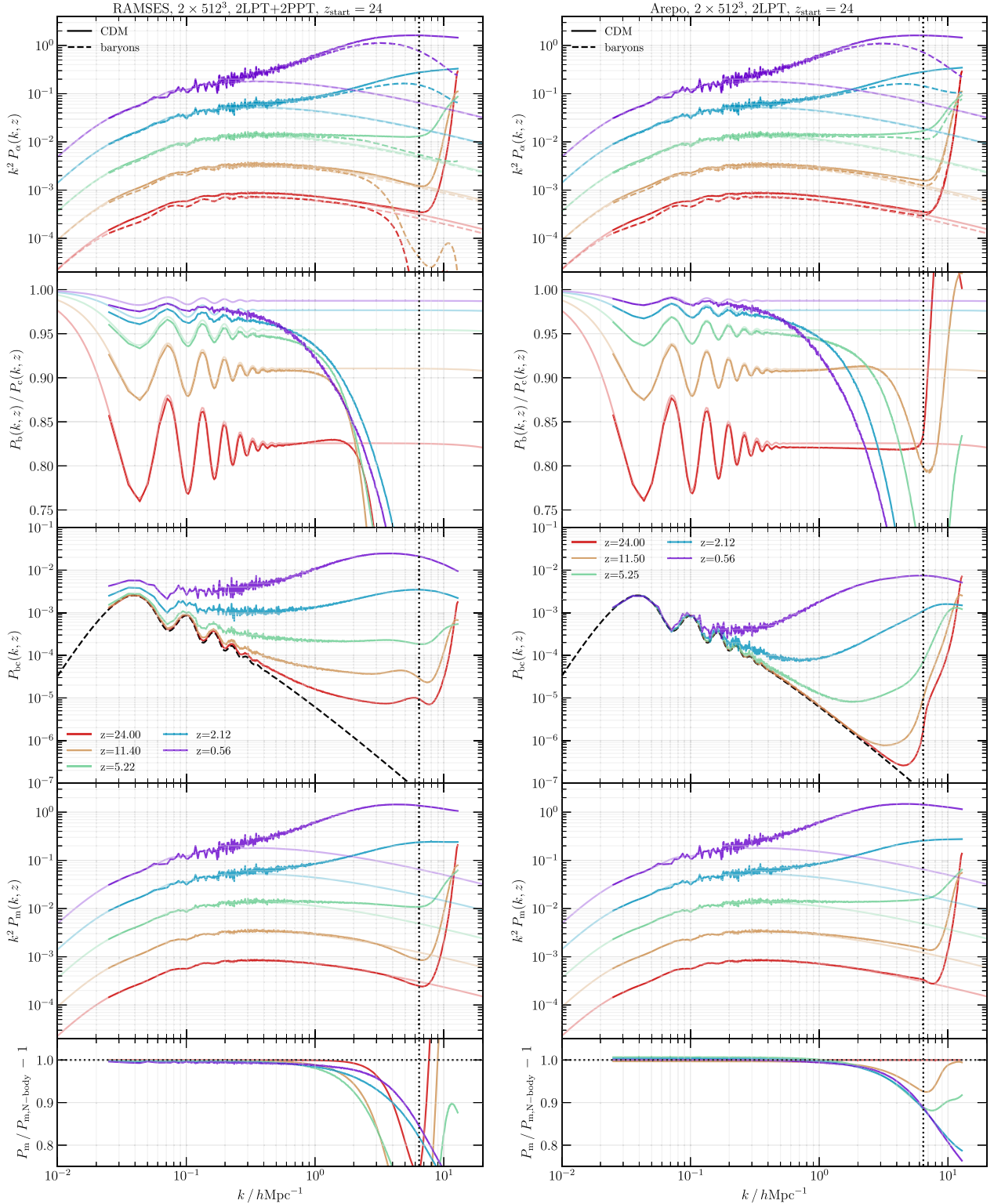
As expected in the presence of collisional processes, the power spectrum of the baryon–CDM density difference  $P_{bc}$ , which is shown in the third panel from the top in Fig. 10, shows a larger amplitude on small scales compared to the evolution in the collisionless simulation shown in Fig. 6. Collisional processes, leading to an isotropic pressure rather than anisotropic stress after shell-crossing (cf. Buehlmann & Hahn 2019), as well as entropy production act to decouple the baryon evolution from the collisionless CDM evolution. Similarly to the collisionless simulations, numerical errors are also particularly pronounced in  $P_{bc}$ , the initial two to three grid-scale power suppression in PPT/RAMSES at  $z = 24$  is visible as a deviation from linear theory peaking at the root-grid Nyquist wavenumber. During quasi-linear evolution, the overall level of error very close to the Nyquist scale is similar for both codes, but extending to larger scales in RAMSES. Also, the per cent level growth error at late times compared to linear theory visible for RAMSES in the  $P_b/P_c$  ratio is visible here as a late time spurious growth.

