The reconstructed power spectrum in the Zeldovich approximation

Shi-Fan Chen,[a] **Zvonimir Vlah**[b] **and Martin White**[a,c]

[a]Department of Physics, University of California,
  Berkeley, CA 94720, U.S.A.
[b]Theory Department, CERN,
  1 Esplanade des Particules, CH-1211 Genéve 23, Switzerland
[c]Department of Astronomy, University of California,
  Berkeley, CA 94720, U.S.A.

  E-mail: shifan_chen@berkeley.edu, zvonimir.vlah@cern.ch, mwhite@berkeley.edu

Abstract. Density-reconstruction sharpens the baryon acoustic oscillations signal by undoing some of the smoothing incurred by nonlinear structure formation. In this paper we present an analytical model for reconstruction based on the Zeldovich approximation, which for the first time includes a complete set of counterterms and bias terms up to quadratic order and can fit real and redshift-space data pre- and post-reconstruction data in both Fourier and configuration space over a wide range of scales. We compare our model to n-body data at $z = 0$ from the `DarkSky` simulation [1], finding sub-percent agreement in both real space and in the redshift-space power spectrum monopole out to $k = 0.4\,h\,\mathrm{Mpc}^{-1}$, and out to $k = 0.2\,h\,\mathrm{Mpc}^{-1}$ in the quadrupole, with comparable agreement in configuration space. We compare our model with several popular existing alternatives, updating existing theoretical results for exponential damping in wiggle/no-wiggle splits of the BAO signal and discuss the usually-ignored effect of higher bias contributions on the reconstructed signal. In the appendices, we re-derive the former within our formalism, present exploratory results on higher-order corrections due to nonlinearities inherent to reconstruction, and present numerical techniques with which to calculate the redshift-space power spectrum of biased tracers within the Zeldovich approximation.

Contents

1 Introduction

Density field reconstruction [2] is a means of improving the determination of the distance-redshift relation using baryon acoustic oscillations (BAO) [3]. The BAO method is a "standard ruler" test which seeks to measure the scale of a feature in the 2-point function whose physical size is known. Comparison with the observed size of this feature gives the angular diameter distance and Hubble parameter as a function of redshift. While the large size of the BAO feature (100 Mpc) makes it relatively immune to systematic effects, nonlinear evolution erases the oscillations on small scales, or broadens the peak in the correlation function, and reduces the accuracy with which the scale can be measured [4–8]. However much of the peak broadening comes from motions sourced by very long wavelength fluctuations [8] which are well measured by surveys aiming to measure BAO. This insight led ref. [2] to propose that

density-field reconstruction could be applied to regain much of the information lost to non-linearities. It has been used in all recent BAO surveys to improve their constraints (e.g. see ref. [9] and references therein).

BAO reconstruction has been studied both numerically [10–12] and analytically [2, 13–27]. Our work builds upon these analytic calculations. Where earlier work made simplifications aimed at highlighting important physical effects, neglected complications such as redshift-space distortions, applied heuristics or otherwise simplified the calculations for explanatory effect, we aim to produce a consistent dynamical theory which can be compared directly to upcoming observational data. Hence we generalize these calculations to also consider the power spectrum and extend the model to include the complete set of quadratic bias terms. To our knowledge this is the first dynamical model with a full bias scheme that can produce consistent real and redshift-space results in both Fourier and configuration space, allowing it to be used for consistent fitting of upcoming data.

There has been significant theoretical work on reconstruction since the first algorithm [2] was suggested. Most recently, a variety of iterative or alternative reconstruction approaches have been developed [15, 20, 22–25]. Though our calculations give some insights into these methods, for near-future experiments and for BAO scales these iterative methods do not lead to significant improvements and so we defer consideration of these more complex algorithms to future work.

The outline of this paper is as follows. Section 2 reviews the formalism of Lagrangian perturbation theory within which we work. Section 3 describes the reconstruction algorithm we seek to model, while section 4 gives our results in Fourier space, comparing to the configuration-space results where appropriate (section 5). We discuss alternative statistics in section 6. To assess the range of validity of our models we compare to N-body simulations in section 7. A comparison with earlier work is given in section 8 before we conclude in section 9. Some technical details are elaborated in the appendices.

