

Biased tracers of two fluids in the Lagrangian picture

Shi-Fan Chen,^a Emanuele Castorina^{a,b} and Martin White^{a,b,c}

^aDepartment of Physics, University of California,
Berkeley, CA 94720, U.S.A.

^bBerkeley Center for Cosmological Physics,
Berkeley, CA 94720, U.S.A.

^cDepartment of Astronomy, University of California,
Berkeley, CA 94720, U.S.A.

E-mail: shifan_chen@berkeley.edu, ecastorina@berkeley.edu, mwhite@berkeley.edu

Received March 4, 2019

Accepted May 22, 2019

Published June 4, 2019

Abstract. We explore Lagrangian perturbation theory (LPT) for biased tracers in the presence of two fluids, focusing on the case of cold dark matter (CDM) and baryons. The presence of two fluids induces corrections to the Lagrangian bias expansion and tracer advection, both of which we formulate as expansions in the three linear modes of the Lagrangian equations of motion. We compute the linear-order two-fluid corrections in the Zeldovich approximation, finding that modifications to the bias expansion and tracer advection both enter as percent-level corrections over a large range of wavenumbers at low redshift and draw parallels with the Eulerian formalism. We then discuss nonlinear corrections in the two-fluid picture, and calculate contributions from the relative velocity effect ($\propto \mathbf{v}_r^2$) at one loop order. Finally, we conduct an exploratory Fisher analysis to assess the impact of two-fluid corrections on baryon acoustic oscillations (BAO) measurements, finding that while modest values of the relative bias parameters can introduce systematic biases in the measured BAO scale of up to 0.5σ , fitting for these effects as additional parameters increases the error bar by less than 30% across a wide range of bias values.

Keywords: baryon acoustic oscillations, cosmological parameters from LSS, cosmological perturbation theory

ArXiv ePrint: [1903.00437](https://arxiv.org/abs/1903.00437)

Contents

1	Introduction	1
2	Linear equations of motion in Lagrangian space	2
2.1	General formalism	3
2.2	Initial conditions and transfer functions	4
3	Lagrangian bias in the two-fluid dynamics	5
3.1	Bias expansion	6
3.2	Modifications to tracer advection	9
4	Galaxy power spectra in the Zeldovich approximation	10
4.1	Analytic form	10
4.2	Cross-spectra of different tracers and IR resummation	13
4.3	Higher order bias	14
5	Degeneracies and bias to BAO	16
6	Conclusions	21
A	Redshift dependence and size of the of bias parameters	23
B	Beyond linear order	23
B.1	Equations of motions	23
B.2	Biasing at second order	25
C	Relative velocity bias terms	25

1 Introduction

Observations of the large-scale structure (LSS) of the universe allow us to shed light on areas of physics ranging from galaxy formation and evolution to fundamental physics. A prime target of present and future LSS surveys is the measurement of baryon-acoustic oscillations (BAO) — the imprints of sound waves in the baryon-photon fluid observed in the cosmic microwave background (CMB) on the observed clustering of galaxies — which can be used as a standard ruler to constrain the expansion of the universe [1]. Upcoming surveys such as DESI [2], EUCLID [3] and WFIRST [4] will provide BAO measurements with higher-than-ever precision, and even more futuristic BAO surveys such as a Stage II 21-cm experiment [5] have been proposed. These next-generation observational campaigns will require us to model the LSS with unprecedented accuracy, at the sub-% level.

One area of recent interest in the field of LSS has been in accounting for the effects induced by the existence of multiple species (cold dark matter, baryons, neutrinos), with similar but distinct clustering properties, using analytic methods. Studies of the perturbative approach to structure formation have traditionally grouped all nonrelativistic species into a “total matter” fluid, whose gravitational collapse is the dominant source of structure on cosmological scales in the late-time universe, but many authors have recently extended

these techniques to include neutrinos [6–10] and baryons [11–17] in the Eulerian framework of Standard Perturbation Theory (SPT). In parallel, the response of galaxy and halo formation to the existence of multiple fluid species has also been subject of extensive investigation [11, 16–22]. Of particular interest are the present-day imprint of relative perturbations between baryons and dark matter on large scales which, being seeded in the same epoch and at the same scales as the baryon acoustic oscillations, has the potential to confound future BAO measurements [11, 22–24]. While these relative perturbations do not grow significantly in time (and relative velocities in fact decay) and are thus small compared to the total-matter growing mode at late times, they amount to coherent supersonic flows post-recombination and could have significant effects on the formation of the first halos and galaxies [11, 13], which are the progenitors of the objects we observe today.

The goal of this work is to formulate perturbation theory and galaxy bias in the presence of multiple fluids within the Lagrangian framework, with a particular focus on the two-fluid baryon-dark matter scenario. Our work is a direct extension of the aforementioned SPT calculations. While Lagrangian Perturbation Theory (LPT) is order-by-order equivalent to SPT, it seamlessly allows a consistent treatment of large scales bulk flows, which are responsible for the final shape and position of the BAO features in the correlation functions or power spectrum [25–31]. The theory can also be extended to handle density field ‘reconstruction’ [32–36]. These features make LPT a natural language for investigating possible distortions to the BAO feature.

This paper is organized as follows. In section 2, we introduce the linear Lagrangian equations of motion and discuss the role of non-gravitational forces such as Compton drag with the CMB. Modifications to Lagrangian galaxy bias and advection in the two-fluid limit are then introduced in section 3. In section 4, we employ the results of the preceding two sections and calculate the lowest-order two-fluid corrections to the galaxy power spectrum in the Zeldovich approximation. Cross spectra and subtleties in the IR resummation are briefly discussed in section 4.2. In section 5 we take up whether the calculated two-fluid corrections can significantly bias BAO measurements, arguing that any such biases can be mitigated by simultaneously fitting for these easily-characterizeable effects. Our conclusions are summarized in section 6.

2 Linear equations of motion in Lagrangian space

In the Lagrangian picture, fluid dynamics is encoded in the displacements $\Psi_\sigma(q)$ of fluid elements of each species, σ , originally situated at Lagrangian positions \mathbf{q} , such that their Eulerian positions at conformal time τ ($d\tau = a^{-1}dt$) are given by [28, 37, 38]

$$\mathbf{x}_\sigma(\mathbf{q}, \tau) = \mathbf{q} + \Psi_\sigma(\mathbf{q}, \tau). \quad (2.1)$$

The subscript $\sigma = \{c, b\}$ denotes the species, either cold dark matter (CDM) or baryons, respectively, whose motion we are tracking. Assuming that initial displacements are infinitesimally small compared to those at the redshifts of interest, the overdensity, δ_σ , of each species at Eulerian position x can be solved for via mass conservation

$$1 + \delta_\sigma(\mathbf{x}, \tau) = \int d^3q \delta_D(\mathbf{x} - \mathbf{q} - \Psi_\sigma(\mathbf{q}, \tau)) = \int d^3q \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{q} - \Psi_\sigma(\mathbf{q}, \tau))}, \quad (2.2)$$

where δ_D is the Dirac delta function. Taylor expanding to first order in displacements yields the familiar result that $\delta_\sigma(x) = -\nabla \cdot \Psi_\sigma(q)$, but, as seen in equation (2.2), one feature of

working in the Lagrangian picture is that the translation into Eulerian quantities, such as the density field, invariably involves nonlinear combinations of Ψ even when only the linear equations of motion are considered.

2.1 General formalism

While CDM particles by assumption experience only the gravitational force, baryons are subject to non-gravitational effects, such as Compton drag and pressure gradients. These effects can be summarized in the equations of motion of the fluid elements

$$\begin{aligned}\ddot{\Psi}_c + \mathcal{H}\dot{\Psi}_c &= -\nabla_x \Phi(\mathbf{q} + \Psi_c) \\ \ddot{\Psi}_b + \mathcal{H}\dot{\Psi}_b &= -\nabla_x \Phi(\mathbf{q} + \Psi_b) + \mathbf{F}_b(\mathbf{q} + \Psi_b),\end{aligned}\tag{2.3}$$

where overdots signify derivatives with respect to τ , $\mathcal{H} = d \ln a / d\tau$ is the conformal Hubble parameter, \mathbf{F}_b is the non-gravitational force per unit mass felt by baryons, and Φ is the gravitational potential at Eulerian position x satisfying Poisson's equation

$$\nabla_x^2 \Phi(\mathbf{x}, \tau) = \frac{3}{2} \Omega_m(\tau) \mathcal{H}^2(\tau) \delta_m(\mathbf{x}, \tau),\tag{2.4}$$

where Ω_m is the total matter mass density and δ_m is the total matter overdensity (see below).

At the linear level, there is no difference between the Eulerian and Lagrangian positions in the above equations of motion, and we will neglect this distinction in the rest of this section unless otherwise stated. Indeed, taking the divergence of equation (2.3) in the linear limit ($x_\sigma \approx q$) directly yields the Euler equation when we map overdensities to displacements and velocities to their derivatives:

$$\delta_\sigma(\mathbf{x}_\sigma) \leftrightarrow -\nabla \cdot \Psi_\sigma(\mathbf{q}), \quad \mathbf{v}_\sigma(x_\sigma) \leftrightarrow \dot{\Psi}_\sigma(\mathbf{q}).\tag{2.5}$$

Note that the first mapping is correct only to linear order, while the second one is exact if the full $\mathbf{x}(\mathbf{q})$ is used. Assuming this translation, the solutions to the Lagrangian equations of motion as described below are essentially identical to those extracted from Boltzmann codes such as CAMB [39] or CLASS [40], provided one chooses post-recombination initial conditions for the Lagrangian displacements.

To solve equation (2.3) in the linear limit, it is convenient to rewrite the baryonic and CDM displacements in terms of a mass-weighted matter component ($\Psi_m = w_c \Psi_c + w_b \Psi_b$), which sources the gravitational potential, and a relative component that characterizes the differential flows between baryons and CDM ($\Psi_r = \Psi_b - \Psi_c$), where we have defined the mass fractions of each species, $w_\sigma = \rho_\sigma / \rho_m$. These are related to the Eulerian quantities $\delta_m = w_b \delta_b + w_c \delta_c$ and $v_r = v_b - v_c$ by $\delta_a = -\nabla \cdot \Psi_a$ and $v_a = \dot{\Psi}_a$, where $a = \{m, r\}$, again at the linear level. The equations of motion in terms of these components are

$$\ddot{\Psi}_m + \mathcal{H}\dot{\Psi}_m = -\nabla \Phi + w_b \mathbf{F}_b\tag{2.6a}$$

$$\ddot{\Psi}_r + \mathcal{H}\dot{\Psi}_r = \mathbf{F}_b.\tag{2.6b}$$

If in addition non-gravitational forces are negligible, the matter and relative components decouple, such that equation (2.6a) can be solved as

$$\Psi_m(\tau) = -\mathbf{m}_+ D_+(\tau) + \mathbf{m}_- D_-(\tau) \approx -\mathbf{m}_+ D_+(\tau),\tag{2.7}$$

where D_+ is the usual linear-theory growth factor. In the last step we have neglected the decaying mode, \mathbf{m}_- , since it is a tiny fraction of the total displacement at all redshifts of interest. For non-gravitational forces, like Compton drag or pressure gradients, direct integration of the linear equations of motion show that the non-gravitational terms make a negligible contribution to the matter component Ψ_m , such that the transfer function at redshifts below $z = 6$ agree with the linear solution in equation (2.7) to within 0.2%, with even better agreement at the lower redshifts of interest in this paper. In the above we have included a minus sign for convenience such that $\delta_{m,0} = \nabla \cdot \mathbf{m}_+$.

We end this subsection by discussing the full solution of the relative displacement when $\mathbf{F}_b = \mathbf{F}_b(\tau)$ is independent of Ψ_r . In this case equation (2.6b) is linear and first order in $\dot{\Psi}_r$ and can be solved as:

$$\dot{\Psi}_r(\tau) = \mathbf{v}_r(\tau_i) \left(\frac{a_i}{a} \right) + \frac{1}{a} \int_{\tau_i}^{\tau} d\tau' a(\tau') \mathbf{F}_b(\tau'), \quad (2.8)$$

where we have set the boundary conditions at initial time τ_i assuming the non-gravitational effects encoded in F_b do not turn on until $\tau > \tau_i$. Equation (2.8) turns out to be an excellent approximation for the large-scale Compton drag electrons experience in the reionization era due to their relative motion with respect to the CMB rest frame, $\mathbf{F}_b = -n_e \sigma_T (\rho_\gamma / \rho_b) a \mathbf{v}_b$, where σ_T is the Thompson scattering cross section, ρ_γ is the photon energy density and n_e the free electron number density. Eq. (2.8) also applies baryonic pressure forces on small scales $\mathbf{F}_b \propto -\nabla \delta_b$ — in both cases the total-matter component may be substituted for the baryonic component (i.e. $\delta_b \approx \delta_m$) at the sub-percent level [17]. In the case of the large-scale Compton drag, assuming $\mathbf{v}_b \simeq \mathbf{v}_m$ yields

$$\dot{\Psi}_r(\tau) = \dot{\Psi}_r(\tau_i) \frac{a(\tau_i)}{a(\tau)} + \left[\frac{1}{a} \int_{\ln(a(\tau_i))}^{\ln(a(\tau))} d \ln(a') \left(n_e(a') \sigma_T \frac{\rho_\gamma(a')}{\rho_b(a')} \right) \frac{f(a') D_+(a')}{a'^2} \right] \frac{\Psi_m(\tau_i)}{D_+(\tau_i)}, \quad (2.9)$$

with $f = dD_+/d \ln(a)$ the linear theory growth factor. The Compton drag thus induces a mixing between the matter and relative components through a numerical prefactor dependent only on the linear growth factor D_+ and reionization history via n_e . Finally, we can integrate (2.9) to yield

$$\Psi_r(\tau) = -\mathbf{r}_+ + \mathbf{r}_- D_r(\tau, \tau_i) + \mathbf{m}_+ D_{\text{CD}}(\tau, \tau_i), \quad D_r(\tau, \tau_i) = \int_{\tau_i}^{\tau} \frac{H_0 d\tau'}{a(\tau')} \quad (2.10)$$

where we can identify $\Psi_r(\tau_i) = -\mathbf{r}_+$, $a(\tau_i) \mathbf{v}_r = H_0 \mathbf{r}_-$, and the Compton-drag kernel D_{CD} is defined as the conformal time integral of the square-bracketed function in (2.9). The linear solutions to both the total-matter and relative components are thus wholly specified by the three modes \mathbf{m}_+ and \mathbf{r}_\pm . Jeans instabilities and baryonic pressure forces affect much smaller scales and won't be further discussed in the remainder of this work.