### 5.1.3 Total matter spectrum and baryon response

Finally, in the second panel from the bottom of Fig. 10, we show the evolution of the total matter spectrum together with its relative deviation from the total matter power spectrum obtained from the purely collisionless two-fluid evolution. While on large scales, an effective pressure arising from shell-crossed shocks plays no role, and the collisional simulations agree perfectly with the collisionless simulation (i.e. within expected numerical errors), we see that at late times  $z \lesssim 2.5$ , both collisional codes predict a fairly rapid suppression in the total matter spectrum at scales  $k \gtrsim 1 h\text{Mpc}^{-1}$ . To compare the collisionless and collisional power spectra, which were output by the respective codes at slightly different snapshot times, we simply rescaled to the output times of RAMSES and AREPO using the linear theory growth factor  $D_+$ . We note that, at this resolution (which is fairly low compared to state-of-the-art galaxy formation simulations), RAMSES predicts a slightly larger suppression  $\sim 15$  per cent at  $k_{\text{Ny}}$  at  $z = 0.56$  compared to AREPO with  $\sim 10$  per cent at  $k_{\text{Ny}}$ , which is possibly due to the smoother ICs and/or advection errors. The suppression shape predicted by the AREPO simulation has a nearly universal shape across all redshifts. We caution that we expect that the precise evolution of the power suppression depends crucially on additional physics such as cooling, UV backgrounds and AGN feedback in more realistic simulations that attempt to model also astrophysical processes. In that sense it is somewhat surprising that the suppression we observe here is quantitatively not all too different at low  $z$  from the range found across state-of-the-art galaxy formation simulations (cf. Chisari et al. 2019, in particular their fig. 3). A suppression that is stronger or affects larger scales, as has been observed in some simulations and is included in recent baryon response models applied to collisionless simulations (cf. e.g. Huang et al. 2019; Schneider et al. 2019; Aricò et al. 2020) clearly requires substantial injection of energy into the baryons beyond just offsetting radiative cooling losses. Due to the absence of additional physics, the physical suppression scale in our simulations is set by gravity alone.

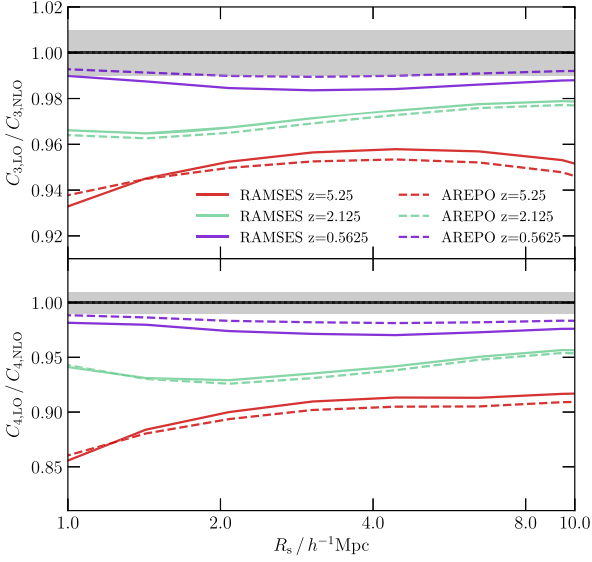
## 5.2 Evolution of non-Gaussianity – cumulants and bispectrum

Finally, we quantify the evolution of non-Gaussianity due to gravitational instability in the density fields. It is well known that convergence in higher order statistics depends sensitively on both



**Figure 10.** Comparison of the evolution of a mixed CDM+baryon fluid in two commonly used cosmological hydrodynamical codes: RAMSES (left-hand panels) and AREPO (right-hand panels). From top to bottom, the figure shows the evolution of the individual baryon and CDM matter power spectra,  $P_b$  and  $P_c$ , the ratio of these two spectra,  $P_b/P_c$ , the spectrum of the difference,  $P_{bc}$ , the total matter power spectrum  $P_m$ , and the ratio of the simulation matter power spectrum to that from a two-fluid collisionless  $N$ -body run. Light lines indicate the linear ‘growing-mode’ PT results at the precise output times of the snapshots. The vertical dotted line indicates the particle Nyquist wavenumber. Note that the initial suppression of baryon power on small scales in the RAMSES simulation is due to the PPT approach (cf. Fig. 5).





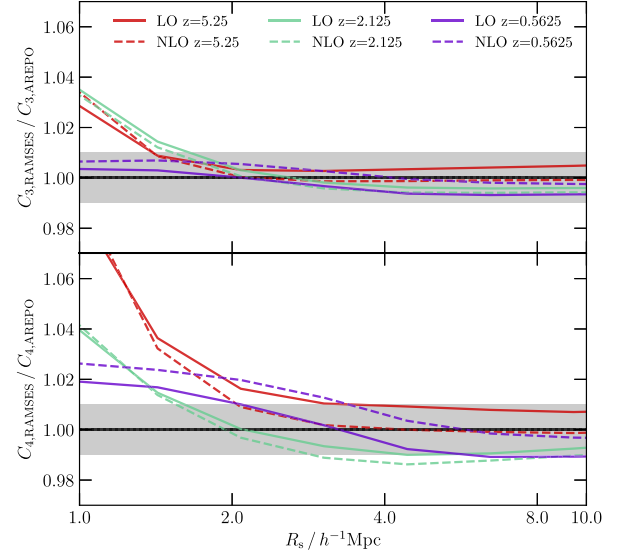
**Figure 11.** Dependence of the third and fourth cumulants,  $C_3$  (top panel) and  $C_4$  (bottom panel) of the baryon density field of the order of the PT used to set-up the initial conditions. We show the ratio of the cumulants obtained from LO simulations (i.e. 1LPT/1PPT, initialized at  $z_{\text{start}} = 49$ ) to those obtained from NLO simulations (i.e. 2LPT/2PPT, initialized at  $z_{\text{start}} = 24$ ) as a function of the scale of the applied top hat filter  $R_{\text{TH}}$ . The results show the transient behaviour well known for collisionless simulations also for hydrodynamic simulations: lower order PT leads to an underestimation of non-Gaussianity, particularly at higher redshift. The improvement brought about by going from LO to NLO is similar at all scales for Lagrangian and Eulerian simulations.