2 Lagrangian perturbation theory

The Lagrangian framework [28–35] describes cosmological structure formation by tracking the displacements $\boldsymbol{\Psi}(\mathbf{q})$ of infinitesimal parcels of the matter fluid from their initial (Lagrangian) positions $\mathbf{q}$. In this picture the present day matter over- and underdensities are a result of the clustering of the displaced Eulerian positions $\mathbf{x}(\mathbf{q}, \tau) = \mathbf{q} + \boldsymbol{\Psi}(\mathbf{q}, \tau)$. The displacements follow the equation of motion $\boldsymbol{\Psi}''(\mathbf{q}) + \mathcal{H}\boldsymbol{\Psi}'(\mathbf{q}) = -\nabla_x \Phi(\mathbf{x})$, where $\Phi(\mathbf{x})$ is the gravitational potential which is in turn sourced by the clustered matter fluid via Poisson's equation $\nabla^2 \Phi(\mathbf{x}, \tau) = \frac{3}{2}\Omega_m(\tau)\mathcal{H}^2(\tau)\delta(\mathbf{x}, \tau)$ with τ the conformal time. This set of equations can be solved perturbatively in terms of the linear overdensity, δ_0, and the first order solution is given by $\boldsymbol{\Psi} = -D(\tau)\nabla\nabla^{-2}\delta_0$, where $D(\tau)$ is the linear growth factor [28].

The Lagrangian picture treats tracer bias and advection separately. Given a biased tracer, a, with initial overdensity $F^a(\mathbf{q}) = F^a\left[\partial^2\Phi(\mathbf{q}), \ldots\right]$, the time-evolved tracer overdensity at conformal time τ is given by number conservation as [34]

$$1 + \delta^a(\mathbf{x}, \tau) = \int d^3\mathbf{q} \; F^a(\mathbf{q}) \; \delta_D\left(\mathbf{x} - \mathbf{q} - \boldsymbol{\Psi}(\mathbf{q}, \tau)\right). \tag{2.1}$$

The cross power spectrum between two biased tracer populations a and b is then

$$P^{ab}(k) = \int d^3\mathbf{q} \; e^{i\mathbf{k}\cdot\mathbf{q}} \; \langle F^a(\mathbf{q}_2)F^b(\mathbf{q}_1) \; e^{i\mathbf{k}\cdot\Delta^{ab}}\rangle_{q=|q_2-q_1|}, \qquad \Delta^{ab} = \boldsymbol{\Psi}^a(\mathbf{q}_2) - \boldsymbol{\Psi}^b(\mathbf{q}_1), \tag{2.2}$$

where we have used that the integrated expectation value can only depend on $\mathbf{q} = \mathbf{q}_2 - \mathbf{q}_1$, due to the translation invariance of the underlying theory. The bias functionals, $F^{a,b}$, can be Taylor expanded in terms of bias coefficients

$$F^a(\mathbf{q}) = 1 + b_1^a \delta_0(\mathbf{q}) + \frac{1}{2} b_2^a \left(\delta_0(\mathbf{q})^2 - \langle \delta_0^2 \rangle \right) + b_s^a \left(s^2(\mathbf{q}) - \langle s^2 \rangle \right) + b_{\nabla^2}^a \nabla_q^2 \delta_0(\mathbf{q}) + \cdots , \quad (2.3)$$

where $s^2 = s_{ij} s_{ij}$ is the square of the shear field, i.e. the traceless part of $\partial\partial\Phi$. Following ref. [36], we also consider contributions from a "derivative bias" b_{∇^2}, i.e. corrections to the bias expansion at scales close to the halo radius R_h proportional to $\nabla^2 \delta_0$; such contributions will, however, be essentially degenerate with counterterms renormalizing nonlinearities in the Zeldovich power spectrum and we will therefore not enumerate them separately in the rest of this work unless otherwise stated.