2.2 Initial conditions and transfer functions

The linear evolution of the density and velocity contrasts can be easily written in terms the CDM and baryon linear transfer functions (output from, e.g. CAMB) as

$$T_{\delta_r}(k) \equiv T_{\delta_b}(k) - T_{\delta_c}(k) \quad \text{and} \quad T_{\theta_r}(k) \equiv T_{\theta_b}(k) - T_{\theta_c}(k) \quad (2.11)$$

where $\theta_{b,c}(k) \equiv -\dot{\delta}_{b,c}(k)$. It is worth noticing that while the velocity field is gauge dependent, velocity differences are not. The transfer function for $\nabla \cdot \mathbf{m}_+$ is simply the present-day matter transfer function T_m and we can furthermore define

$$\begin{aligned} T_{\nabla \cdot \mathbf{r}_+}(k) &\equiv T_b(k, z_i) - T_c(k, z_i) \\ T_{\nabla \cdot \mathbf{r}_-}(k) &\equiv [(1 + z_i)H_0]^{-1} \left(T_{\theta_b}(k, z_i) - T_{\theta_c}(k, z_i) \right). \end{aligned} \quad (2.12)$$

These three functions specify the solution for the Ψ_m , Ψ_r and $\dot{\Psi}_r$ at any $z < z_i$. The choice of z_i is somewhat arbitrary but choosing redshifts before the onset of reionization has the advantage of separating the effects of gravity from Compton drag. This choice also justifies the normalization in eq. (2.12), since \mathbf{r}_- is independent of redshift. In the remainder of the paper we assume $z_i = 20$.

In addition to the above, we will show below that calculating the power spectrum at some redshift z in the Lagrangian picture requires linear-theory spectra of the relative displacement at that redshift, which will typically include corrections from Compton drag. These can be calculated via equations (2.7) and (2.8) to give

$$T_{\nabla \cdot \Psi_r}(k, z) = T_{\nabla \cdot \mathbf{r}_+}(k) + D_r(z, z_i)T_{\nabla \cdot \mathbf{r}_-}(k) + D_{\text{CD}}(z, z_i)T_{\nabla \cdot m_+}(k). \quad (2.13)$$

Sample solutions of the equation of motion in eq. (2.6) when F_b is given by Compton drag with the CMB are shown in figure 1. After reionization most of large scale power in the relative velocity transfer function, $T_{\nabla \cdot \dot{\Psi}_r}$, is provided by the Compton drag, which in turn affects the evolution of the relative baryon-dark matter density at large scales (see top panels in figure 1). Figure 1 also justifies the approximations we used to compute the drag forces, as one can see by the excellent agreement with the full CAMB output. Other non-gravitational effects like pressure terms (Jeans instability) and radiative transfer effects [41–44], can be written in a similar form.

Ratios of the transfer functions to the total matter one are shown in figure 2. We notice that the relative density perturbation is much larger than the relative velocity one, by a factor of a hundred at least, and the two relative components have the same behavior with wave-number k at small and large scales. Nonetheless \mathbf{r}_+ and \mathbf{r}_- have significant differences in shape around the BAO scales and therefore will have to be treated separately from the point of view of the galaxy bias expansion.

3 Lagrangian bias in the two-fluid dynamics

In the Lagrangian approach, galaxy bias is assumed to arise as the response of the overdensity of galaxies, or the precursors thereof, to the variation of the initial conditions encoded in the fields $\{\Psi_\sigma(\mathbf{q})\}$ of the various species, and then transported via advection to their present-day positions $\mathbf{x}(q, t) = \mathbf{q} + \Psi_g(\mathbf{q}, t)$. Thus, when computing the density of a biased tracer the number-conservation equation (2.2) is modified to

$$1 + \delta_g(\mathbf{x}, \tau) = \int d^3q F_g[\mathbf{q} | \{\Psi_\sigma(\mathbf{q})\}] \delta_D[\mathbf{x} - \mathbf{q} - \Psi_g(\mathbf{q}, \tau)]. \quad (3.1)$$

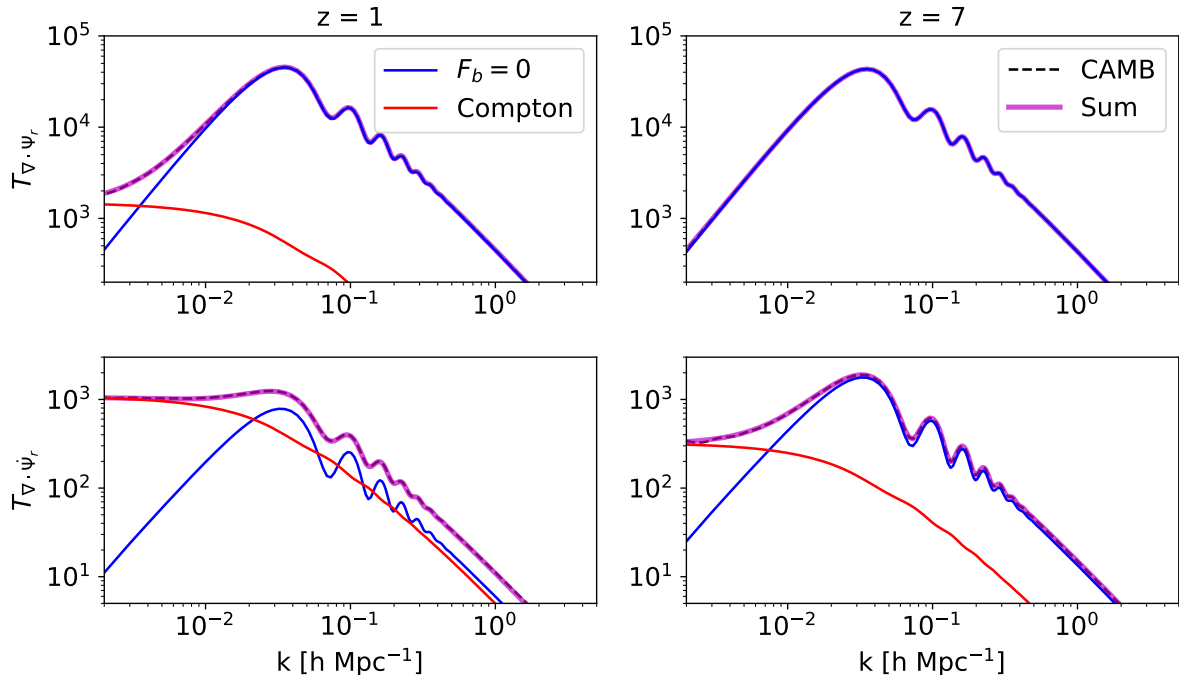


Figure 1. Transfer functions for the relative component from equation (2.13) at $z = 1$ (left column) and $z = 7$ (right column). These transfer functions solve equation (2.8). The top row shows the transfer functions for $\nabla \cdot \Psi_r$, i.e. the relative density. The bottom row shows the transfer functions for $\nabla \cdot \dot{\Psi}_r$, i.e. the relative velocity divergence. The free-falling ($F_b = 0$) and Compton drag contributions are shown separately, the effect of Compton drag on the relative velocity is immediately apparent even right after reionization ($z_{\text{re}} = 7.90$) at $z = 7$, whereas the relative displacement is dominated by the $F_b = 0$ contribution at all but the largest scales shown. Unlike the Compton contribution, which is flat at large scales, the primordial ($F_b = 0$) contributions fall off as k^2 towards low wavenumbers, reflecting the origin of relative perturbations in pre-recombination baryonic pressure forces. At low redshifts, the solutions to the Lagrangian equations of motion, with initial conditions set at $z_i = 20$, are in excellent quantitative agreement with the results from CAMB (black dashed lines, barely visible on the plot as they lie below the purple lines).

The standard picture of (local) Lagrangian bias, outlined above, has been developed in the 1-fluid case by many authors, see for example [45–53] and [54] for a recent review on galaxy bias. In this section our focus will be on extending these arguments to the case of multiple fluids, and in particular to the two-fluid case. In the presence of two fluids, the form of equation (3.1) raises two questions: (1) the form of the response F_g and (2) whether biased tracers follow the dark matter, baryons, or a combination thereof. We address these in turn.

3.1 Bias expansion

The initial tracer overdensity, $F_g[\mathbf{q}|\{\Psi_\sigma(\mathbf{q})\}]$, is defined to be a functional encoding the physics of gravitational collapse and galaxy formation at some Lagrangian position \mathbf{q} . Since the galaxy density field is a scalar quantity under rotations, F_g will also be a scalar. We will assume this functional is local, in the sense gravitational collapse depends only on the value of the fields within a characteristic patch of size R_h , which then flows coherently on large scales with Ψ_g [54]. In the fluid limit, these conditions imply that the system is wholly specified — albeit in some complex, nonlinear way — by the species overdensities, $\delta_\sigma(\mathbf{q})$, velocities,

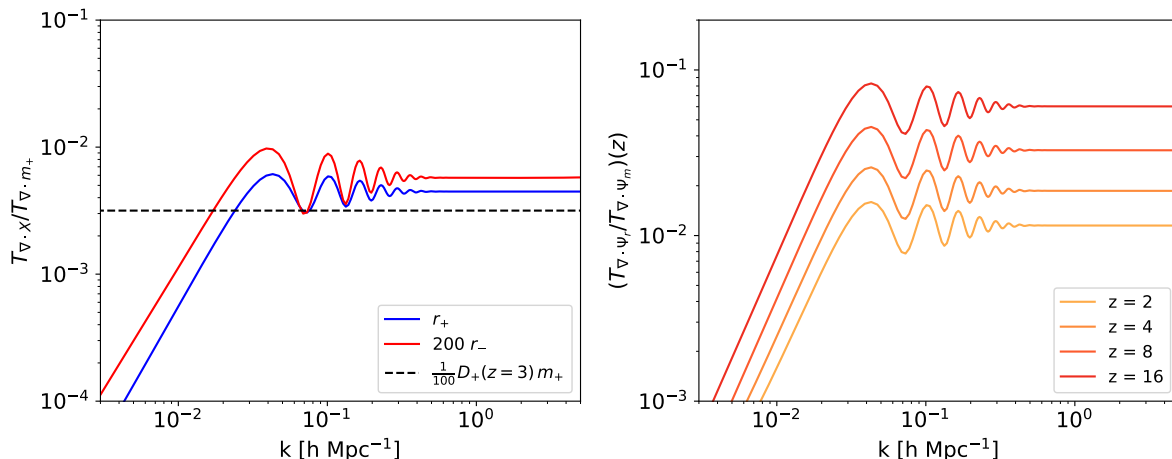


Figure 2. Relative to total-matter-component transfer function ratios. (Left) Transfer function ratios between the initial fields m_+ and r_{\pm} defined at $z = 20$. The so-normalized constant r_+ , which roughly corresponds to the relative overdensity mode, is a percent level contribution relative to the total-matter growing mode m_+ . The decaying mode r_- , which corresponds roughly to the relative velocity, enters at significantly below the percent level. Note however that our definition somewhat exaggerates its smallness by “redshifting” it to $z = 0$. The equivalent ratio for one percent of the growing mode at $z = 3$ is plotted for comparison in black. (Right) Transfer function ratios between the evolved relative and total matter displacements at redshifts $z = 2$ – 6 . While the relative displacement is a percent level effect at low redshifts ($z = 2$), it enters at close to the ten percent level at higher redshifts ($z = 16$).

$\mathbf{v}_\sigma(\mathbf{q})$, and the gravitational potential,¹ $\Phi(\mathbf{q})$, at some initial time τ_i . The condition that F_g is local — or rather, nonlocal with width R_h — can be equivalently (but more conveniently) expressed by requiring F_g depend only on the initial fields and their spatial derivatives, with n^{th} derivatives suppressed by n powers of R_h [54].

In addition to the assumption of locality, the form of F_g is strongly restricted by various symmetries. General relativity requires that all physical quantities be diffeomorphism invariant, which in our case reduces to rotational invariance and invariance under generalized Galilean transformations [55]:

$$\mathbf{q} \rightarrow \mathbf{q}, \quad \Psi_\sigma \rightarrow \Psi_\sigma + \mathbf{n}(\tau), \quad \Phi \rightarrow \Phi - \mathbf{x} \cdot (\ddot{\mathbf{n}} + \mathcal{H}\dot{\mathbf{n}}), \quad (3.2)$$

where \mathbf{n} are time-dependent but spatially constant vector fields.

Rotational invariance simply requires that only contracted scalar quantities enter the bias; the restrictions placed on the bias expansion by general Galilean invariance are more subtle, and it is here that the two-fluid case diverges from the single-fluid case. Under this symmetry, densities remain unchanged — for instance that at first order $\delta_\sigma(\mathbf{q}) = -\nabla \cdot \Psi_\sigma(\mathbf{q})$ — while velocities get boosted by a spatially constant amount (leaving $\partial \mathbf{v}$ invariant) and the gravitational potential changes in a spatially linear way (leaving $\partial \Phi$ invariant). In the single-fluid regime, where only one set of densities and velocities exist, this directly implies that velocities can only enter with at least one spatial derivative, and the gravitational potential

¹The gravitational potential Φ , while not independent of δ_m , depends on the total matter density in a very non-local way. To make our bias expansion local, and be able to truncate the derivative expansion at a reasonable order, we thus include it as a standalone quantity here.

can only enter as second (spatial) derivatives and beyond. The single-fluid overdensity, which is unchanged under the transformation, can enter at any order.

The presence of two or multiple fluids relaxes some of the above restrictions. In particular, since all particle velocities are boosted by the same amount (\mathbf{n}') under a Galilean transformation, the relative velocity $\mathbf{v}_r = \mathbf{v}_b - \mathbf{v}_c$ remains invariant and can thus enter the bias expansion at zeroth order in derivatives. The total matter velocity, \mathbf{v}_m , on the other hand, is boosted and can thus still only enter at the derivative level. These two quantities form an equivalent basis to the individual species velocities and there is no loss of generality in defining the bias expansion in terms of them. We may similarly write terms involving species densities, which can enter separately, in the total matter and relative density basis. In general relativity the gravitational potential is unaffected by the number of species as a consequence of the equivalence principle, i.e. gravitational interactions are universal. The full set of physical fields that can enter F_g in the two fluid case is then

$$F_g = F_g[\delta_\sigma, \mathbf{v}_\sigma, \Phi] \equiv F_g[\delta_m, \delta_r, \partial\mathbf{v}_m, \mathbf{v}_r, \partial\partial\Phi, \dots], \quad (3.3)$$

where the dots stand for higher derivative operators. To first order in the fields we can therefore write²

$$1 + \delta_g(\mathbf{q}) = 1 + b_1\delta_m + b_r\delta_r + b_\theta\theta_r + \dots \quad (3.4)$$

which is similar to the Eulerian linear theory expression in [16]. This is not surprising, since at first order $\mathbf{q} \simeq \mathbf{x}$, however we will see below that differential advection can introduce further terms degenerate with the initial Lagrangian bias terms above, such that the Eulerian relative-component bias will in general be a combination of these terms.