the starting time of simulation and the truncation order in the PT expansion (e.g. Munshi et al. 1994; Scoccimarro 1998; Crocce et al. 2006; Michaux et al. 2020). Due to the absence of higher order PT schemes for baryon-CDM simulations, such tests have not been made in the two-fluid case to our knowledge. Here, we specifically consider two summary statistics: (1) the third and fourth cumulants (i.e. skewness and kurtosis) of the baryon density field as a function of filtering scale covering the range between the non-linear and the linear scales of our simulations, and (2) the equilateral bispectrum, i.e. the harmonic version of the three-point correlation function, of the baryon and the CDM density field.

### 5.2.1 Cumulant statistics of the baryon density field

To quantify the improvement brought about by going to higher order PT when generating initial conditions, we first study the influence of first versus second order PT on the one-point statistics of the smoothed baryon density field. Specifically, we investigate ratios of the  $C_3$  and  $C_4$  cumulants, as defined in equation (35), between the LO runs (i.e. using either 1LPT or 1PPT with a starting time of  $z_{\text{start}} = 49$ ) and the NLO runs (which use 2LPT/2PPT with a starting time of  $z_{\text{start}} = 24$ ). Note that we vary both the starting time and the order since the NLO start is too late for first-order PT. An even earlier start might improve agreement, but usually comes at the cost of larger numerical errors in the solution (see e.g. Michaux et al. 2020).

The results are shown for  $C_3$  in the top panel, and for  $C_4$  in the bottom panel of Fig. 11 for simulations run with both RAMSES (solid lines) and AREPO (dashed lines). In agreement with previous studies for the total matter density field (e.g. Crocce et al. 2006, see in particular their fig. 5), we find that first-order ICs underestimate



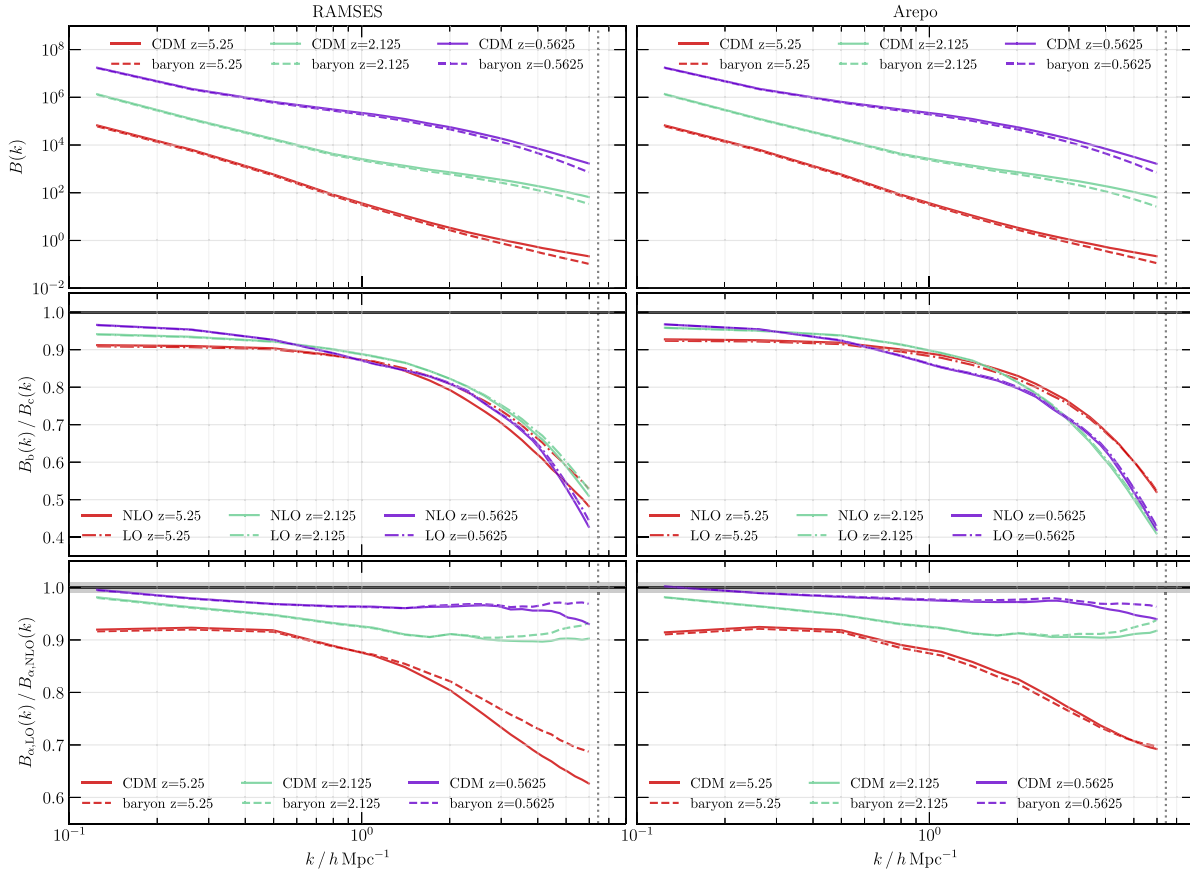
**Figure 12.** Ratio of the third and fourth cumulants  $C_3$  (top panel) and  $C_4$  (bottom panel) of the baryon density field between the RAMSES and the AREPO simulations as a function of the scale of the applied top hat filter  $R_{\text{TH}}$ . Results from the LO runs (1LPT/1PPT,  $z_{\text{start}} = 49$ ) are shown as solid lines, those from the NLO runs (2LPT/2PPT,  $z_{\text{start}} = 24$ ) are shown as dashed lines for three output times indicated by the different colours. Agreement between the two codes is at the few per cent level.