In this work our focus will be on modelling reconstruction within the Zeldovich approximation [28, 37], which keeps only the linear order term in the dynamics of $\boldsymbol{\Psi}$ but re-sums the effects of the displacement to all orders in a Galilean-invariant manner (this is true for reconstruction also if we take it to mean that all displacements transform the same way). This is specifically accomplished by evaluating the exponential in equation (2.2) via the cumulant expansion, and evaluating the bias expansion using functional derivatives (see e.g. refs. [34–36]). Following standard techniques, as outlined in the references above, the resulting expression for the cross spectrum is

$$\begin{aligned}
P^{ab}(k) = \int d^3\mathbf{q} \; e^{i\mathbf{k}\cdot\mathbf{q}} \; e^{-\frac{1}{2}k_i k_j A_{ij}^{ab}} &\left[1 + \alpha_0 k^2 + i b_1^b \mathbf{k} \cdot U^a + i b_1^a \mathbf{k} \cdot U^a + b_1^a b_1^b \xi_L + \frac{1}{2} b_2^a b_2^b \xi_L^2 \right. \\
&- \frac{1}{2} k_i k_j \left(b_2^b U_i^a U_j^a + b_2^a U_i^b U_j^b + 2 b_1^a b_1^b U_i^a U_j^b \right) + i k_i (b_2^b b_1^a U_i^a + b_1^b b_2^a U_i^b) \, \xi_L \\
&- \frac{1}{2} k_i k_j (b_s^a \Upsilon_{ij}^b + b_s^a \Upsilon_{ij}^b) + i k_i (b_1^a b_s^b V_i^{ab} + b_1^b b_s^a V_i^{ba}) \\
&\left. + \frac{1}{2} (b_2^a b_s^b + b_2^b b_s^a) \chi^{12} + b_s^a b_s^b \zeta + \cdots \right]
\end{aligned} \quad (2.4)$$

where we have defined[1] the quadratic two point functions

$$A_{ij}^{ab} = \langle \Delta_i^{ab} \Delta_j^{ab} \rangle, \; U_i^b = \langle \Delta_i^{ab} \delta_0(\mathbf{q}_2) \rangle, \; \xi_L = \langle \delta_0(\mathbf{q}_2) \delta_0(\mathbf{q}_1) \rangle \quad (2.5)$$

and shear correlators

$$\zeta = \langle s^2(\mathbf{q}_2) s^2(\mathbf{q}_1) \rangle, \; \Upsilon_{ij}^b = \langle \Delta_i^{ab} \Delta_j^{ab} s^2(\mathbf{q}_2) \rangle, \; V_i^{ab} = \langle \Delta_i^{ab} \delta_0(\mathbf{q}_2) s^2(\mathbf{q}_1) \rangle, \; \chi^{12} = \langle \delta_0^2(\mathbf{q}_1) s^2(\mathbf{q}_2) \rangle. \quad (2.6)$$

Note that in the above calculations we have, without loss of generality, associated tracers a and b with Lagrangian positions $\mathbf{q}_2$ and $\mathbf{q}_1$, respectively. The quantities in equation (2.4) with a and b swapped can also be calculated by swapping the positions $\mathbf{q}_1 \leftrightarrow \mathbf{q}_2$. As an example, $U^b = -\langle \boldsymbol{\Psi}^b(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \rangle$ is the two-point function between the displacement of tracer b and the matter overdensity. The vector and tensor two point functions defined above can be decomposed via rotational symmetry into scalar components, e.g. $A_{ij} = X(q)\delta_{ij} + Y(q)\hat{q}_i\hat{q}_j$ and $U_i = U(q)\hat{q}_i$. Formulae for these functions, expressed as Hankel transforms of power spectra, are given in appendix A. Finally, we include the contribution $\alpha_0 k^2$ in the square

---

[1] These are generalizations of the similar auto-spectrum quantities defined in refs. [35–37].

brackets of equation (2.4) as the lowest-order counterterm renormalizing sensitivities to small-scale power in A_{ij} — in practice this simply modifies the matter contribution $P_{\text{Zel}}(k)$ ($\propto 1$ in the square brackets) to $(1 + \alpha_0 k^2) P_{\text{Zel}}(k)$ (see e.g. refs. [38, 39]). Each term in equation (2.4) can be evaluated as Hankel transforms (see e.g. ref. [36]) using the identities given at the end of [38], which we carry out using the `mcfit` package.[2]