Finally, since F_g is defined as a functional on the initial conditions which can be chosen to be sufficiently early that they lie deep in the linear-theory regime, its form can be further simplified and expressed purely in terms of the initial modes \mathbf{m}_+ and \mathbf{r}_\pm . In the single fluid case, this restriction leads to the simplification that all bias terms can be written in terms of spatial derivatives of the total matter displacement $\mathbf{m}_+ \sim \Psi_m$; this is a direct consequence that, up to time-dependent constant factors, $\delta_m \sim \partial\Psi_m$, $\mathbf{v}_m \sim \Psi_m$ and $\partial\partial\Phi \sim \partial\Psi$ in linear theory. In the two-fluid case these terms must be supplemented by those involving the relative modes. Specifically, including the \mathbf{v}_r dependence requires the inclusion of terms proportional to \mathbf{r}_- and including δ_r dependence similarly requires terms proportional to $\nabla \cdot \mathbf{r}_+$. Equation (3.4) can thus be re-expressed as:

$$F_g(\mathbf{q}) = b_1\delta_m + b_+\nabla \cdot \mathbf{r}_+ + b_-\nabla \cdot \mathbf{r}_- + \dots \quad (3.5)$$

We therefore have a direct correspondence in the bias expansion between the initial modes expressed in Eulerian and Lagrangian space. Notice that the bias expansion defined above is complete, in the sense that it contains all possible operators compatible with the symmetries of the problem. In particular, while \mathbf{r}_\pm are defined at a particular initial redshift z_i , in the linear regime this dependence amounts to a simple linear transformation and can be absorbed into the definition of the bias parameters (appendix A).

Finally, an additional complication arises when halo formation is affected by Compton drag. As pointed out by [17], by picking out the local CMB rest frame such that the drag force $\propto \mathbf{v}_b$, we lose the gauge redundancy of Galilean transformations. This will in general produce heretofore forbidden terms such as those proportional to the matter-component velocity \mathbf{v}_m .

²A list of bias terms up to second order is given in appendix B.2.

However, the terms thus generated are required by rotational invariance to enter at second order and beyond. For the remainder of this paper we will thus neglect these contributions, which are subdominant to the already sub-percent level contributions we study.

Whereas there exists quite a large literature on measuring and predicting, using approximate physical models, the value of the bias parameters in one-fluid scenarios, less attention has been devoted to the multi-fluid case. From an effective field theory perspective the dimensionless parameters should be of order unity, but in reality the actual value of the bias parameters is tracer-dependent and can be quite a bit larger or smaller. In this work we will assume, unless otherwise noted, that typical values are given by $b_+ \simeq 1$ and $b_- \simeq 6.8$ derived in [16] using a spherical collapse model. These numbers are consistent with the non-detection of relative bias effects in BOSS DR12 by ref. [23], who find e.g. $b_+ = -1.0 \pm 2.5$ to within one sigma when fitting for b_+ , b_- and c_- (section 4.3) across all redshift bins, with large systematic biases measured in dark-matter only simulations that had to be subtracted.

3.2 Modifications to tracer advection

Once the initial, biased tracer overdensity is set, the overdensity at later times is set by the tracer “fluid” advecting from initial (Lagrangian) q to final positions $q + \Psi_g$ along trajectories described by the tracer equation of motion

$$\ddot{\Psi}_g + \mathcal{H}\dot{\Psi}_g = -\nabla\Phi + \mathbf{F}_{b,g}, \quad (3.6)$$

where we have included a non-gravitational term, $F_{b,g}$, to account for the possibility that tracers feel non-gravitational forces. Such non-gravitational contributions may arise, for example, from the Compton drag on the baryonic component of galaxies, or from various galaxy formation processes. Since such contributions are always local in space and time, we will assume the above equation satisfies the same symmetries of eq. (3.2), i.e. the force acting on galaxies depends only on density fields and velocity gradients.

Equation (3.6) can be solved by subtracting the equation of motion of the total matter displacement (equation (2.6a)) and defining $\Psi_{r,g} = \Psi_g - \Psi_m$. Neglecting the baryonic contributions such that the tracers’ dynamics are governed only by gravity, and assuming that the initial tracer displacements are a weighted average of the baryonic and CDM displacements, i.e. $\Psi_{g,i} = \Psi_{m,i} + f_g \Psi_{r,i}$, this immediately yields the time evolution

$$\Psi_g(\tau) = \Psi_m(\tau) + f_g[\Psi_r(\tau)]_{\text{CD}=0}, \quad (3.7)$$

where the relative displacement is evaluated assuming zero Compton drag. Note that if we assume that the tracer field is made of objects composed of the same mass fractions of baryons and CDM as the total matter content of the universe, i.e. with $f_g = 0$, equation (3.7) reduces to the trajectory of the matter component. Similarly, objects composed purely of baryons or the CDM will (at the linear level) follow the baryon or CDM displacements, respectively.

We can alternatively think of eq. (3.7) as a bias expansion of the galaxy displacements in terms of the underlying fields, since Ψ_m and Ψ_r are the only two linear operators allowed by symmetries at lowest order in spatial derivatives. If the tracer flow is purely gravitational, as assumed above, the equivalence principle further restricts the coefficient of the total matter displacement — which encapsulates the motion due to the gravitational potential — to be exactly 1 at all times. However, this restriction can be broken by baryonic contributions ($\propto \mathbf{F}_{b,g}$) such as the Compton drag. As seen in the second term on the right hand side of equation (2.9), the acceleration due to Compton drag generates displacements proportional

to Ψ_m ; this contribution, on top of the aforementioned gravitational displacements, can lead to an expansion $\Psi_g = (1 + \alpha_{\text{CD}})\Psi_m + f_g\Psi_r + \dots$ for some nonzero coefficient α_{CD} due to Compton drag, where the total-matter coefficient deviates from unity. Consequences of this modified expansion for the power spectrum are considered at the end of section 4.1 and in figure 6. Other baryonic forces, such as pressure forces at small scales, can similarly be included as further terms ($\Psi_g \ni c_s^2 \nabla \delta_b$) in this expansion.

4 Galaxy power spectra in the Zeldovich approximation

4.1 Analytic form

From equation (3.1), the power spectrum at redshifts z for a biased tracer can be computed as

$$P_{gg}(k, z) = \int d^3q \, e^{i\mathbf{k}\cdot\mathbf{q}} \left\langle F_g[\mathbf{q}_1] F_g[\mathbf{q}_2] e^{i\mathbf{k}\cdot(\Psi_g(\mathbf{q}_1, z) - \Psi_g(\mathbf{q}_2, z))} \right\rangle_{q=|\mathbf{q}_2 - \mathbf{q}_1|}, \quad (4.1)$$

where the subscripts denote quantities evaluated at two points separated by q in Lagrangian space. It is important to note that the bias functions F_g are evaluated in terms of the linear modes m_+, r_\pm defined at the initial redshift z_i . In the Zeldovich approximation displacements are solved to linear order but the full mapping between initial and final times is kept. This amounts to keeping the displacement correlators exponentiated in what follows [46]. We will adopt the bias expansion in equation (3.5). We evaluate integrals involving F_g by functional differentiation in the usual manner [46, 47, 56]: we include a term (e.g. λX) in the exponential for each argument, X , of F_g and evaluate terms like X^n via $\partial^n / \partial \lambda^n$ of $\exp[\lambda X]$.

Under the above assumptions our task reduces to evaluating

$$e^{i\mathcal{M}} \equiv \left\langle \exp \left(i\mathbf{k} \cdot \Delta_g(z) + \lambda_{\delta_{m,1}} \delta_{m,1} + \lambda_{+,1} \nabla \mathbf{r}_{+,1} + \lambda_{-,1} \nabla \mathbf{r}_{-,1} + (1 \leftrightarrow 2) \right) \right\rangle \quad (4.2)$$

with numerical subscripts referring to Lagrangian coordinates, q_1 and q_2 , and

$$\Delta_g = \Psi_{g,1} - \Psi_{g,2} = \Psi_{m,1} - \Psi_{m,2} + f_g(\Psi_{r,1} - \Psi_{r,2}) \equiv \Delta_m + f_g \Delta_r \quad (4.3)$$

The function $e^{i\mathcal{M}}$ can be evaluated using the cumulant theorem as the exponential of the connected components. The Zeldovich approximation assumes linear dynamics, such that only quadratic terms survive

$$\begin{aligned} e^{i\mathcal{M}} = \exp \Bigg\{ & -\frac{1}{2} k_i k_j A_{ij}^{mm} - f_g k_i k_j A_{ij}^{rm} - \frac{f_g^2}{2} k_i k_j A_{ij}^{rr} \\ & + i\mathbf{k} \cdot ((\lambda_{\delta_{m,1}} + \lambda_{\delta_{m,2}})(U_{mm} + f_g U_{rm}) \\ & + (\lambda_{+,1} + \lambda_{+,2})(U_{m+} + f_g U_{r+}) + (\lambda_{-,1} + \lambda_{-,2})(U_{m-} + f_g U_{r-})) \\ & + (\lambda_{\delta_{m,1}} \lambda_{+,2} + (1 \leftrightarrow 2)) \xi_{\delta_m \nabla \mathbf{r}_+} + (\delta_m, \nabla \mathbf{r}_-) + (\delta_m, \delta_m) \\ & + (\nabla \mathbf{r}_+, \nabla \mathbf{r}_+) + (\nabla \mathbf{r}_+, \nabla \mathbf{r}_-) + (\nabla \mathbf{r}_-, \nabla \mathbf{r}_-) \Bigg\}, \end{aligned} \quad (4.4)$$

where we have defined

$$A_{ij}^{ab} = \left\langle \Delta_i^a(z) \Delta_j^b(z) \right\rangle, \quad U_i^{a\pm} = \left\langle \Delta_i^a(z) \nabla \cdot \mathbf{r}_\pm(\mathbf{q}_1) \right\rangle, \quad \xi_{ab} = \left\langle a(\mathbf{q}_1) b(\mathbf{q}_2) \right\rangle, \quad (4.5)$$

noting that the Δ 's carry an implicit redshift dependence while the other fields do not. For the total-matter component this redshift dependence is a direct growth factor scaling and

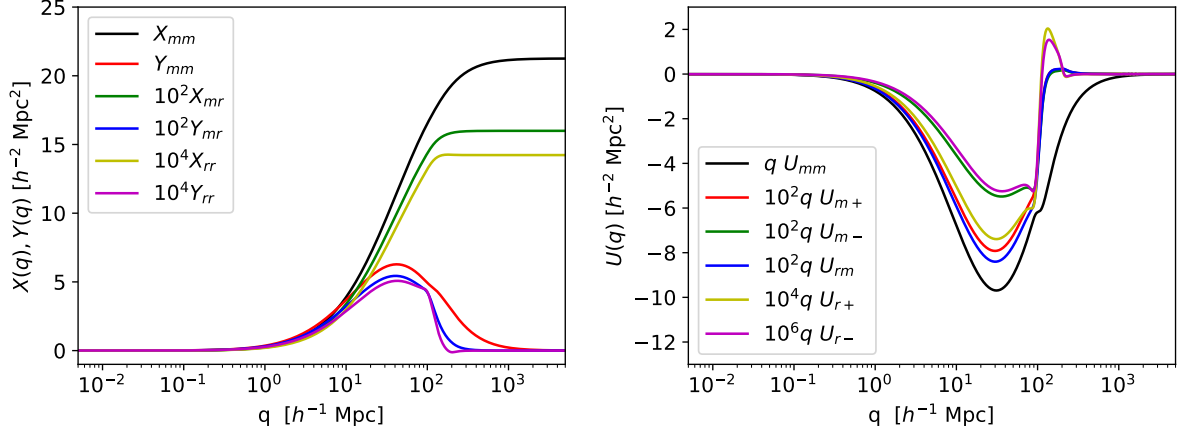


Figure 3. Correlation functions entering the galaxy power spectrum in eq. (4.7) at $z = 1.2$. Left panel: the displacement auto- and cross-correlation functions between the different components. Right panel: bias-weighted, displacement correlation functions. Correlation functions involving the relative component exhibit abrupt features around $q \sim 10^2 h^{-1} \text{ Mpc}$, reflecting the baryon acoustic oscillation scale.

we will for convenience take the linear field's value as evaluated at the observed redshift $\delta_m = -D_m(z)\nabla \cdot \mathbf{m}_+$. The paired parentheses denote terms similar to the preceding except with the indicated pair of variables. For example, in the third line

$$(\delta_m, \nabla \mathbf{r}_-) \equiv (\lambda_{\delta_m,1} \lambda_{-,2} + (1 \leftrightarrow 2)) \xi_{\delta_m \nabla \mathbf{r}_-}(q) \quad (4.6)$$

and when the elements of a pair are repeated the term should be divided by a symmetry factor of two.

Figure 3 shows the different correlation functions entering the above calculation. Since the correlation function of the different displacements fields, $A_{ij}^{ab}(q)$, is a tensor, we can decompose it as $A_{ij}^{ab}(q) = X^{ab}(q)\delta_{ij}^K + Y^{ab}(q)\hat{q}_i\hat{q}_j$, and the functions $X(q)$'s and $Y(q)$'s are shown in the left panel of figure 3. Clearly the galaxy displacements are dominated by the total matter component, with the relative terms contributing much less than a % to the bulk flows. This fact will enable us to treat the terms proportional to f_g perturbatively, as they will be much smaller than one for wavenumbers below the nonlinear scale defined by $k^2\Sigma^2 \lesssim 1$, where the Zeldovich r.m.s. displacement is $\Sigma \propto X_{mm}(q \rightarrow \infty)$. The same conclusions apply for the bias weighted displacements $U(q)$'s, shown on the right hand panels in figure 3, where $U_m(q) \gg U_{\pm}(q)$.