the degree of non-Gaussianity also in the baryon field. Errors are systematically larger at higher redshift ( $z \gtrsim 2$ ), approaching more than 10 per cent on all scales at  $z \gtrsim 5$  for our simulation set-up. The overall improvement brought about by going from LO to NLO is virtually identical (within better than one per cent) for both RAMSES and AREPO, which is a clear validation of our PPT approach for Eulerian finite-volume methods. It is particularly interesting to see that advection errors of the Eulerian code, which are relatively prominent in the power spectrum on small scales, are not reflected here.

To compare more accurately the results obtained with the two codes, we show in Fig. 12 explicitly the ratio of the cumulants measured in the RAMSES and the AREPO simulations. While we find that both  $C_3$  and  $C_4$  agree when smoothed on large scales to about one per cent, RAMSES shows a consistently larger amount of non-Gaussianity on small scales  $R \lesssim 2h^{-1}\text{Mpc}$ . One can speculate that on small scales particle noise or poor sampling in underdense regions in the Lagrangian code could impact these results, and/or that advection errors in the Eulerian code could lead to decreased variance, while possibly higher order cumulants are better retained, so that these normalized cumulants appear boosted. While we only note this systematic discrepancy here, it is certainly worthwhile to investigate its origin and detailed dependence on resolution and/or code parameters in future work in light of precision predictions of the Ly  $\alpha$  forest.

### 5.2.2 Bispectrum evolution in the two-fluid system

Finally, we analyse the evolution of the baryon and CDM density bispectrum, and its dependence on PT order and simulation code. We focus here for simplicity only on the equilateral bispectrum. Our main results are shown in Fig. 13, which presents results for the NLO simulations obtained with RAMSES in the left-hand panels, and for AREPO in the right-hand panels. The CDM (solid line) and baryon (dashed line) bispectra are shown for our usual three



**Figure 13.** Equilateral bispectra of the baryon and CDM density fields for the RAMSES (left-hand panels) and AREPO (right-hand panels) simulations. *Top panels:* The bispectra at  $z = 5.25, 2.125$ , and  $0.5625$  for the CDM (solid lines) and baryon (dashed lines) density fields. *Middle panels:* Ratio between the baryon and CDM bispectra for the three different redshifts for the LO (i.e. 1LPT/1PPT,  $z_{\text{start}} = 49$ ) and NLO (i.e. 2LPT/2PPT,  $z_{\text{start}} = 24$ ) simulations. *Bottom panels:* Ratio of the LO and NLO bispectra for CDM (solid) and baryons (dashed) for the three redshifts.

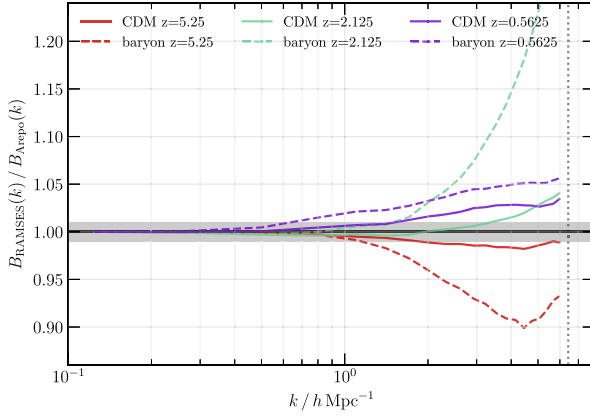
redshifts in the top most panel. Similarly as for the baryon and CDM power spectra, one observes that the baryon bispectrum is suppressed relative to the CDM one, particularly so on small scales. To make this more explicit, the middle panels of the figure show the ratio of the baryon to CDM bispectra,  $B_b(k)/B_c(k)$  revealing a relatively time independent suppression comparable in amplitude to the suppression in the power spectrum (cf. second panels from top in Fig. 10). Note that for the power ratio  $P_b(k)/P_c(k)$ , we observed a significant evolution of the suppression in the AREPO run, growing from smaller to larger scales over time.