The Lagrangian formalism allows a straightforward translation between real and redshift space via a mapping of the Lagrangian displacements. In particular, assuming the plane-parallel approximation[3] and working in the Zeldovich approximation, quantities in redshift space are given simply by substituting $\mathbf{\Psi}_i \to \mathbf{\Psi}_i^R = R_{ij} \mathbf{\Psi}_j$ [34]. Here $R_{ij} = \delta_{ij} + f \hat{n}_i \hat{n}_j$, where $\hat{n}$ denotes the line-of-sight direction and $f = d \ln D / d \ln a$ is the linear-theory growth rate. To lowest order, transforming into redshift space requires the inclusion of a second counterterm dependent on the line-of-sight angle $\nu = \hat{k} \cdot \hat{n}$. We can see this explicitly, for example, in the UV-sensitive zero-lag term in A_{ij}, which gains an angular dependence

$$k_i k_j \langle (\mathbf{\Psi}_i + \hat{n}_i \hat{n}_l \dot{\mathbf{\Psi}}_l)(\mathbf{\Psi}_j + \hat{n}_j \hat{n}_m \dot{\mathbf{\Psi}}_m) \rangle \equiv k^2 (X(0) + (2\dot{X}(0) + \ddot{X}(0))\nu^2), \tag{2.7}$$

where $\dot{\mathbf{\Psi}}$ is the velocity in Hubble units equal to $f\mathbf{\Psi}$ in the Zeldovich approximation;[4] roughly speaking, we need one angle-independent counterterm $\alpha_0 k^2$ to absorb the UV dependence of $X(0)$ and another $\alpha_2 k^2 \nu^2$ to absorb the UV dependence of the velocities. The complete set of counterterms in redshift space thus makes a contribution of the form $(\alpha_0 + \alpha_2 \nu^2) k^2 P_{ZA}(\mathbf{k})$; since $P_{ZA}(\mathbf{k})$ is equal to $(1 + f\nu^2)^2 P_L(k)$ to linear order, an equivalent viewpoint — which we will adopt in this work — is to have constant counterterms $\bar{\alpha}_0 k^2$ and $\bar{\alpha}_2 k^2$ for the monopole and quadrupole, respectively, where the barred counterterms are linear combinations of the unbarred quantities.

3 Reconstruction algorithm

In this section we describe two possible methods for reconstruction in redshift space, both built around the Zeldovich approximation. The standard procedure for reconstruction was developed in ref. [2] and involves displacing both observed galaxies and a spatially uniform distribution by a calculated shift field, χ, then taking the relative density contrast between the two sets of particles as the reconstructed density field. For a suitably chosen χ, this can reduce the effect of large scale (IR) bulk flows that "blur" the BAO feature. However there is no consensus in the community on the correct procedure for handling redshift-space distortions: the implementation in ref. [43] chose to multiply χ by $1 + f$ in the line-of-sight direction for δ_d but not for δ_s. This 'undoes' the supercluster infall effect [44] and reduces the $\ell > 0$ moments of the 2-point function on large scales. Ref. [16] suggested a symmetric treatment of δ_d and δ_s, which recovers linear theory on large scales. This is more natural from the point of view of perturbation theory and better behaved near the boundaries, but is less often implemented on data. A number of other choices were explored in ref. [18] but in this work we will restrict our attention to the two methods described above.

The reconstruction procedure consists of the following steps [2]:

1. Smooth the observed galaxy density field δ_g with a kernel $\mathcal{S}$ to filter out small scale (high k) modes, which are difficult to model. We use a Gaussian smoothing of scale

---

[2]https://github.com/eelregit/mcfit.

[3]This should be an excellent approximation on BAO scales [40], but if necessary the formalism can be modified to include "wide-angle" effects [41].

[4]See refs. [36, 42] for a more detailed exposition of the "dot notation."

R_s, specifically $\mathcal{S}(k) = \exp[-(kR_s)^2/2]$, though none of our analytic results will depend specifically on this choice. For galaxy surveys Gaussian smoothing has been universally adopted (though with different conventions for R_s) but in other contexts it may be advantageous to implement a Wiener filter instead (e.g. ref. [19]).