Working to linear order in the power spectrum we then have that the galaxy-galaxy power spectrum is given by

$$P_{gg}(k) = \int d^3q e^{i\mathbf{k}\cdot\mathbf{q}} e^{-\frac{1}{2}k_i k_j A_{ij}^{mm}} \left[1 - f_g k_i k_j A_{ij}^{rm} - \frac{f_g^2}{2} k_i k_j A_{ij}^{rr} \right. \\ + 2ik \cdot (b_1 U_{mm} + b_+ U_{m+} + b_- U_{m-}) \\ + 2f_g ik \cdot (b_1 U_{rm} + b_+ U_{r+} + b_- U_{r-}) \\ + b_m^2 \xi_{\delta_m \delta_m} + 2b_m b_+ \xi_{\nabla \mathbf{r}_+ \delta_m} + 2b_m b_- \xi_{\nabla \mathbf{r}_- \delta_m} \\ \left. + b_+^2 \xi_{\nabla \mathbf{r}_+ \nabla \mathbf{r}_+} + 2b_+ b_- \xi_{\nabla \mathbf{r}_+ \nabla \mathbf{r}_-} + b_-^2 \xi_{\nabla \mathbf{r}_- \nabla \mathbf{r}_-} + \mathcal{O}(P_L^2) \right]. \quad (4.7)$$

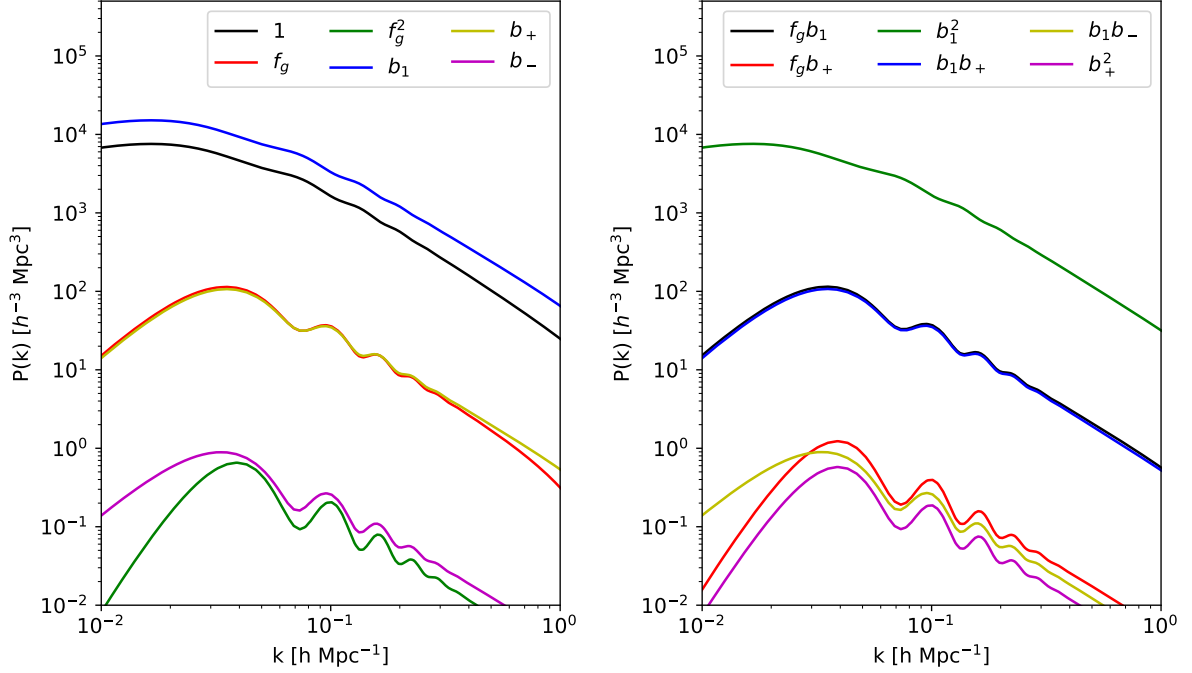


Figure 4. Different contributions to the galaxy power spectrum in the Zeldovich approximation, eq. (4.7), at $z = 1.2$. Terms proportional to $b_+ b_-$, $f_g b_-$, and b_-^2 have been omitted as they are two orders of magnitude smaller than the smallest contributions shown. Many terms, such as those involving f_g and b_+ , are essentially degenerate.

Figure 4 shows the different contributions to the galaxy power spectrum in the Zeldovich approximation at $z = 1.2$. The leading corrections to the total-matter power spectrum come at the roughly percent level from terms in equation (4.7) linear in \mathbf{r}_+ , i.e. in b_+ and f_g . These contributions are essentially degenerate, with differences due to the dynamical evolution of Ψ_r in the f_g term, as we will discuss in the next paragraph. Corrections quadratic in \mathbf{r}_+ or linear in \mathbf{r}_- enter at roughly the same size four orders of magnitude below the total-matter contributions.

An interesting consequence of the advection of biased tracers with $|f_g| > 0$ is the appearance of relative bias terms even if none were present in the initial Lagrangian bias expansion. To see this, we can take the low- k limit of eq. (4.7), neglecting for the moment non-gravitational contributions to $\Psi_r(q)$, and obtain up to $\mathcal{O}(P(k))$

$$\begin{aligned}
 P_{gg}(k, z) = & (1 + b_1)^2 P_{\delta_m \delta_m}(k) \\
 & + 2(1 + b_1)(b_+ + f_g) P_{m \nabla \mathbf{r}_+}(k) + 2(1 + b_1)(b_- + f_g D_r(z)) P_{m \nabla \mathbf{r}_-}(k) \\
 & + (b_+ + f_g)^2 P_{\nabla \mathbf{r}_+ \nabla \mathbf{r}_+}(k) + (b_- + f_g D_r(z))^2 P_{\nabla \mathbf{r}_- \nabla \mathbf{r}_-}(k) \\
 & + 2(b_+ + f_g)(b_- + f_g D_r(z)) P_{\nabla \mathbf{r}_+ \nabla \mathbf{r}_-}(k).
 \end{aligned} \tag{4.8}$$

We immediately recognize the familiar expression for the Eulerian linear bias, $b_1^E = 1 + b_1$, and that the relative density and velocity bias terms get renormalized by terms proportional to f_g . To make further contact with the existing literature employing the Eulerian formulation of the equations of motion [16, 17], we can identify the relative baryon dark-matter density perturbation δ_r with the divergence of \mathbf{r}_+ , $\delta_r \equiv \nabla \cdot \mathbf{r}_+$, and the relative baryon dark-matter velocity divergence θ_r with the divergence of \mathbf{r}_- , $\theta_r \equiv (1 + z)H_0 \nabla \cdot \mathbf{r}_-$. This implies that

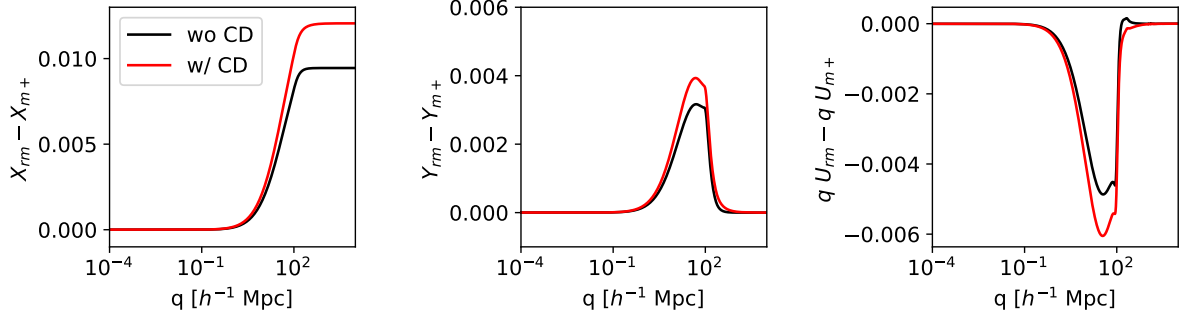


Figure 5. Comparison of two point functions with (red) and without (black) contributions from Compton drag. While the differences are small (cf. figure 3), they are non-negligible at large scales. The contributions from \mathbf{r}_+ have been subtracted off for ease of comparison.

the bias parameters in [16, 17] associated to the Eulerian fields are $b_{\delta_r}^E = b_+ + f_g$ and $b_{\theta_r}^E = (1+z)^{-1}H_0^{-1}(b_- + f_g D_r(z))$. Note that the referenced overdensities and velocities are those defined at the initial redshift z_i so should not be directly substituted for their Eulerian counterparts; for more details about the mapping of bias parameters from some initial time z_i to Eulerian coordinates see appendix A.

A final caveat occurs when the non-gravitational forces on the tracer, $\mathbf{F}_{b,g}$ are nonzero. The integrated effect of such forces on $\Psi_{r,g}$ must then be accounted for. For example, when dealing with baryons and dark matter, the effects of Compton drag on large scales are non-negligible. In this case, since the Compton drag force is proportional to the total-matter displacement, the two-point functions in eq. (4.5) involving Δ^r will gain a contribution proportional to Δ^m (figure 5). Such contributions can be non-negligible at large scales and can dominate in the contributions to the power spectrum proportional to f_g at low wavenumber (figure 6). Importantly, terms proportional to b_{\pm} are unaffected since they are related only to the primordial modes \mathbf{r}_{\pm} , breaking the degeneracy between f_g and b_+ . Since the difference between these terms is proportional to the total-matter component, this difference can alternatively be absorbed into the total-matter bias b_m [17]. Comparisons of these terms with and without Compton drag are shown in figure 6. Comparing the f_g contribution with and without Compton drag we see, as expected, that renormalizing the linear total-matter bias b_1 to include a contribution proportional to $f_g D_{CD}(z)$ (purple dotted curve) is sufficient to account for the non-gravitational Compton drag contributions.

4.2 Cross-spectra of different tracers and IR resummation

So far we have dealt only with tracer auto-spectra. The situation for cross-spectra is complicated by the non-cancellation of the IR-exponent at small separations, q . For two generic fluids, X and Y , such that $\Psi_{X,Y} = \Psi_m + f_{X,Y}\Psi_r$, the cross spectrum will take the form as in equation (2.2):

$$P_{XY}(k) = \int d^3q e^{i\mathbf{k}\cdot\mathbf{q}} e^{-\frac{1}{2}k_i k_j A_{ij}^{XY}} \left[\dots \right], \quad (4.9)$$

where the exponentiated two-point function A^{XY} is given by

$$A_{ij}^{XY}(q) = \left(\langle \Psi_i^X \Psi_j^X \rangle + \langle \Psi_i^Y \Psi_j^Y \rangle - 2\langle \Psi_i^X \Psi_j^Y \rangle \right) + \left(2\langle \Psi_i^X \Psi_j^Y \rangle - 2\langle \Psi_i^X(\mathbf{q}) \Psi_j^Y(0) \rangle \right), \quad (4.10)$$

and expectation values of point operators are displayed without arguments. Both terms in parentheses on the r.h.s. of equation (4.10) are well-defined and invariant under generalized

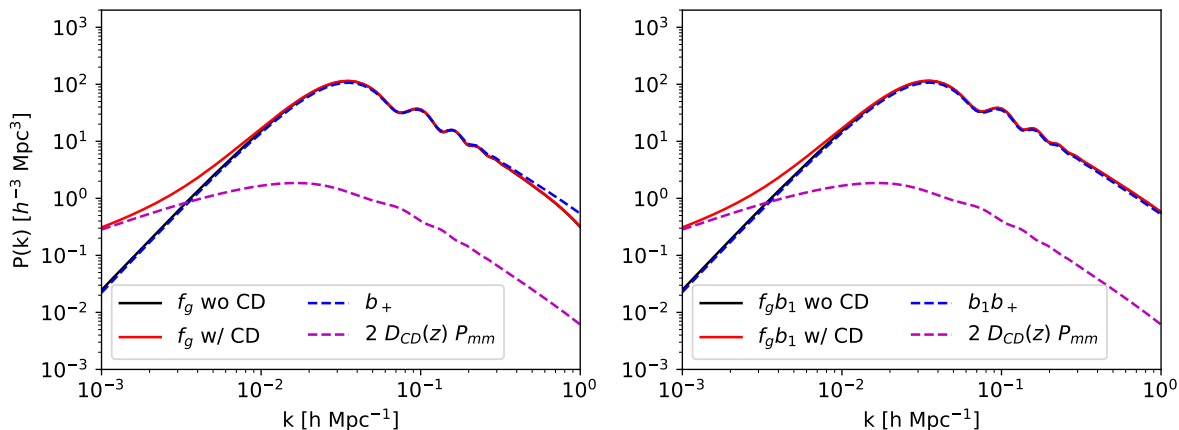


Figure 6. Comparison of terms involving b_+ (blue dashed) and f_g with and without Compton drag (red and black). The two are largely degenerate in the latter case, but with Compton drag the f_g terms are dominated by a contribution proportional to the total-matter power spectrum at large scales, which can alternatively be renormalized into the matter bias b_1 , shown separately as a dashed magenta curve. The left panel shows contributions due to contracting the relative components ($f_g \Psi_r$ or $b_+ \nabla \cdot \mathbf{m}_+$) with the total matter displacement Ψ_m , while the right panel shows contractions with the total matter bias $b_1 \delta_m$.

Galilean transformations; however the second term vanishes as $q \rightarrow 0$ while the first does not.³ As first noted in ref. [15], this is in contrast to the single-fluid case where A_{ij} had to vanish at small scales due to Galilean invariance.

In principle, the non-cancellation discussed above will introduce a large scale damping in the power spectrum at scales proportional to the difference $|\Psi^X - \Psi^Y|^2$. However, since $\Psi^{X,Y}$ are both expected to have the same coefficient in the total-matter component (i.e. unity) this difference squared will generically be proportional to $(f_X - f_Y)^2 \mathcal{O}(\Psi_r^2)$, and thus is suppressed by about four orders of magnitude relative to the Zeldovich displacement, Σ^2 , at the redshifts with which we are concerned ($z < 10$). On the other hand, while differential streaming is expected to damp cross spectra negligibly even if f_X is of order unity, as discussed in the previous section it will still generate an observable effect degenerate with the relative bias b_+ .

4.3 Higher order bias

Thus far we have not discussed the fact that any perturbative model should be considered an effective field theory, working up to some scale Λ [57–59]. This forces us to introduce a set of counterterms that remove the small scale sensitivities of the perturbative calculations. For instance all the $A_{ij}(q)$ terms contain a zero-lag piece computed at zero separation, i.e. $q = 0$, where perturbation theory breaks down. In the single fluid case, this UV-sensitivity is renormalized to lowest order in the power spectrum by a counterterm $c_s k^2 P_{ZA}(k)$ [29, 60], where the free parameter c_s has to be matched to simulations or data. The same structure of the counterterms appears in the two fluid scenario: for instance, the $A_{ij}^{ab}(q)$ required to calculate auto and cross spectra feature the same UV-sensitive contributions as $q \rightarrow 0$, requiring one value of c_s^a for each species. In principle, terms in the equations of motion due to the relative component will add additional UV sensitivities to our predictions; in practice, however, such

³A similar non-cancellation occurs in the modeling of BAO reconstruction, where the cross-term between the ‘displaced’ and ‘shifted’ fields exhibits the same behavior [33, 35].

contributions are subdominant in the dynamics of the relative component and negligible for the total-matter component (appendix B). To the extent that these contributions can be ignored, then, the two-fluid equations of motion can be renormalized identically to the single fluid case with one set of counterterms for each species or tracer. As counterterms have minor impact on BAO scales, and are anyway fitted to the data in both the single and multiple fluid cases, we do not include them in the Fisher calculation in the next section.

We have equally refrained from discussing bias beyond linear order. As in the equations of motion, contributions beyond first order in the linear power spectrum proportional only to the total-matter component can be added consistently as in the single-fluid case, and we will ignore small nonlinear contributions proportional to one or more powers of the relative component.⁴ However, one exception must be made: operators involving the relative-velocity between the baryon and dark matter squared, which, despite being at second order in the relative component, can be non-negligible due to their distinct dimensional scaling [11, 16, 18, 22]. Such contributions were the focus of the first studies of bias [11, 18, 22] in the two-fluid picture, and we will show how their calculation fits naturally into the Lagrangian framework. For a discussion of other second order bias operators see appendix B.2.