We also compare the bispectra obtained from the NLO IC runs against those with only LO (first-order) initial conditions in the bottom panels of Fig. 13. The impact on the component bispectra is as one would expect similar to the single fluid case, where one observes also a strongly suppressed bispectrum at high redshifts for first order ICs except when very early starting times are used to initialize the simulations (e.g. Crocce et al. 2006; McCullagh, Jeong & Szalay 2015) – the well-known transient. The bispectra from early start and with low-order ICs are however more impacted by the accumulation of discreteness errors (Michaux et al. 2020), which one would expect to be even more dramatic in the baryon component if a diffusive Eulerian scheme is used. Since we do not disentangle starting redshift and PT order, discreteness and truncation errors are somewhat convoluted here. Comparing the AREPO and RAMSES results for NLO versus LO ICs, one can still draw a few interesting conclusions (and we leave more thorough investigations

to future work): The LO/NLO bispectrum ratios for baryons are very similar for AREPO and RAMSES, on all scales, showing that numerical diffusion due to the longer integration time in LO versus NLO is not important. In the particle component, we see however a stronger suppression in RAMSES, which could be a consequence of the effectively lower force resolution at early times due to the AMR scheme. At the same time, particle noise might induce spurious effects. Note that in the absence of exact solutions or at least a full convergence study, all such conclusions are speculative.

In summary, the excellent agreement between the improvement between LO and NLO for both 2LPT and 2PPT ICs clearly validates the PPT approach for higher order ICs for baryons for Eulerian codes. Foreman et al. (2020) have previously studied the baryon bispectrum in a full ‘physics’ galaxy formation simulation including cooling and AGN feedback, however starting from ICs where baryons trace CDM perfectly. It will be interesting to compare these results with our adiabatic runs as well as with a more realistic astrophysical simulation that takes our new ICs into account.

As a last comparison, we show the explicit ratios of bispectra between the NLO IC simulations performed with RAMSES and with AREPO. Due to the slight difference in snapshot times, we divide all bispectra by the value in the first  $k$ -bin. The result is shown in Fig. 14. We caution that in order to establish which results are converged, one would have to conduct a rigorous resolution test. Here, we are more interested at the level of typical differences due to the different methods. We find for the earliest snapshot, at  $z =$



**Figure 14.** Ratio of the equilateral bispectra of the baryon and CDM density fields between the RAMSES and AREPO simulations as a function of the triangle scale  $k$ . We show only results from the NLO simulations, indicating a significantly boosted bispectrum amplitude on small scales in RAMSES compared to AREPO. Note that in this plot, we have re-normalized the respective amplitudes so that the ratio is unity in the smallest  $k$ -bin to suppress contribution from slightly different output times and growth in the two codes (cf. Fig. 12).

5.25, a  $\sim 2$  per cent suppression of the CDM bispectrum close to the Nyquist wavenumber in the RAMSES run compared to AREPO, and a much larger suppression of the baryon bispectrum by up to 10 per cent close to the Nyquist wavenumber in the baryon bispectrum. This difference is consistent with a similar suppression visible also in the power spectrum and owed to an effectively smoother IC in the baryons and arguably also additional advection errors. What is more curious is that at late times, this is reversed, and we observe a higher amplitude in the bispectrum in the RAMSES run, particularly so for baryons, where the effect is very significant at intermediate redshifts  $z \sim 2$ . This behaviour is consistent with a similar difference in the cumulants when the baryon density field is smoothed on relatively small scales reported above (cf. Fig. 12) – and we have already speculated about possible reasons there. Still, the results for the RAMSES and AREPO appear consistent at the few per cent level, which is remarkable in light of the very different approaches to evolving baryons that these codes adopt. The often mentioned advection errors incurred by Eulerian schemes seem to have much less influence on higher order statistics.

## 6 SUMMARY AND CONCLUSIONS

With the increasing precision of current and upcoming cosmological observations, the long-standing problem of how to generate accurate initial conditions for cosmological simulations that model the distinct non-linear evolution of both CDM and baryons has become more pressing. In this paper, we present the numerical implementation of a novel approach to set up initial conditions for two-fluid cosmological simulations, and validate our implementation and its performance based on various summary statistics. In brief, our new approach

- (i) provides higher order Lagrangian ( $n$ LPT) ICs for two gravitationally coupled fluids in the cold limit, by restricting to a generalization of the ‘growing-mode’ solutions of standard LPT;
- (ii) applies a field-theoretic approach to LPT to initialize Eulerian simulations using PPT (Uhlemann et al. 2019; Rampf et al. 2021);
- (iii) relies on backscaling the late-times input fields to initialization time (instead of a forward approach), thereby improving

accuracy at low  $z$ , and having sub-per-cent errors at  $z \lesssim 24$  on most scales of interest; and

(iv) prevents the typical problematic excitation of spurious growth well known for two-fluid  $N$ -body systems, even when very small gravitational softening is used.