2. Compute the shift, χ, by dividing the smoothed galaxy density field by a bias factor b and linear RSD factor [44] and then take the inverse gradient. Assuming linear theory with scale-independent bias and supercluster infall holds on large scales, the calculated shift field should approximate the negative smoothed Zeldovich displacement. In a simulation with a periodic box, these first two steps can be implemented using FFTs as

$$\chi_{\mathbf{k}} = -\frac{i\mathbf{k}}{k^2}\mathcal{S}(k) \left(\frac{\delta_g(\mathbf{k})}{b + f\nu^2} \right) \approx -\mathcal{S}(k)\mathbf{\Psi}^{(1)}(\mathbf{k}) \tag{3.1}$$

where the bias factor is related to the Lagrangian first-order bias by $b = 1 + b_1$ and we have defined the line-of-sight angle $\nu = \hat{n} \cdot \hat{k}$. For non-periodic data the relevant differential equation can be solved by multigrid[5] or by linear algebra techniques [43] or iteratively using FFTs [45].

3. Move the galaxies by $\chi_d = \mathbf{R}\chi$ and compute the "displaced" density field, δ_d.

4. Shift an initially spatially uniform distribution of particles by

   - **Rec-Sym**: $\chi_s = \mathbf{R}\chi$, i.e. the same amount as the observed galaxies, or,
   - **Rec-Iso**: the un-redshifted $\chi_s = \chi$.

   to form the "shifted" density field, δ_s. Note that we have borrowed the nomenclature of ref. [18] for the latter, which "isotropizes" the reconstructed field on large scales. For the former we use "Rec-Sym" to indicate the symmetry of the treatment of δ_d and δ_s.

5. The reconstructed density field is defined as $\delta_r \equiv \delta_d - \delta_s$ with power spectrum $P_r(k) \propto \langle |\delta_r^2| \rangle$.

Throughout we shall assume that the fiducial cosmology and halo bias are properly known during reconstruction (see e.g. refs. [27, 46] for relaxation of this assumption), and take the approximation in eq. (3.1) to be exact. For further discussion of this point see refs. [17, 21]. The procedure in real space can be straightforwardly obtained by setting $f = 0$, in which case Rec-Sym and Rec-Iso become equivalent. Taking the limit $\mathcal{S} \to 0$ or $\chi \to 0$ returns the 'raw' spectrum, before reconstruction.

4 Reconstructed power spectrum

There has been significant earlier work on modeling density-field reconstruction within perturbation theory [2, 13–27]. In particular ref. [16] presented a calculation of the configuration-space two-point function (the correlation function) under the assumption of Zeldovich dynamics and that $\chi = -\mathcal{S}\mathbf{\Psi}$. In this paper we generalize that calculation to a more complete bias model (see section 5), including all terms allowed by symmetries up to quadratic order as well as a proper set of counterterms, and we show how to implement the model in Fourier

---

[5]https://github.com/martinjameswhite/recon_code.

space. We have explicitly checked that the Hankel transform of our Fourier-space expressions matches the direct configuration-space calculation to 1% in all terms, and we release code which makes consistent predictions for both statistics with a common set of parameters. To our knowledge this is the first calculation which provides self-consistent predictions in both spaces, uses a dynamical rather than a heuristic model, works in redshift space and has a full set of bias and counterterms.

Our focus in this section will be to model the reconstructed power spectrum using Lagrangian perturbation theory in both real and redshift space (the expression for the 'propagator' is given in appendix B for completeness). Following the algorithm outlined above, the reconstructed power spectrum in real space is given by $P_{\text{recon}} = P^{dd} + P^{ss} - 2P^{ds}$. Within the Lagrangian framework we can write the displaced density field as