At second order in the bias expansion we can write

$$F_g[\Psi_m, \Psi_r | \mathbf{q}] \supset b_v \sigma_{v_r}^2 \frac{[\mathbf{v}_b(\mathbf{q}) - \mathbf{v}_c(\mathbf{q})]^2}{\sigma_{v_r}^2} = b_v \sigma_{v_r}^2 \frac{\mathbf{r}_-(\mathbf{q})^2}{\sigma_{r_-}^2} \equiv c_- [\mathbf{r}_-(\mathbf{q})]^2 \quad (4.11)$$

where $\sigma_{v_r}^2$ is the 1-point variance of the relative velocities and $\sigma_{r_-}^2 = (1+z)^{-2} H_0^{-2} \sigma_{v_r}^2$. As several authors [11, 16] have pointed out, baryon-dark matter relative velocities can be quite large at the time when the first halos and galaxies form, which could result in a large value of b_v for their late time descendants. The value of $b_v \sigma_{v_r}^2$ can be as large as 0.01, which will make this contribution at second order in the power spectrum larger than the b_- terms, even on linear scales. It is however worth remembering that a value of $b_v \sigma_{v_r}^2 \simeq 10^{-5}$ is also plausible, which would substantially reduce the importance of this contribution.

To consistently compute the power spectrum contributions due to $c_- \sim b_{v^2}$ we must go beyond the Zeldovich approximation. Up to 1-loop in Lagrangian perturbation theory we have to compute 4 new terms to properly include the new bias parameter c_- . Beyond these, terms proportional to c_-^2 can be safely neglected as they are $\mathcal{O}(P_{\nabla \mathbf{r}_-}^2)$. For the same reason we drop all the terms proportional to $b_{\pm} c_-$, as well as contributions of the relative component to the equations of motion. This leaves us with contributions proportional to c_- , $b_1 c_-$, $b_2 c_-$, and $b_{s^2} c_-$.

The first of these, proportional to c_- , contains a 1-loop contribution and is given by

$$P_{gg}(k) \supset c_- \int d^3 q e^{i\mathbf{k} \cdot \mathbf{q}} e^{-\frac{1}{2} k_i k_j A_{ij}^{mm}} \left(2i k_i \mathcal{U}_i(q) - \frac{1}{2} k_i k_j A_{ik}^{m-} A_{jk}^{m-} \right), \quad (4.12)$$

where we have defined $A_{ij}^{m-} = \langle \Delta_{m,i}(\mathbf{r}_{-,2} - \mathbf{r}_{-,1})_j \rangle$ and the 1-loop contribution from the second-order Lagrangian displacement $\Psi^{(2)}$ enters as

$$\mathcal{U}_i(q) \equiv \left\langle \Delta^{(2)} \mathbf{r}_{-,1}^2 \right\rangle = \hat{q}_i \int \frac{dk}{2\pi^2} k^2 Q_{v^2}(k) j_1(kq) \quad (4.13)$$

The kernel Q_{v^2} is derived in appendix C.

⁴A proper accounting of such terms would in addition require solving the relative-component equations of motion to beyond linear order, which is beyond our present scope.

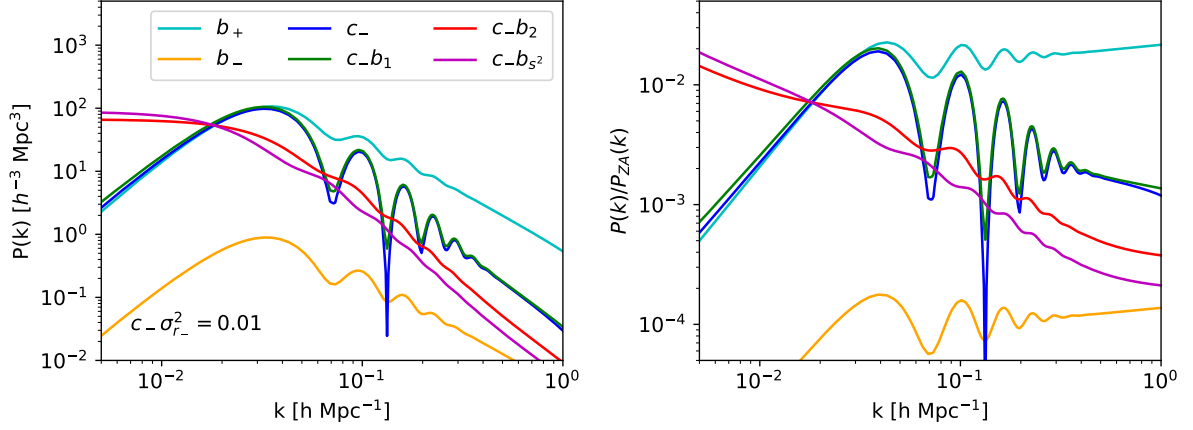


Figure 7. Contributions to the Zeldovich galaxy power spectrum from relative velocity bias at second order. All biases are set to unity except for c_- , which is set such that $b_v \sigma_{v_r}^2 = 0.01$ — in this case, the contributions from $b_{v,2}$ are seen to be quite comparable to those from b_+ , and moreover exhibit BAO “wiggles” far more prominently than does the regular ZA contribution.

The remaining terms do not contain loop contributions and follow straightforwardly from evaluating the second and third cumulants in eq. (4.1) within the Zeldovich approximation. These are those proportional to the first order bias:

$$P_{gg}(k) \supset 2ik_i b_1 c_- \int d^3q e^{i\mathbf{k}\cdot\mathbf{q}} e^{-\frac{1}{2}k_i k_j A_{ij}^{mm}} A_{ij}^{m-}(q) U_j^{-m}(q), \quad (4.14)$$

second order bias:

$$P_{gg}(k) \supset 2b_2 c_- \int d^3q e^{i\mathbf{k}\cdot\mathbf{q}} e^{-\frac{1}{2}k_i k_j A_{ij}^{mm}} U_i^{m-}(q) U_i^{m-}(q), \quad (4.15)$$

and shear

$$P_{gg}(k) \supset 4b_{s2} c_- \int d^3q e^{i\mathbf{k}\cdot\mathbf{q}} e^{-\frac{1}{2}k_i k_j A_{ij}^{mm}} W_{ijk}^{s-}(q) W_{ijk}^{s-}(q), \quad (4.16)$$

where we have defined the 2-point functions $U^{-m} \equiv \langle \mathbf{r}_-(q) \delta_m(0) \rangle = U^{m-}$ and $W_{ijk}^{s-}(q) = \langle s_{ij}(q) \mathbf{r}_{-,k}(0) \rangle$. Details of the above calculation can be found in appendix C.

The contributions proportional to c_- and their comparison with the 1-piece in eq. (4.7) and with the b_{\pm} ones computed in the previous section is shown in figure 7 for $z = 1.2$, assuming $b_v \sigma_{v_r}^2 = 0.01$. The c_- terms are indeed larger than the b_- terms on most scales, but still subdominant compared to the b_+ terms. Notably, the c_- terms feature significantly larger oscillatory features than contributions from b_{\pm} , with minima that differ from maxima by more than an order of magnitude.

5 Degeneracies and bias to BAO

Baryon acoustic oscillations (BAO) in the photon-baryon fluid before combination imprint a characteristic clustering scale in the distribution of galaxies that can be used as a standard ruler to constrain the cosmic expansion history [61]. In general this method is regarded as highly robust as it probes very large scales which are largely unaffected by astrophysical processes. However, relative component contributions to the two-point function also occur

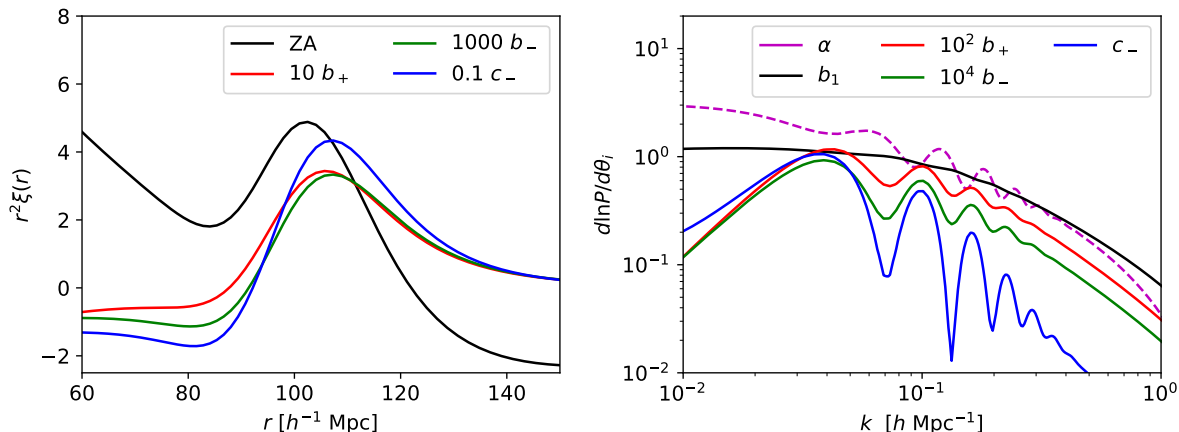


Figure 8. (Left) Contributions to the $z = 1.2$ correlation function from the various relative component biases, multiplied by constant factors for ease of comparison. All contributions have prominent features at the BAO scale, reflecting their origin in early-universe acoustic oscillations. (Right) Derivatives of the power spectrum with respect to these parameters and the BAO scale parameter α at $z = 1.2$, with $b_m = 0.5$, $b_2 = 0.2$, $b_+ = 1$, $b_- = 7$ and $c_- \sigma_{r-}^2 = 0.01$. Despite the fact that all these templates feature prominent oscillations, they nonetheless possess distinct scale dependence. Note that some of the derivatives have been multiplied by powers of ten for ease of comparison.

on very large scales and their oscillatory features, although arising from the same physical process of the standard BAO features in the matter density power spectrum, could bias our estimates of the distance scale if not properly taken into account [22–24]. Indeed, as shown in the left panel of figure 8, all the relative component contributions we have considered show distinct features around the BAO peak.

The extent to which contributions from the relative component can contaminate measurements of the BAO scale can be estimated using the Fisher matrix formalism [62]. The galaxy overdensity has a covariance that is diagonal in Fourier space and given by the power spectrum plus shot noise, $\hat{P}_{gg} = P_{gg}(k) + \bar{n}^{-1}$; for the parameters $\{\theta_i\}$, the Fisher matrix is given by

$$F_{ij} = V_{\text{obs}} \int \frac{d^3k}{(2\pi)^3} \frac{1}{2} \frac{\partial \ln \hat{P}_{gg}(k)}{\partial \theta_i} \frac{\partial \ln \hat{P}_{gg}(k)}{\partial \theta_j}, \quad (5.1)$$

where V_{obs} is the observed volume. For simplicity we neglect redshift space distortions and focus only on the isotropic BAO signal, though we will comment on how our Lagrangian analysis can be naturally extended to redshift space in the final paragraph. We model the power spectrum using the two-fluid Zeldovich terms derived above and include matter contributions up to one loop (see e.g. [52]), including contributions from the quadratic Lagrangian bias b_2 . We consider only scales between $k_{\text{min}} = 10^{-2} h \text{ Mpc}^{-1}$ and $k_{\text{max}} = 0.25 h \text{ Mpc}^{-1}$, and fiducial value of $b_1 = 0.53$ and $b_2 = 0.2$. The number density of galaxies is $\bar{n} = 4.2 \times 10^{-4} h^3 \text{ Mpc}^{-3}$ and we assume $V = 5 h^{-3} \text{ Gpc}^3$. These numbers are chosen to be similar to what galaxy surveys like DESI [2] or Euclid [3] are expected to measure, and in particular are based off the expected DESI ELG population at $z = 1.25$ in a bin of width $\Delta z = 0.1$ and 14,000 square degrees of observation.

To quantify the potential impact of the relative component on standard BAO analyses, we will compare two models of the power spectrum within the Fisher formalism: the “correct” model \mathbf{M}_1 , which is a function of all total-matter and relative component biases, and the

nested “standard” model \mathbf{M}_0 , wherein the relative component biases are set to zero (i.e. $b_{\pm}, c_- = 0$). The observed power spectrum is in addition a function of the BAO scaling parameter α such that

$$P_{gg}(k, z, \alpha, \mathbf{M}) = \alpha^{-3} P_{gg}\left(\frac{k}{\alpha}, z, \mathbf{M}\right). \quad (5.2)$$

The derivative of the baseline galaxy power spectrum with respect to the parameters is shown in figure 8. These templates all show oscillatory features of roughly the same frequency as the BAO scale but exhibit distinguishable scale dependence. For reference, applying eq. (5.1) returns sub-% error on the BAO scale, with $\sigma_{\alpha} = 0.9\%$, for the standard analysis using \mathbf{M}_0 .

We can now compute the systematic shifts in α that would be incurred by neglecting the relative component, i.e. by fitting to \mathbf{M}_0 . For convenience, we will split the parameters in \mathbf{M}_1 into $\theta = (\phi_a, \psi_{\sigma})$, where ϕ_a with Latin indices are the BAO scale and total-matter parameters and ψ_{σ} with Greek indices are the relative component biases, such that \mathbf{M}_0 is given by $\theta = (\phi_a, \psi_{\sigma} = 0)$. In this language the shift in α and b_1 due to using the standard model can be calculated to first order as [63]

$$\delta\theta_a = -(F_0)_{ab}^{-1} G_{b\sigma} \delta\psi_{\sigma}, \quad a, b = \alpha, b_1, \sigma = b_{\pm}, c_-. \quad (5.3)$$

Here F_0 and G are respectively diagonal and off-diagonal blocks of the full Fisher matrix $F = F(\theta_0)$ calculated at the best fit parameters θ_0 for the full model \mathbf{M}_1 , such that $F_{0,ab} = F_{ab}$ and $G_{b\sigma} = F_{b\sigma}$, and $\delta\psi$ is the deviation of ψ in the standard cold dark matter only model \mathbf{M}_0 from \mathbf{M}_1 , i.e. $\delta\psi = -\psi_0$.

As a simple first example, we consider a toy-model Universe in which the only relative contribution is b_+ . Figure 9 compares the “true” power spectrum, $P_1(k)$, assuming $b_+ = 5$, with best fits to the power spectrum in a dark matter only universe $P_0(k)$, described by the model parameters \mathbf{M}_0 , where the values of α, b_1, b_2 are shifted from their true values according to equation (5.3). Different values of the maximum wave number k_{\max} included in the Fisher calculation are shown with different lines. For $k_{\max} = 0.1 h \text{ Mpc}^{-1}$, we find a significant departure in phase between the two models, compared to higher limiting wavenumbers, as evident from the phase of the residual in the bottom panel. Beyond $k_{\max} = 0.15 h \text{ Mpc}^{-1}$ there are sufficient BAO wiggles that the phase of the residuals are essentially locked. We caution that the same exercise repeated with both matter and relative terms in the Zeldovich approximation can lead to wide swings in the BAO scale $\delta\alpha$ as a function of k_{\max} . This can be understood as follows: at $k \gtrsim 0.1 h \text{ Mpc}^{-1}$, b_+ contributes both oscillatory behavior and a broadband shape identical to the total matter component. The latter is essentially an amplitude change and can be roughly cancelled by a shift δb_m , which it is thus fixed independently of k_{\max} . This then requires $\delta\alpha$ to shift with k_{\max} as more oscillations are included until the oscillations in \mathbf{r}_+ relative to \mathbf{m}_+ are damped at large k (figure 2). This broadband effect is ameliorated by including nonlinear terms for BAO measurements, but the partial degeneracy of b_+ with the power spectrum amplitude likely implies that ignoring two-fluid effects may affect measurement of the amplitude of the power spectrum (though this effect will also be partially mitigated by redshift-space distortions).