The theoretical foundations are presented in the companion paper Rampf et al. (2021), and summarized in Sections 2.1–2.4 in this article. All methods are implemented in the IC generator MONOFONIC MUSIC-2 (i.e. the single resolution, non-‘zoom’ version of MUSIC-2), which we make publicly available.<sup>5</sup>

We validate the quality of our initial conditions in two steps, first by considering the purely gravitational collisionless evolution of the two-fluid system using the  $N$ -body method (specifically GADGET-2, Section 4), and in a second step using two commonly used cosmological  $N$ -body+collisional hydrodynamics codes, specifically the Eulerian RAMSES code and the AREPO moving mesh code (Section 5). The respective main results are as follows.

### 6.1.1 Collisionless simulations

Using a suite of collisionless two-fluid simulations, including up to 3LPT ICs, our conclusions based on an extensive analysis of power spectra are that

- (i) Erroneous growth due to discreteness, previously observed in two-fluid simulations, is absent when initial mass variations instead of displacement perturbations are used (see Fig. 6). These mass variations are small (per cent level), vary mostly on large scales, and are independent of the starting redshift.
- (ii) Residual discreteness and truncation (LPT transient) errors in two-fluid systems are similar to those in single-fluid systems and confined to scales close to the particle Nyquist wavenumber (cf. Michaux et al. 2020).
- (iii) Therefore, late starting times with high-order LPT yield the best accuracy (before shell-crossing), by optimizing the impact of perturbative truncation errors versus discreteness errors.

Furthermore, we confirm a previously reported non-linear suppression in the baryon to CDM power ratio in collisionless simulations (Somogyi & Smith 2010; Angulo et al. 2013) also in the absence of different gravitational softening for baryons and CDM.

### 6.1.2 Hydrodynamic simulations

For the fully hydrodynamic simulations, we use our novel PPT approach up to second order to set up the baryon initial conditions on a regular mesh for the Eulerian finite-volume code RAMSES, while AREPO is initialized with our novel two-fluid LPT, identical to the collisionless simulations. The CDM  $N$ -body particles in RAMSES are of course also initialized using LPT. Using an analysis of CDM, baryon and total matter density power spectra, cumulants of the baryon density field, and CDM and baryon density bispectra, we validate the performance of both the  $n$ LPT and  $n$ PPT ICs between  $z \sim 5$  and 0.5. Since the main purpose of this paper is to present and validate the numerical implementation of our novel PT approaches, our conclusions based on hydrodynamic simulations do not include a rigorous resolution study. Also a study of the impact of the numerous

<sup>5</sup>Available from <https://bitbucket.org/ohahn/monofonic>.

parameters of each code on the results is beyond the scope of this work. Several of the conclusions below should therefore be followed up with more rigorous convergence tests in future work. Our main findings based on the hydrodynamic simulations are as follows:

(i) The PPT approach to set up Eulerian baryon ICs leads to a natural suppression of power on the smallest scales due to the finite effective ‘Jeans’-scale associated with the  $\hbar$ -parameter of this method (which is set mostly by resolution). In contrast, two-fluid LPT is perfectly cold and has no such scale.

(ii) Despite this initial suppression, the late time ( $z \lesssim 2.5$ ) baryon evolution and all power spectra agree well (i.e. within a few per cent) between the Eulerian and the moving mesh runs.

(iii) The improvement brought about by second-order over first-order PT in setting up ICs is virtually identical for the LPT and the PPT initial conditions, as demonstrated by our study of higher order cumulants and bispectra. The impact is similar to the improvements seen for single-fluid total-matter PT (Crocce et al. 2006), and most important for higher order statistics.

(iv) Finite pressure in our non-radiative two fluid simulations leads to a very similar suppression of the total matter spectrum compared to the collisionless simulations at late times, independent of the simulation code, and in broad agreement with previous results based on ‘full-physics’ simulations (Chisari et al. 2019).

(v) We find some interesting differences in the amount of small-scale non-Gaussianity between the RAMSES and AREPO simulations that possibly warrant further investigation.

In conclusion, we presented the numerical implementation of the ‘growing mode’ two-fluid LPT/PPT approach discussed in detail in the companion paper Rampf et al. (2021), and validated its performance for both Eulerian and Lagrangian hydrodynamic codes. We believe that the presented improvements are on par with the necessary increase in the precision of cosmological simulations, in particular when probing baryons at increasingly higher redshifts.

## ACKNOWLEDGEMENTS

We thank our referee Stéphane Colombi for various valuable suggestions that helped to improve the presentation of our results. We thank Michaël Michaux for his bispectrum scripts, Raul Angulo, Bruno Marcos, Fabian Schmidt, and Romain Teyssier for useful discussions, and Volker Springel for comments on an earlier version of the manuscript. OH acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, Grant agreement no. 679145 (COSMO-SIMS). CR is a Marie Skłodowska-Curie Fellow and acknowledges funding from the People Programme (Marie Curie Actions) of the European Union Horizon 2020 Programme under Grant Agreement No. 795707 (COSMO-BLOW-UP). This work was granted access to the HPC resources of TGCC/CINES under the allocation A0060410847 attributed by GENCI (Grand Equipement National de Calcul Intensif). We thank the authors of RAMSES, AREPO, CLASS, and BSKIT for making their software publicly available.