$$
\begin{aligned}
1 + \delta_d(\mathbf{r}) &= \int d^3\mathbf{x} \; (1 + \delta(\mathbf{x})) \; \delta_D \left[\mathbf{r} - \mathbf{x} - \chi_d(\mathbf{x}) \right] \\
&= \int d^3\mathbf{x} \int d^3\mathbf{q} \; F(\mathbf{q}) \; \delta_D \left[\mathbf{x} - \mathbf{q} - \boldsymbol{\Psi}(\mathbf{q}) \right] \; \delta_D \left[\mathbf{r} - \mathbf{x} - \chi_d(\mathbf{x}) \right] \\
&= \int d^3\mathbf{q} \; F(\mathbf{q}) \; \delta_D \left[\mathbf{r} - \mathbf{q} - \boldsymbol{\Psi}(\mathbf{q}) - \chi_d(\mathbf{q} + \boldsymbol{\Psi}(\mathbf{q})) \right],
\end{aligned}
\tag{4.1}
$$

where we performed the $\mathbf{x}$ integral using the first δ-function to go from the second to third lines. Importantly while the fluid displacement, $\boldsymbol{\Psi}$, is evaluated at the Lagrangian position, $\mathbf{q}$, the shift field is evaluated at the shifted Eulerian position, $\mathbf{q} + \boldsymbol{\Psi}$. The above equalities hold both when the pre-reconstruction coordinate, $\mathbf{x}$, is in real or redshift space, with the implicit substitution of $\boldsymbol{\Psi} \to \mathbf{R}\boldsymbol{\Psi}$ in the latter case, as long as the appropriate shift field χ_d is chosen. The expression for the shifted density can be similarly derived or found by setting $\boldsymbol{\Psi}(\mathbf{q}) = 0$ and $\chi_d \to \chi_s$ in the above expression. In Fourier space this translates to

$$
\begin{aligned}
(2\pi)^3 \delta_D(\mathbf{k}) + \delta_d(\mathbf{k}) &= \int d^3\mathbf{q} \; e^{-i\mathbf{k}\cdot\mathbf{q}} \; F(\mathbf{q}) \; e^{-i\mathbf{k}\cdot\left[\boldsymbol{\Psi}(\mathbf{q}) + \chi_d(\mathbf{q} + \boldsymbol{\Psi}(\mathbf{q})) \right]} \\
(2\pi)^3 \delta_D(\mathbf{k}) + \delta_s(\mathbf{k}) &= \int d^3\mathbf{q} \; e^{-i\mathbf{k}\cdot\mathbf{q}} \; e^{-i\mathbf{k}\cdot\chi_s(\mathbf{q})}.
\end{aligned}
\tag{4.2}
$$

Below we will make the approximation $\chi(\mathbf{q} + \boldsymbol{\Psi}) \approx \chi(\mathbf{q})$. The nonlinearities from the Lagrangian-to-Eulerian mapping can be understood as a perturbation series in $\boldsymbol{\Psi}/R$, where R is the smoothing scale, and we explore their consequences in appendix E (see also refs. [17, 21] and the discussion in ref. [16]). Within this approximation we can treat the displaced and shifted field as tracers with displacements

$$
\boldsymbol{\Psi}^d = \boldsymbol{\Psi} + \chi_d, \quad \boldsymbol{\Psi}^s = \chi_s,
\tag{4.3}
$$