The same formalism can be applied to more realistic bias models. In the upper left panel of figure 10 we consider the case when the observed power spectrum contains nonzero values b_{\pm} and $c_- = 0$, and forecast the shifts in α due to the wrong assumption of $b_{\pm} = 0$. Due to the small size of the b_- contributions (see figure 4), we expect shifts in BAO inferred distances to be dominated by b_+ , and this is indeed what we find, contours of constant $\delta\alpha$ are almost

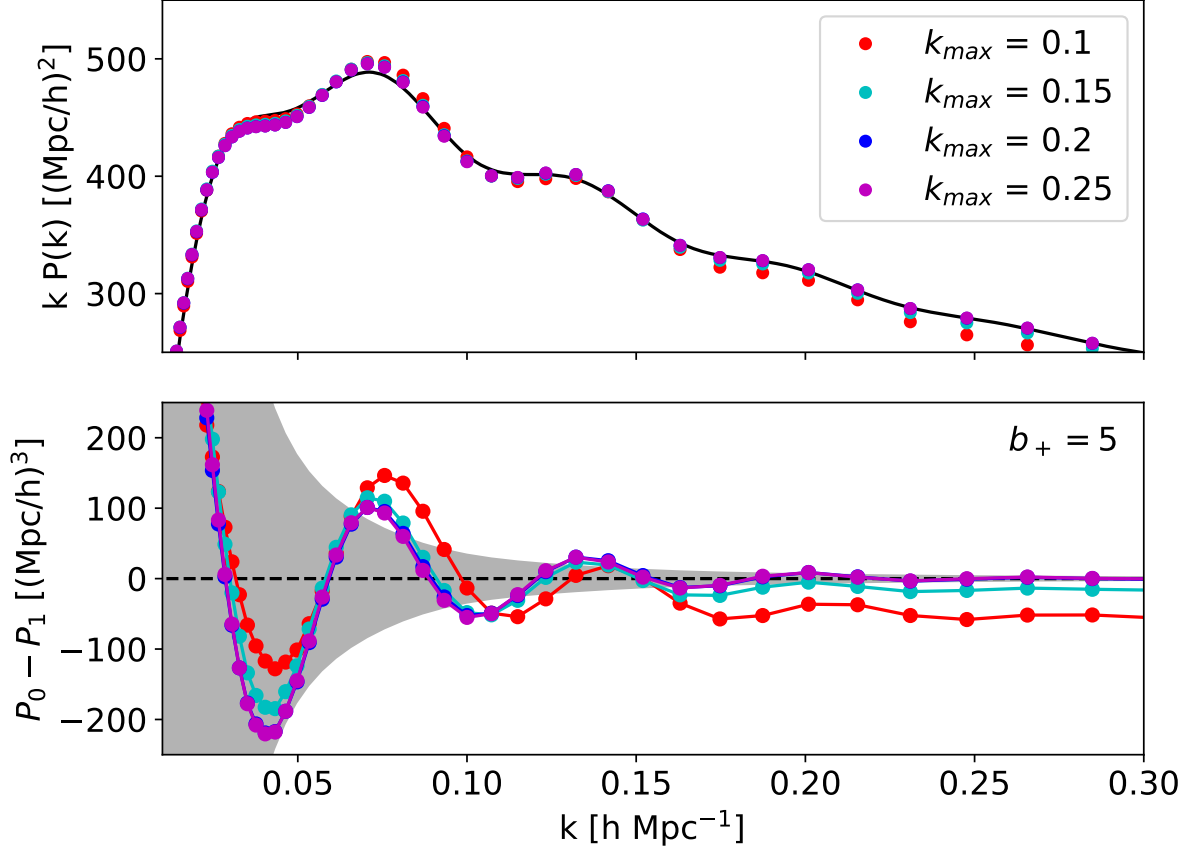


Figure 9. (Top) Best fit power spectra using the total-matter-component-only model, \mathbf{M}_0 , for a universe where $b_+ = 5$ with varying maximum fitted wave numbers k_{max} . (Bottom) Residuals of the above fits, compared to expected errors ($\Delta \ln k = 0.06$), shaded in gray. Fitting over too narrow a range ($k_{\text{max}} = 0.1 h \text{ Mpc}^{-1}$) results in a highly biased phase, while fits using larger wave number ranges covering more than one BAO wiggle are essentially in phase. The remaining oscillating residuals significantly exceed the expected error and are due to lack-of-fit for the oscillations in the relative component.

independent of b_- even when $|b_-| = 10$. On the other hand, we see that values of $b_+ \sim 5$ shift the measured α by up to 0.4%, close to half of the error on α expected when using \mathbf{M}_0 .

However, the physics behind the relative components is quite well understood and can be easily included in Fisher forecasts or power spectrum analyses. Indeed, as seen in figure 8, the templates for the various relative biases and α have distinct shape and could be distinguishable depending on the noise level of the measurements. The upper right plot in figure 10 shows the increase in σ_α induced by marginalizing over b_\pm in universes where b_\pm and b_v are not necessarily nonzero.⁵ The total loss of constraining power is modest, with less than 10% worse error bars even after marginalizing over two extra parameters. In both the computations of the shifts in α and the increase of $\sigma(\alpha)$, the volume of the survey does not enter, and the final results depend only on the shot noise levels.

In the lower set of plots in figure 8, we repeat the same exercise described above including c_- as an extra free parameter. Since b_- is irrelevant for the final results we set it to zero

⁵The nonzero $b_{\pm,v}$ produce shifts in the measured α, b_m when using \mathbf{M}_0 , which must be taken account when computing σ_α . To first order, the shifted Fisher matrix is given by F_0 .

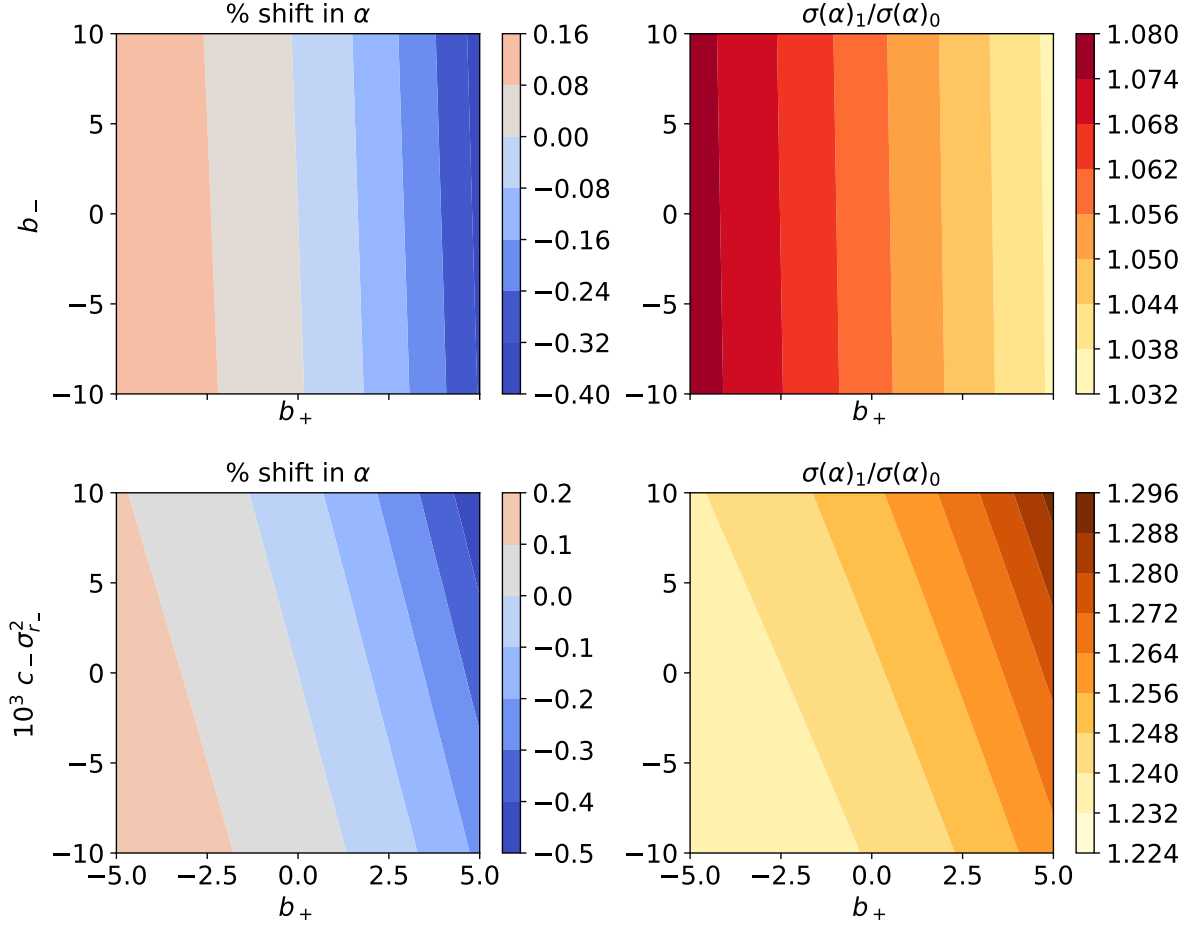


Figure 10. (Top Left) Shift in measured α when neglecting relative component biases as a function b_{\pm} in the absence of c_- . While b_- contributes negligibly, $b_+ = 5$ produces a shift up to a 0.4%. (Top Right) Ratio of error bars in α when marginalizing over b_{\pm} vs. when they are kept fixed at zero, such that the best-fit value of α is biased in the latter case. In the latter case the forecast takes into account the shift away from the true value due to incorrect model assumptions. (Bottom Row) Same as the above, but with c_- added as a nonzero parameter in \mathbf{M}_1 . We have set the true $b_- = 0$ for convenience but marginalize over it to calculate uncertainties. While even $c_- \sigma_{r-}^2 = 0.01$ contributes only a tenth of a percent to the shift in α , the error bars are inflated relative to the top row by up to twenty percent. We assume $k_{\max} = 0.25 h \text{ Mpc}^{-1}$ throughout.

(but still marginalized over it). We find that b_+ and c_- are anti-correlated, with larger shifts compared to the b_{\pm} case, but $\delta\alpha/\alpha \leq 0.5\%$ in all cases. Marginalizing over the extra parameter c_- results in a 20–30% increase in $\sigma(\alpha)$, which is still benign for BAO constraints. Our results therefore advocate for the implementation of relative component biases, at least of b_+ and c_- , in standard BAO data analysis of the galaxy power spectrum or correlation function.

Finally, in figure 11 we investigate the detectability of the two-fluid effects in the same setup. On their own, both b_+ and c_- become 1σ detectable at the upper end of our explored parameter ranges, shown as the red lines in figure 11. However, once all three relative bias parameters are marginalized over, the black set of curves in figure 11, neither will be detectable within our fiducial volumes, with c_- in particular at 0.1σ , well out of reach even if all the DESI redshift bins are combined.

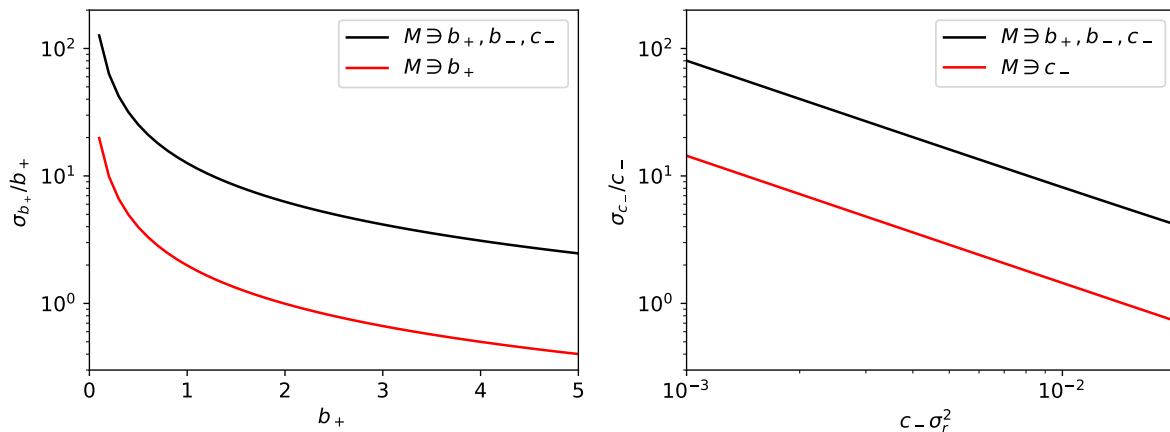


Figure 11. Constraints on b_+ and c_- in our fiducial setup if only each respective parameter can be varied (black), and if all relative parameters are simultaneously marginalized over (red). Notably, when the full model is taken into account detecting the relative velocity effect (c_-) will require up to ten times more signal to noise.

6 Conclusions

The large scale structure of the universe, whose formation is dominated by the dynamics of gravitational collapse, is one of the premier probes into fundamental physics. At subleading order, the presence of multiple particle species, broadly categorized into cold dark matter, baryons and neutrinos, with distinct properties beyond their shared gravitational attraction, can present additional features in this structure, which will become increasingly important as future surveys push to higher precision. In particular, relative perturbations between baryons and cold dark matter are prominent at the same scale as baryon acoustic oscillations and have the potential to cause systematic biases in future BAO measurements.

In this paper, we develop the Lagrangian formalism to calculate the clustering of biased tracers in the presence of multiple fluids, focusing specifically on the two-fluid scenario with dark matter and baryons. The Eulerian description of two-fluid dynamics has been studied extensively in the past and we make contact with previous work as appropriate throughout the text. LPT includes an automatic resummation over long-wavelength bulk flows and is thus able to accurately capture the shape of BAO features for biased tracers. In addition, LPT naturally maps bias terms from their initial Lagrangian positions to advected Eulerian positions, in contrast to Eulerian theory in which advective terms must be put in by hand, thereby simplifying the treatment of bias as responses to linear initial perturbations.

The presence of two fluids introduces terms beyond those encountered in traditional single fluid cosmological perturbation theory, with modifications in both the bias expansion and tracer advection. In the former, the generalized Galilean invariance that restricted the bias to contain only second derivatives of the gravitational potential in the single fluid case, allows terms including relative overdensities and velocities between different species. In the latter, initial relative displacements between various species are preserved under free fall and present an additional source of bias. Large scale non-gravitational forces such as Compton drag induced by the CMB can introduce additional corrections. We formulate modifications to tracer bias and advection in terms of three initial modes, constants of motion in the linear equations of motion, which roughly correspond to the initial total-matter displacement field and the relative displacement and velocity fields between dark matter and baryons.