## DATA AVAILABILITY

The software implementing all described techniques to generate initial conditions is freely available at <https://bitbucket.org/ohahn/monofonic>. The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Ahn K., 2016, *ApJ*, 830, 68
- Almgren A. S., Bell J. B., Lijewski M. J., Lukić Z., Van Andel E., 2013, *ApJ*, 765, 39
- Angulo R. E., Pontzen A., 2016, *MNRAS*, 462, L1
- Angulo R. E., Hahn O., Abel T., 2013, *MNRAS*, 434, 1756
- Aricò G., Angulo R. E., Hernández-Monteagudo C., Contreras S., Zennaro M., Pellejero-Ibañez M., Rosas-Guevara Y., 2020, *MNRAS*, 495, 4800
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Phys. Rep.*, 367, 1
- Bernardeau F., van de Rijdt N., Vernizzi F., 2012, *Phys. Rev. D*, 85, 063509
- Binney J., Tremaine S., 2008, *Galactic Dynamics*, 2nd edn. Princeton Series in Astrophysics, Princeton Univ. Press, Princeton
- Bird S., Feng Y., Pedersen C., Font-Ribera A., 2020, *J. Cosmology Astropart. Phys.*, 2020, 002
- Blas D., Lesgourgues J., Tram T., 2011, *J. Cosmol. Astropart. Phys.*, 2011, 034
- Blazek J. A., McEwen J. E., Hirata C. M., 2016, *Phys. Rev. Lett.*, 116, 121303
- Brandbyge J., Rampf C., Tram T., Leclercq F., Fidler C., Hannestad S., 2017, *MNRAS*, 466, L68
- Bryan G. L. et al., 2014, *ApJS*, 211, 19
- Buehlmann M., Hahn O., 2019, *MNRAS*, 487, 228
- Chen S.-F., Castorina E., White M., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 006
- Chisari N. E. et al., 2019, *Open J. Astrophys.*, 2, 4
- Colombi S., Jaffe A., Novikov D., Pichon C., 2009, *MNRAS*, 393, 511
- Crocce M., Pueblas S., Scoccimarro R., 2006, *MNRAS*, 373, 369
- Dalal N., Pen U.-L., Seljak U., 2010, *J. Cosmol. Astropart. Phys.*, 2010, 007
- Dubois Y. et al., 2014, *MNRAS*, 444, 1453
- Durrer R., 2008, *The Cosmic Microwave Background*. Cambridge Univ. Press, Cambridge
- Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, *ApJS*, 57, 241
- Emberson J. D., Frontiere N., Habib S., Heitmann K., Larsen P., Finkel H., Pope A., 2019, *ApJ*, 877, 85
- Fialkov A., Barkana R., Tseliakhovich D., Hirata C. M., 2012, *MNRAS*, 424, 1335
- Fidler C., Tram T., Rampf C., Crittenden R., Koyama K., Wands D., 2017, *J. Cosmol. Astropart. Phys.*, 2017, 043
- Foreman S., Coulton W., Villaescusa-Navarro F., Barreira A., 2020, *MNRAS*, 498, 2887
- Greif T. H., White S. D. M., Klessen R. S., Springel V., 2011, *ApJ*, 736, 147
- Guth A. H., Hertzberg M. P., Prescod-Weinstein C., 2015, *Phys. Rev. D*, 92, 103513
- Hahn O., Abel T., 2011, *MNRAS*, 415, 2101
- Hockney R. W., Eastwood J. W., 1981, *Computer Simulation Using Particles*, Computer Simulation Using Particles. McGraw-Hill, New York
- Hopkins P. F., 2015, *MNRAS*, 450, 53
- Hu W., Dodelson S., 2002, *ARA&A*, 40, 171
- Huang H.-J., Eifler T., Mandelbaum R., Dodelson S., 2019, *MNRAS*, 488, 1652
- Joyce M., Marcos B., 2007, *Phys. Rev. D*, 75, 063516
- Joyce M., Marcos B., Gabrielli A., Baertschiger T., Sylos Labini F., 2005, *Phys. Rev. Lett.*, 95, 011304
- Klypin A. A., Shandarin S. F., 1983, *MNRAS*, 204, 891
- Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, *ApJS*, 111, 73
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Ludlow A. D., Schaye J., Schaller M., Bower R., 2020, *MNRAS*, 493, 2926
- Marcos B., Baertschiger T., Joyce M., Gabrielli A., Sylos Labini F., 2006, *Phys. Rev.*, D73, 103507
- Matsubara T., 2015, *Phys. Rev. D*, 92, 023534
- McCullagh N., Jeong D., Szalay A. S., 2015, *MNRAS*, 455, 2945
- Michaux M., Hahn O., Rampf C., Angulo R. E., 2020, *MNRAS*, 500, 663
- Munshi D., Sahní V., Starobinsky A. A., 1994, *ApJ*, 436, 517
- Naoz S., Barkana R., 2005, *MNRAS*, 362, 1047
- Naoz S., Yoshida N., Barkana R., 2011, *MNRAS*, 416, 232