where the Zeldovich displacements should be understood as being in redshift space for the displaced field and in either redshift or real space for the shifted field depending on the method used. In this picture the "displaced" tracer has the same bias functional as the original galaxies ($F^d \equiv F^g$) while the "shifted" tracer is unbiased ($F^s \equiv 1$). A straightforward consequence of the reconstruction procedure is that, like that of any discrete tracer, the shift field autospectrum will contain an independent shot noise term $P^{ss}_{\text{SN}} = 1/n_s$, where n_s is the number density of the uniform random particles. The full shot noise contribution to the reconstructed spectrum is the sum of the galaxy and random particle shot noises.
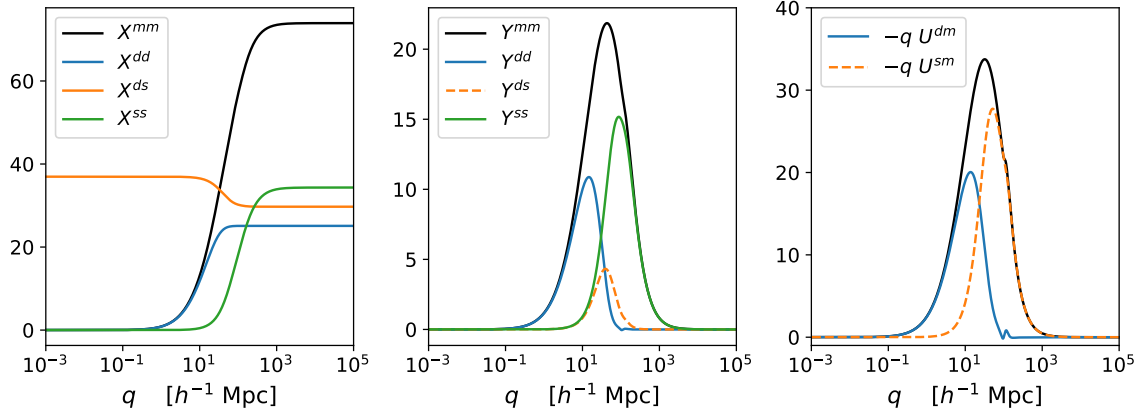
Figure 1. Lagrangian space two point functions used to compute reconstructed power spectra. Dashed quantities have been multiplied by an overall negative sign, and reflect that the shifted field is defined to be negatively correlated with the underlying matter field. Roughly speaking, the shifted and displaced correlators reproduce the general trend for the total matter correlators, shown in black, on large and small scales, respectively. An exception is X^{ds}, whose non-vanishing value on small scales reflect that the point values of $\mathbf{\Psi}^d$ and $\mathbf{\Psi}^s$ differ exactly by the Zeldovich displacement. Note also the small but visible features around $q = 100\,h^{-1}$ Mpc, i.e. the BAO scale.

4.1 Real space

In real space both the displaced and shifted fields are moved by the same, smoothed negative Zeldovich displacement, $\chi_d = \chi_s = -S \star \mathbf{\Psi}$, such that in Fourier space

$$\mathbf{\Psi}^d(\mathbf{k}) = [1 - \mathcal{S}(k)]\,\mathbf{\Psi}(\mathbf{k}), \quad \mathbf{\Psi}^s(\mathbf{k}) = -\mathcal{S}(k)\mathbf{\Psi}(\mathbf{k}). \tag{4.4}$$

and the auto- and cross-spectra can be calculated using equation (2.4) and the correlators in appendix A, using linear theory spectra

$$P_L^{dd}(k) = [1 - \mathcal{S}(k)]^2\,P_L(k), \quad P_L^{ds}(k) = -\mathcal{S}(k)\,[1 - \mathcal{S}(k)]\,P_L(k), \quad P_L^{ss}(k) = \mathcal{S}(k)^2 P_L(k) \tag{4.5}$$

as well as tracer-matter power spectra

$$P_L^{dm}(k) = [1 - \mathcal{S}(k)]\,P_L(k), \qquad P_L^{sm}(k) = -\mathcal{S}(k)P_L(k). \tag{4.6}$$

Note that the shifted field is negatively correlated with both the matter and displaced fields by-construction, since the random particles are displaced in the opposite direction of the (smoothed) Zeldovich displacement.

The Lagrangian space two-point correlation functions required to calculate the pre- and post-reconstruction power spectra, normalized to their present-day values, are shown in figure 1. For simplicity we have excluded the shear correlators and refer readers to appendix A for further details. The components X and Y describe correlation functions of two displacements, while the U's involve those with only one displacement, such that the former are Hankel transforms of the linear tracer-tracer spectra, while the latter involve the linear tracer-matter spectra. As expected, the Y's and U's for the displaced and shifted fields contain the behavior of the full matter contribution and small and large scales, respectively, and cross correlations between the shifted field and the displaced or matter fields is negative.