We explicitly calculate the galaxy auto-power spectrum in the Zeldovich approximation within this formalism. Cross correlations between the relative modes introduce eight terms linear in the power spectrum — however, those quadratic in the relative component are suppressed by four orders of magnitude relative to the single fluid terms at low redshifts relevant for the next generation of galaxy surveys. Comparing to the Eulerian result explicitly to first order in the power spectrum, we find that the Eulerian relative component bias corresponds to linear mixtures of the Lagrangian bias, with modifications to the tracer advection entering both the Eulerian relative overdensity bias and the Eulerian relative velocity divergence bias. We then take up the calculation of cross spectra, finding a large scale damping due to an IR noncancellation in the relative component that is nonetheless negligibly small on perturbative scales. We also briefly discuss higher order corrections to the equations of motion in the presence of two fluids from an effective theory point of view, and perform an example one loop calculation for the relative velocity effect ($\propto \mathbf{v}_r^2$).

We conduct an exploratory analysis into whether two-fluid effects can cause systematic biases in measurements of the BAO scale. Taking the example of DESI ELGs at $z = 1.25$, we show that while ignoring two-fluid effects can lead to systematic shifts in the measured BAO scale as large as half a sigma, properly marginalizing over these effects induces less than ten percent loss in precision for a wide range of bias values. Since the scale dependence of the underlying physics is well understood, these results advocate for including two-fluid terms at linear order in future analyses. The dominant relative bias term ($\propto b_+$) does not fall quadratically with the growth factor like the total-matter contributions, and we therefore expect the relative bias signal as a fraction of total power to scale with redshift as $D_+^{-1}(z)$ and become proportionally more significant for surveys (such as the proposed Stage II 21-cm survey [5]) at higher redshifts. Studies of more highly biased tracers such as DESI quasars [2], for which the total-matter contributions are correspondingly larger, will on the other hand be less influenced by the relative bias for similar reasons.

While the Lagrangian picture is a natural playground for their study, in this paper we have opted not to study redshift space distortions (RSD). We note, however, that of the two relative components, \mathbf{r}_+ is dominant but stationary while \mathbf{r}_- is so small as to be essentially negligible — two-fluid impacts should thus have a relatively small impact on RSD. However, as noted in the previous section, since the dominant relative component contribution b_+ is somewhat degenerate with the overall power spectrum amplitude, it is possible that two-fluid effects could hinder the accuracy of $f\sigma_8$ measurements beyond the percent level. We will return to this issue in future work.

Acknowledgments

We thank Jonathan Blazek and Jahmour Givans for helpful comments. SC is supported by the National Science Foundation Graduate Research Fellowship (Grant No. DGE 1106400) and by the UC Berkeley Theoretical Astrophysics Center Astronomy and Astrophysics Graduate Fellowship. M.W. is supported by the U.S. Department of Energy and by NSF grant number 1713791. This work made extensive use of the NASA Astrophysics Data System and of the `astro-ph` preprint archive at [arXiv.org](https://arxiv.org).

A Redshift dependence and size of the of bias parameters

The bias expansion in equation (3.5) has an implicit dependence on the initial redshift z_i that must be taken into account to reach consistent conclusions. Since the initial conditions mix at most linearly, no information can be lost by choosing one initial time τ_i over another; for example, the sensitivity of halos to the relative velocity divergence after reionization, which contains a contribution from the total matter overdensity (eq. (2.9)), can be directly accounted for by calibrating the bias parameter for δ_m at an earlier redshift.

As a simple example we consider the redshift dependence of the relative components in the sourceless ($F_b = 0$) case. If we set our initial time at τ'_i instead of τ_i we will get

$$\Psi_r(\tau) = \left(-\mathbf{r}_+ + \mathbf{r}_- D_r(\tau'_i, \tau_i) \right) + \mathbf{r}_- D_r(\tau, \tau'_i) \equiv -\mathbf{r}'_+ + \mathbf{r}'_- D_r(\tau, \tau'_i). \quad (\text{A.1})$$

Re-expanding F_g at τ'_i thus yields

$$F_g(\mathbf{q}) = b_1 \delta_m + b'_+ \nabla \cdot \left(\mathbf{r}_+ - \mathbf{r}_- D_r(\tau'_i, \tau_i) \right) + b'_- \nabla \cdot \mathbf{r}'_- + \dots \quad (\text{A.2})$$

Since b' and b apply to the same field configurations at different times, they must yield the same initial overdensity F_g — this requirement can be satisfied by enforcing the differential equations

$$\frac{db_+}{d\tau} = 0, \quad \frac{db_-}{d\tau} = \frac{b_+}{a(\tau)}. \quad (\text{A.3})$$

Intriguingly, the presence of a relative overdensity bias can “generate” a relative velocity bias at later times. This can be understood as follows: the relative overdensity at late times is a linear combination of the relative overdensity and velocities at earlier times. Similar, though more complicated, versions of this relation hold when $F_b \propto \mathbf{m}_+$, in which case mixing of all three initial fields must be taken into account.

B Beyond linear order

B.1 Equations of motions

In this appendix we derive the equations of motion beyond linear order in the two-fluid scenario, and show that the nonlinear contribution of \mathbf{r}_\pm to the total-matter component are quadratically suppressed, and that the nonlinear relative component is always sourced by at least one component of \mathbf{r}_\pm .

The Lagrangian equations of motion at higher order can be found by taking the real-space divergence of both equations in (2.3). To do so we make use of the identities

$$1 + \delta^a(x, t) = \left| \frac{\partial x^a}{\partial q} \right|^{-1} = \mathcal{J}^a(q, t)^{-1} \quad (\text{B.1})$$

$$\nabla_{x^a} \cdot V = \left[\frac{\partial x^a}{\partial q} \right]_{ij}^{-1} \frac{\partial V_i}{\partial q_j} = [\delta_{ij} + \Psi_{i,j}^a]^{-1} V_{i,j}, \quad (\text{B.2})$$

where the negative powers in the second line denote matrix inverses, to account for the coordinate transformations between Lagrangian coordinates q and the fluid trajectories for each species $x^a = q + \Psi^a$, as well as the standard matrix identities $(I + A)^{-1} = I - A + A^2 - A^3 + \mathcal{O}(A^4)$ and $\det(I + A) = 1 + \text{Tr}[A] + \frac{1}{2}(\text{Tr}[A]^2 - \text{Tr}[A^2]) + \mathcal{O}(A^3)$. We will neglect the

effects of Compton drag, which affects the relative displacement at a few percent level even at late times and on linear scales and thus enter into our final power spectra at the same order of magnitude as the relative component squared, and assume potential flow.

The above equations imply that the Lagrangian equations of motion for the fluid displacements of a species a at n^{th} order takes the generic form

$$D\Psi_{i,i}^{(a,n)} = - \sum_{m=1}^{n-1} F^{(a,n-m)} D\Psi^{(a,m)} + \frac{3}{2} \mathcal{H}^2 \Omega_m \sum_{a'} w_{a'} \left(\frac{\mathcal{J}-1}{\mathcal{J}} \right)^{(a',n)}, \quad (\text{B.3})$$

where the superscript (a, n) denotes the species and order of each term, the derivative operator D is defined such that $DX = X'' + \mathcal{H}X'$, and the $F^{(a,n)}$'s are kernels composed of displacements of the species a at order n and below. Switching to the total matter and relative components, we have

$$D\Psi_{i,i}^{(m,n)} = - \sum_{a'} w_{a'} \sum_{m=1}^{n-1} F^{(a',n-m)} D\Psi^{(a',m)} + \frac{3}{2} \mathcal{H}^2 \Omega_m \sum_{a'} w_{a'} \left(\frac{\mathcal{J}-1}{\mathcal{J}} \right)^{(a',n)} \quad (\text{B.4})$$

$$D\Psi_{i,i}^{(r,n)} = - \left(\sum_{m=1}^{n-1} F^{(c,n-m)} D\Psi^{(c,m)} - \sum_{m=1}^{n-1} F^{(b,n-m)} D\Psi^{(b,m)} \right). \quad (\text{B.5})$$

We can derive some elementary properties of these equations without solving for their particular forms using symmetry arguments. Noting that the r.h.s. of (B.4) is symmetric under species index exchange ($b \leftrightarrow c$) while (B.5) is antisymmetric, we can conclude that (1) the first relative contribution to the total matter EOM at each order must be of order $\mathcal{O}(\mathbf{r}_{\pm}^2)$, and all subsequent contributions suppressed by further even powers of the linear relative component and (2) the relative component is always sourced by at least one power of \mathbf{r}_{\pm} , since the total matter component is even under this swap while the relative component is even. Note that (1) implies that the dynamics of the total-matter displacement are affected by the relative component only at the percent-of-a-percent level, and (2) implies that the nonlinear relative displacement is never less suppressed in \mathbf{r}_{\pm} than the linear solution. A similar result occurs in Eulerian theory, as described in [15].

For completeness, the explicit second and third order equations of motion for the relative component are, up to first order in the linear relative perturbation

$$\begin{aligned} D\Psi_{i,i}^{(r,2)} &= \Psi_{j,i}^{(r,1)} D\Psi_{i,j}^{(m,1)} \\ D\Psi_{i,i}^{(r,3)} &= \Psi_{j,i}^{(r,1)} D\Psi_{i,j}^{(m,2)} + \Psi_{j,i}^{(m,1)} D\Psi_{i,j}^{(r,2)} + \Psi_{j,i}^{(r,2)} D\Psi_{i,j}^{(m,1)} \\ &\quad - \left(\Psi_{j,k}^{(r,1)} \Psi_{k,i}^{(m,1)} + \Psi_{j,k}^{(m,1)} \Psi_{k,i}^{(r,1)} \right) D\Psi_{i,j}^{(m,1)}, \end{aligned} \quad (\text{B.6})$$

which are in-line with the symmetry arguments outlined above. The equations of motion for the total matter component at second and third order can be similarly verified to be simply the equations of motion in the one-fluid case to this order in the relative component.

B.2 Biasing at second order

Below we list all contributions to the bias expansion up to second order in the initial fields omitting derivative corrections:

$$\begin{aligned}
F_g = & b_1 \delta_m + b_{\delta_r} \delta_r + b_{\theta_r} \theta_r \\
& + \frac{1}{2} b_2 \delta_m^2 + b_{s^2} s_{ij} s_{ij} + b_{\delta_m \delta_r} \delta_m \delta_r + b_{\delta_m \theta_r} \delta_m \theta_r + b_{v_r \partial \delta_m} (\mathbf{v}_r)_i \partial_i \delta_m + b_{s \partial v} \partial_i (\mathbf{v}_r)_j s_{ij} \\
& + b_{v_r^2} \mathbf{v}_r^2 + \dots
\end{aligned} \tag{B.7}$$

In the main body of this paper we consider relative bias terms up to first order in the power spectrum, since even these represent only percent level effects, with the exception of the relative velocity effect $\propto \mathbf{v}_r^2$, which has a distinct scaling. As noted in the text, we note that the presence of Compton drag can introduce additional terms due to loss of gauge redundancy; we refer readers to the extensive discussion in [17].

C Relative velocity bias terms

In this appendix we provide details for the contributions of the relative velocity bias b_{v^2} at $\mathcal{O}(P^2)$ to the galaxy power spectrum. These contributions require the calculation of two new 2-point functions, the one-loop correlation between matter displacements and the squared relative velocity, and the correlation function between the shear field s_{ij} and the relative velocity. We describe these in turn.

The second order solution to the total-matter displacement (correct up to first order in the relative component) is given by

$$\Psi_i^{(2)}(k) = \frac{1}{2} \frac{3}{7} \frac{i k_i}{k^2} \int \frac{d^3 p}{(2\pi)^3} \left[1 - \left(\frac{(k-p) \cdot p}{|k-p||p|} \right)^2 \right] \delta_{m,0}(p) \delta_{m,0}(k-p), \tag{C.1}$$

and more simply the “normalized” relative velocity at first order is given by

$$\mathbf{r}_{-,i}(k) = \frac{-i k_i}{k^2} (\nabla \cdot \mathbf{r}_-)(k). \tag{C.2}$$

From this we can calculate the two-point function

$$\begin{aligned}
\langle \Psi^{(2)}(q) \mathbf{r}_-^2(0) \rangle = & \frac{3}{7} \int \frac{d^3 k}{(2\pi)^3} e^{ik \cdot q} \left(\frac{-i k_i}{k^2} \right) \frac{d^3 p}{(2\pi)^3} \\
& \left[1 - \left(\frac{(k-p) \cdot p}{|k-p||p|} \right)^2 \right] \frac{p \cdot (k-p)}{p^2 (k-p)^2} P_{\delta_m \nabla \mathbf{r}_-}(p) P_{\delta_m \nabla \mathbf{r}_-}(k-p),
\end{aligned} \tag{C.3}$$

which can be simplified to give

$$\langle \Psi^{(2)}(q) v_{r,0}^2(0) \rangle = \hat{q} \int \frac{d^3 k}{(2\pi)^3} e^{ik \cdot q} Q_{v^2}(k), \tag{C.4}$$

where the kernel is defined as

$$Q_{v^2}(k) \equiv \frac{3}{7} \int_0^\infty dr P_{m-}(kr) \int_{-1}^1 \frac{dx}{4\pi^2} \frac{r(x-r)(1-x^2)}{(1+r^2-2rx)^2} P_{m-}\left(k\sqrt{1+r^2-2rx}\right). \tag{C.5}$$

Next, the shear-velocity correlation function W_{ijk}^{s-} is given in Fourier space by

$$W_{ijk}^{s-}(q) = i \int \frac{d^3k}{(2\pi)^2} e^{ik \cdot q} \left(\frac{k_i k_j k_k}{k^4} - \frac{1}{3} \delta_{ij} \frac{k_k}{k^2} \right) P_{m-}(k) \equiv \tilde{W}_{ijk}^{s-}(q) - \frac{1}{3} \delta_{ij} U_k^{m-}(q), \quad (\text{C.6})$$

where in the last equality we have split W_{ijk}^{s-} into a totally-symmetric piece and a familiar piece proportional to U^{m-} . The former can be decomposed into scalar components

$$\tilde{W}_{ijk}^{s-}(q) = \mathcal{A}(q) \hat{q}_i \hat{q}_j \hat{q}_k + \mathcal{B}(q) (\hat{q}_i \delta_{jk} + \hat{q}_j \delta_{ki} + \hat{q}_k \delta_{ij}), \quad (\text{C.7})$$

with the scalar components defined as spherical Bessel transformations:

$$\mathcal{A}(q) = \int \frac{dk}{2\pi^2} k j_3(kq) P_{m-}(k) \quad (\text{C.8})$$

$$\mathcal{B}(q) = - \int \frac{dk}{2\pi^2} k \frac{1}{5} (j_1(kq) + j_3(kq)) P_{m-}(k). \quad (\text{C.9})$$