O’Leary R. M., McQuinn M., 2012, *ApJ*, 760, 4  
 Planck Collaboration VI, 2020, *A&A*, 641, A6  
 Porqueres N., Hahn O., Jasche J., Lavaux G., 2020, *A&A*, 642, A139  
 Rampf C., 2012, *J. Cosmol. Astropart. Phys.*, 1212, 004  
 Rampf C., Hahn O., 2021, *MNRAS letters*, 501, L71  
 Rampf C., Villone B., Frisch U., 2015, *MNRAS*, 452, 1421  
 Rampf C., Uhlemann C., Hahn O., 2021, *MNRAS*, in press  
 Schaller M., Gonnet P., Chalk A. B. G., Draper P. W., 2016, preprint (arXiv:1606.02738)  
 Schaye J. et al., 2015, *MNRAS*, 446, 521  
 Schmidt F., 2016, *Phys. Rev. D*, 94, 063508  
 Schneider A., Teyssier R., 2015, *J. Cosmol. Astropart. Phys.*, 2015, 049  
 Schneider A., Teyssier R., Stadel J., Chisari N. E., Le Brun A. M. C., Amara A., Refregier A., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 020  
 Scoccimarro R., 1998, *MNRAS*, 299, 1097  
 Scoccimarro R., 2000, *ApJ*, 544, 597  
 Sefusatti E., Crocce M., Scoccimarro R., Couchman H. M. P., 2016, *MNRAS*, 460, 3624  
 Shandarin S. F., Zeldovich Ya. B., 1989, *Rev. Mod. Phys.*, 61, 185  
 Slepian Z., Eisenstein D. J., 2015, *MNRAS*, 448, 9  
 Slepian Z. et al., 2018, *MNRAS*, 474, 2109  
 Somogyi G., Smith R. E., 2010, *Phys. Rev. D*, 81, 023524  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Springel V., 2010, *MNRAS*, 401, 791  
 Springel V. et al., 2018, *MNRAS*, 475, 676  
 Taruya A., Nishimichi T., Jeong D., 2018, *Phys. Rev. D*, 98, 103532  
 Teyssier R., 2002, *A&A*, 385, 337  
 Tomlinson J., Jeong D., Kim J., 2019, *AJ*, 158, 116  
 Tsaliakhovich D., Hirata C., 2010, *Phys. Rev. D*, 82, 083520  
 Uhlemann C., Rampf C., Gosenca M., Hahn O., 2019, *Phys. Rev. D*, 99, 083524  
 Valkenburg W., Villaescusa-Navarro F., 2017, *MNRAS*, 467, 4401  
 van Leer B., 1979, *J. Comput. Phys.*, 32, 101  
 Wadsley J. W., Keller B. W., Quinn T. R., 2017, *MNRAS*, 471, 2357  
 Weinberger R., Springel V., Pakmor R., 2020, *ApJS*, 248, 32  
 Weltman A. et al., 2020, *PASA*, 37, e002  
 White S. D. M., 1996, in Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds, *Cosmology and Large Scale Structure*. Elsevier, North Holland, p. 349  
 Yoo J., Dalal N., Seljak U., 2011, *J. Cosmol. Astropart. Phys.*, 2011, 018  
 Yoshida N., Sugiyama N., Hernquist L., 2003, *MNRAS*, 344, 481  
 Zel’dovich Y. B., 1970, *A&A*, 500, 13  
 Zennaro M., Bel J., Villaescusa-Navarro F., Carbone C., Sefusatti E., Guzzo L., 2017, *MNRAS*, 466, 3244  
 Zheligovsky V., Frisch U., 2014, *J. Fluid Mech.*, 749, 404

## APPENDIX: CODE ACCURACY PARAMETERS

We carried out all simulations with the parameter settings listed in Table A1 for RAMSES, and in Table A2 for AREPO (and GADGET-2 where applicable).

**Table A1.** RAMSES code parameter values used in this paper.

RUN_PARAMS	
nsubcycle	1, 2
AMR_PARAMS	
levelmin	9
levelmax	17
HYDRO_PARAMS	
courant_factor	0.8
slope_type	2
pressure_fix	.true.
scheme	‘muscl’
riemann	hllc
REFINE_PARAMS	
m.refine	4., 10*8.
interpol_var	1
interpol_type	0

**Table A2.** AREPO code parameter values used in this paper. The time-step and force accuracy/softening parameters also apply to the collisionless GADGET-2 runs, where baryons are however treated as Type2 particles instead of Type0. Also for the  $2 \times 256^3$  GADGET-2 run, the softening used was twice larger.

TypeOfTimestepCriterion	0
ErrTolIntAccuracy	0.025
CourantFac	0.8
MaxSizeTimestep	0.01
TypeOfOpeningCriterion	1
ErrTolTheta	0.7
ErrTolForceAcc	0.0025
SofteningComovingType0	0.025
SofteningComovingType1	0.025
SofteningMaxPhysType0	0.25
SofteningMaxPhysType1	0.25
GasSoftFactor	2.5
SofteningTypeOfPartType0	0
SofteningTypeOfPartType1	1
MinimumComovingHydroSoftening	0.025
CellShapingSpeed	0.5
CellMaxAngleFactor	2.25
ReferenceGasPartMass	0
TargetGasMassFactor	1
RefinementCriterion	1
DerefinementCriterion	1

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