References

- [1] D.H. Weinberg, M.J. Mortonson, D.J. Eisenstein, C. Hirata, A.G. Riess and E. Rozo, *Observational Probes of Cosmic Acceleration*, *Phys. Rept.* **530** (2013) 87 [[arXiv:1201.2434](#)] [[INSPIRE](#)].
- [2] DESI collaboration, *The DESI Experiment Part I: Science, Targeting and Survey Design*, [arXiv:1611.00036](#) [[INSPIRE](#)].
- [3] L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **21** (2018) 2 [[arXiv:1606.00180](#)] [[INSPIRE](#)].
- [4] WFIRST collaboration, *WFIRST Science Investigation Team “Cosmology with the High Latitude Survey” Annual Report 2017*, [arXiv:1804.03628](#) [[INSPIRE](#)].
- [5] COSMIC VISIONS 21 CM collaboration, *Inflation and Early Dark Energy with a Stage II Hydrogen Intensity Mapping experiment*, [arXiv:1810.09572](#) [[INSPIRE](#)].
- [6] S. Saito, M. Takada and A. Taruya, *Nonlinear power spectrum in the presence of massive neutrinos: perturbation theory approach, galaxy bias and parameter forecasts*, *Phys. Rev. D* **80** (2009) 083528 [[arXiv:0907.2922](#)] [[INSPIRE](#)].
- [7] M. Shoji and E. Komatsu, *Massive Neutrinos in Cosmology: Analytic Solutions and Fluid Approximation*, *Phys. Rev. D* **81** (2010) 123516 [Erratum *ibid.* **D 82** (2010) 089901] [[arXiv:1003.0942](#)] [[INSPIRE](#)].
- [8] D. Blas, M. Garny, T. Konstandin and J. Lesgourgues, *Structure formation with massive neutrinos: going beyond linear theory*, *JCAP* **11** (2014) 039 [[arXiv:1408.2995](#)] [[INSPIRE](#)].
- [9] E. Castorina, C. Carbone, J. Bel, E. Sefusatti and K. Dolag, *DEMNI: The clustering of large-scale structures in the presence of massive neutrinos*, *JCAP* **07** (2015) 043 [[arXiv:1505.07148](#)] [[INSPIRE](#)].
- [10] L. Senatore and M. Zaldarriaga, *The Effective Field Theory of Large-Scale Structure in the presence of Massive Neutrinos*, [arXiv:1707.04698](#) [[INSPIRE](#)].
- [11] N. Dalal, U.-L. Pen and U. Seljak, *Large-scale BAO signatures of the smallest galaxies*, *JCAP* **11** (2010) 007 [[arXiv:1009.4704](#)] [[INSPIRE](#)].
- [12] G. Somogyi and R.E. Smith, *Cosmological perturbation theory for baryons and dark matter I: one-loop corrections in the RPT framework*, *Phys. Rev. D* **81** (2010) 023524 [[arXiv:0910.5220](#)] [[INSPIRE](#)].

- [13] D. Tseliakhovich and C. Hirata, *Relative velocity of dark matter and baryonic fluids and the formation of the first structures*, *Phys. Rev. D* **82** (2010) 083520 [[arXiv:1005.2416](#)] [[INSPIRE](#)].
- [14] F. Bernardeau, N. Van de Rijdt and F. Vernizzi, *Power spectra in the eikonal approximation with adiabatic and nonadiabatic modes*, *Phys. Rev. D* **87** (2013) 043530 [[arXiv:1209.3662](#)] [[INSPIRE](#)].
- [15] M. Lewandowski, A. Perko and L. Senatore, *Analytic Prediction of Baryonic Effects from the EFT of Large Scale Structures*, *JCAP* **05** (2015) 019 [[arXiv:1412.5049](#)] [[INSPIRE](#)].
- [16] F. Schmidt, *Effect of relative velocity and density perturbations between baryons and dark matter on the clustering of galaxies*, *Phys. Rev. D* **94** (2016) 063508 [[arXiv:1602.09059](#)] [[INSPIRE](#)].
- [17] F. Schmidt and F. Beutler, *Imprints of Reionization in Galaxy Clustering*, *Phys. Rev. D* **96** (2017) 083533 [[arXiv:1705.07843](#)] [[INSPIRE](#)].
- [18] J. Yoo, N. Dalal and U. Seljak, *Supersonic Relative Velocity Effect on the Baryonic Acoustic Oscillation Measurements*, *JCAP* **07** (2011) 018 [[arXiv:1105.3732](#)] [[INSPIRE](#)].
- [19] E. Castorina, E. Sefusatti, R.K. Sheth, F. Villaescusa-Navarro and M. Viel, *Cosmology with massive neutrinos II: on the universality of the halo mass function and bias*, *JCAP* **02** (2014) 049 [[arXiv:1311.1212](#)] [[INSPIRE](#)].
- [20] M. LoVerde, *Halo bias in mixed dark matter cosmologies*, *Phys. Rev. D* **90** (2014) 083530 [[arXiv:1405.4855](#)] [[INSPIRE](#)].
- [21] J.B. Muñoz and C. Dvorkin, *Efficient Computation of Galaxy Bias with Neutrinos and Other Relics*, *Phys. Rev. D* **98** (2018) 043503 [[arXiv:1805.11623](#)] [[INSPIRE](#)].
- [22] J. Blazek, J.E. McEwen and C.M. Hirata, *Streaming velocities and the baryon-acoustic oscillation scale*, *Phys. Rev. Lett.* **116** (2016) 121303 [[arXiv:1510.03554](#)] [[INSPIRE](#)].
- [23] F. Beutler, U. Seljak and Z. Vlah, *Constraining the relative velocity effect using the Baryon Oscillation Spectroscopic Survey*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2723 [[arXiv:1612.04720](#)] [[INSPIRE](#)].
- [24] Z. Slepian et al., *Constraining the baryon-dark matter relative velocity with the large-scale three-point correlation function of the SDSS BOSS DR12 CMASS galaxies*, *Mon. Not. Roy. Astron. Soc.* **474** (2018) 2109 [[arXiv:1607.06098](#)] [[INSPIRE](#)].
- [25] N. Padmanabhan and M. White, *Calibrating the Baryon Oscillation Ruler for Matter and Halos*, *Phys. Rev. D* **80** (2009) 063508 [[arXiv:0906.1198](#)] [[INSPIRE](#)].
- [26] B.D. Sherwin and M. Zaldarriaga, *The Shift of the Baryon Acoustic Oscillation Scale: A Simple Physical Picture*, *Phys. Rev. D* **85** (2012) 103523 [[arXiv:1202.3998](#)] [[INSPIRE](#)].
- [27] N. McCullagh and A.S. Szalay, *Nonlinear Behavior of Baryon Acoustic Oscillations from the Zel'dovich Approximation Using a Non-Fourier Perturbation Approach*, *Astrophys. J.* **752** (2012) 21 [[arXiv:1202.1306](#)] [[INSPIRE](#)].
- [28] M. White, *The Zel'dovich approximation*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 3630 [[arXiv:1401.5466](#)] [[INSPIRE](#)].
- [29] R.A. Porto, L. Senatore and M. Zaldarriaga, *The Lagrangian-space Effective Field Theory of Large Scale Structures*, *JCAP* **05** (2014) 022 [[arXiv:1311.2168](#)] [[INSPIRE](#)].
- [30] Z. Vlah, U. Seljak, M.Y. Chu and Y. Feng, *Perturbation theory, effective field theory and oscillations in the power spectrum*, *JCAP* **03** (2016) 057 [[arXiv:1509.02120](#)] [[INSPIRE](#)].
- [31] L. Senatore and G. Trevisan, *On the IR-Resummation in the EFTofLSS*, *JCAP* **05** (2018) 019 [[arXiv:1710.02178](#)] [[INSPIRE](#)].

- [32] D.J. Eisenstein, H.-j. Seo, E. Sirko and D. Spergel, *Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak*, *Astrophys. J.* **664** (2007) 675 [[astro-ph/0604362](#)] [[INSPIRE](#)].
- [33] N. Padmanabhan, M. White and J.D. Cohn, *Reconstructing Baryon Oscillations: A Lagrangian Theory Perspective*, *Phys. Rev. D* **79** (2009) 063523 [[arXiv:0812.2905](#)] [[INSPIRE](#)].
- [34] Y. Noh, M. White and N. Padmanabhan, *Reconstructing baryon oscillations*, *Phys. Rev. D* **80** (2009) 123501 [[arXiv:0909.1802](#)] [[INSPIRE](#)].
- [35] M. White, *Reconstruction within the Zeldovich approximation*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 3822 [[arXiv:1504.03677](#)] [[INSPIRE](#)].
- [36] B.D. Sherwin and M. White, *The Impact of Wrong Assumptions in BAO Reconstruction*, *JCAP* **02** (2019) 027 [[arXiv:1808.04384](#)] [[INSPIRE](#)].
- [37] Ya. B. Zeldovich, *Gravitational instability: An Approximate theory for large density perturbations*, *Astron. Astrophys.* **5** (1970) 84 [[INSPIRE](#)].
- [38] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, *Phys. Rept.* **367** (2002) 1 [[astro-ph/0112551](#)] [[INSPIRE](#)].
- [39] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [[astro-ph/9911177](#)] [[INSPIRE](#)].
- [40] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **07** (2011) 034 [[arXiv:1104.2933](#)] [[INSPIRE](#)].
- [41] A. Pontzen, *Scale-dependent bias in the baryonic-acoustic-oscillation-scale intergalactic neutral hydrogen*, *Phys. Rev. D* **89** (2014) 083010 [[arXiv:1402.0506](#)] [[INSPIRE](#)].
- [42] S. Gontcho A Gontcho, J. Miralda-Escudé and N.G. Busca, *On the effect of the ionizing background on the Ly α forest autocorrelation function*, *Mon. Not. Roy. Astron. Soc.* **442** (2014) 187 [[arXiv:1404.7425](#)] [[INSPIRE](#)].
- [43] P.U. Sanderbeck, V. Irsic, M. McQuinn and A. Meiksin, *Estimates for the impact of Ultraviolet Background fluctuations on galaxy clustering measurements*, *Mon. Not. Roy. Astron. Soc.* **485** (2019) 5059 [[arXiv:1810.12321](#)] [[INSPIRE](#)].
- [44] G. Cabass and F. Schmidt, *A new scale in the bias expansion*, *JCAP* **19** (2019) 031 [[arXiv:1812.02731](#)] [[INSPIRE](#)].
- [45] P. Catelan, F. Lucchin, S. Matarrese and C. Porciani, *The bias field of dark matter halos*, *Mon. Not. Roy. Astron. Soc.* **297** (1998) 692 [[astro-ph/9708067](#)] [[INSPIRE](#)].
- [46] T. Matsubara, *Resumming Cosmological Perturbations via the Lagrangian Picture: One-loop Results in Real Space and in Redshift Space*, *Phys. Rev. D* **77** (2008) 063530 [[arXiv:0711.2521](#)] [[INSPIRE](#)].
- [47] T. Matsubara, *Nonlinear perturbation theory with halo bias and redshift-space distortions via the Lagrangian picture*, *Phys. Rev. D* **78** (2008) 083519 [Erratum *ibid.* **D 78** (2008) 109901] [[arXiv:0807.1733](#)] [[INSPIRE](#)].
- [48] T. Baldauf, V. Desjacques and U. Seljak, *Velocity bias in the distribution of dark matter halos*, *Phys. Rev. D* **92** (2015) 123507 [[arXiv:1405.5885](#)] [[INSPIRE](#)].
- [49] V. Desjacques, M. Crocce, R. Scoccimarro and R.K. Sheth, *Modeling scale-dependent bias on the baryonic acoustic scale with the statistics of peaks of Gaussian random fields*, *Phys. Rev. D* **82** (2010) 103529 [[arXiv:1009.3449](#)] [[INSPIRE](#)].
- [50] E. Castorina, A. Paranjape, O. Hahn and R.K. Sheth, *Excursion set peaks: the role of shear*, [[arXiv:1611.03619](#)] [[INSPIRE](#)].

- [51] C. Modi, E. Castorina and U. Seljak, *Halo bias in Lagrangian Space: Estimators and theoretical predictions*, *Mon. Not. Roy. Astron. Soc.* **472** (2017) 3959 [[arXiv:1612.01621](#)] [[INSPIRE](#)].
- [52] Z. Vlah, E. Castorina and M. White, *The Gaussian streaming model and convolution Lagrangian effective field theory*, *JCAP* **12** (2016) 007 [[arXiv:1609.02908](#)] [[INSPIRE](#)].
- [53] M. Schmittfull, M. Simonović, V. Assassi and M. Zaldarriaga, *Modeling Biased Tracers at the Field Level*, [arXiv:1811.10640](#) [[INSPIRE](#)].
- [54] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rept.* **733** (2018) 1 [[arXiv:1611.09787](#)] [[INSPIRE](#)].
- [55] B. Horn, L. Hui and X. Xiao, *Lagrangian space consistency relation for large scale structure*, *JCAP* **09** (2015) 068 [[arXiv:1502.06980](#)] [[INSPIRE](#)].
- [56] J. Carlson, B. Reid and M. White, *Convolution Lagrangian perturbation theory for biased tracers*, *Mon. Not. Roy. Astron. Soc.* **429** (2013) 1674 [[arXiv:1209.0780](#)] [[INSPIRE](#)].
- [57] D. Baumann, A. Nicolis, L. Senatore and M. Zaldarriaga, *Cosmological Non-Linearities as an Effective Fluid*, *JCAP* **07** (2012) 051 [[arXiv:1004.2488](#)] [[INSPIRE](#)].
- [58] J.J.M. Carrasco, M.P. Hertzberg and L. Senatore, *The Effective Field Theory of Cosmological Large Scale Structures*, *JHEP* **09** (2012) 082 [[arXiv:1206.2926](#)] [[INSPIRE](#)].
- [59] L. Senatore, *Bias in the Effective Field Theory of Large Scale Structures*, *JCAP* **11** (2015) 007 [[arXiv:1406.7843](#)] [[INSPIRE](#)].
- [60] Z. Vlah, M. White and A. Aviles, *A Lagrangian effective field theory*, *JCAP* **09** (2015) 014 [[arXiv:1506.05264](#)] [[INSPIRE](#)].
- [61] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *Phys. Rev. D* **98** (2018) 030001 [[INSPIRE](#)].
- [62] M. Tegmark, *Measuring cosmological parameters with galaxy surveys*, *Phys. Rev. Lett.* **79** (1997) 3806 [[astro-ph/9706198](#)] [[INSPIRE](#)].
- [63] A.N. Taylor, T.D. Kitching, D.J. Bacon and A.F. Heavens, *Probing dark energy with the shear-ratio geometric test*, *Mon. Not. Roy. Astron. Soc.* **374** (2007) 1377 [[astro-ph/0606416](#)] [[INSPIRE](#)].