The $X(q)$ components however, especially the cross-correlation X^{ds}, display more subtle behavior. In particular, we have

$$A_{ij}^{ds}(\mathbf{q}) \stackrel{q \to 0}{=} \langle \mathbf{\Psi}_i^d \mathbf{\Psi}_j^d \rangle + \langle \mathbf{\Psi}_i^s \mathbf{\Psi}_j^s \rangle - 2\langle \mathbf{\Psi}_i^d \mathbf{\Psi}_j^s \rangle \equiv \Sigma^2 \delta_{ij}, \tag{4.7}$$

such that $X^{ds}(q) \to \Sigma^2$ as $q \to 0$. This is because, when evaluated at the same point, $\mathbf{\Psi}^d - \mathbf{\Psi}^s = \mathbf{\Psi}$, i.e. the difference between the displaced the shifted displacements is none other than the original Zeldovich displacement. This in turn implies that the cross spectrum is damped at small scales $\propto \exp[-k^2\Sigma^2/2]$ due to the nonzero displacement between the displaced and shifted fields. Similar behavior is seen in the evaluation of unequal-time correlation functions [47] and the baryon-cold dark matter cross-correlation [48, 49], though the physical mechanisms are of course different. At large scales, we similarly have

$$A_{ij}^{ds}(\mathbf{q}) \stackrel{q \to \infty}{=} \langle \mathbf{\Psi}_i^d \mathbf{\Psi}_j^d \rangle + \langle \mathbf{\Psi}_i^s \mathbf{\Psi}_j^s \rangle \equiv \left(\Sigma_{dd}^2 + \Sigma_{ss}^2 \right) \delta_{ij}, \tag{4.8}$$

such that X^{ds} asymptotes to the average of X^{dd} and X^{ss} at large separations. For completeness, we give explicit expressions for the displaced and shifted X^{ab} here:

$$X^{dd}(q) = \frac{2}{3} \int \frac{dk}{2\pi^2} \Big[1 - \Big(j_0(kq) + j_2(kq) \Big) \Big] \Big(1 - \mathcal{S}(k) \Big)^2 P_L(k)$$

$$X^{ds}(q) = \frac{2}{3} \int \frac{dk}{2\pi^2} \Big[\frac{1}{2} \Big((1 - \mathcal{S}(k))^2 + \mathcal{S}(k)^2 \Big) + \mathcal{S}(k) \Big(1 - \mathcal{S}(k) \Big) \Big(j_0(kq) + j_2(kq) \Big) \Big] P_L(k)$$

$$X^{ss}(q) = \frac{2}{3} \int \frac{dk}{2\pi^2} \Big[1 - \Big(j_0(kq) + j_2(kq) \Big) \Big] \mathcal{S}^2(k) P_L(k). \tag{4.9}$$

The corresponding expressions for Y^{ab} can be directly obtained by calculating -3 times the j_2 components. As we shall discuss further in section 8, the signs for the Bessel function coefficients in our expression for X^{ds} differ from those in ref. [26]. We note also, as has been emphasized before [13], each of the three contributions to P_{recon} has a different damping factor which can only be roughly approximated by a single Gaussian term.

The lowest-order bias terms in the reconstructed real-space power spectrum at $z = 0$ are shown figure 2. The pure-matter piece (i.e. the "1" in equation (2.4)) is the only term that includes contributions from all three combinations of d and s, while the b_1^2 piece consists of only the dd contribution. While each piece individually differs from the linear power spectrum, compared to the pre-reconstruction power spectrum, the Zeldovich approximation predicts that the post-reconstruction power spectrum largely recovers the oscillatory features in the linear spectrum, as seen in the lower panels of figure 2. We note that the structure of the breakdown into P^{dd}, P^{ds} and P^{ss} shown in figure 2 proceeds similarly in the higher-order bias contributions: bias terms like b_1^2, that are products of two bias parameters (e.g. $b_1 b_2$, $b_2 b_s$, ...), do not involve any displacements ($\mathbf{\Psi}$) and can thus only enter in the autospectrum of the biased "d" tracer P^{dd}, while those like b_1 that involve only one bias parameter (e.g. b_2, b_s) involve two-point functions with one displacement contracted and thus contribute to the cross spectrum P^{ds} as only one of the constituent tracers needs to be biased. The autospectrum P^{ss} does not contain any bias terms.

Figure 3 shows all the contributions to the reconstructed galaxy power spectrum and correlation function, up to the quadratic bias and shear terms. As seen in the top panel, the reduced damping in the lowest-order bias term "wiggles," barely visible in log-log plots of the reconstructed power spectrum, translate to significantly sharper and less shifted BAO features
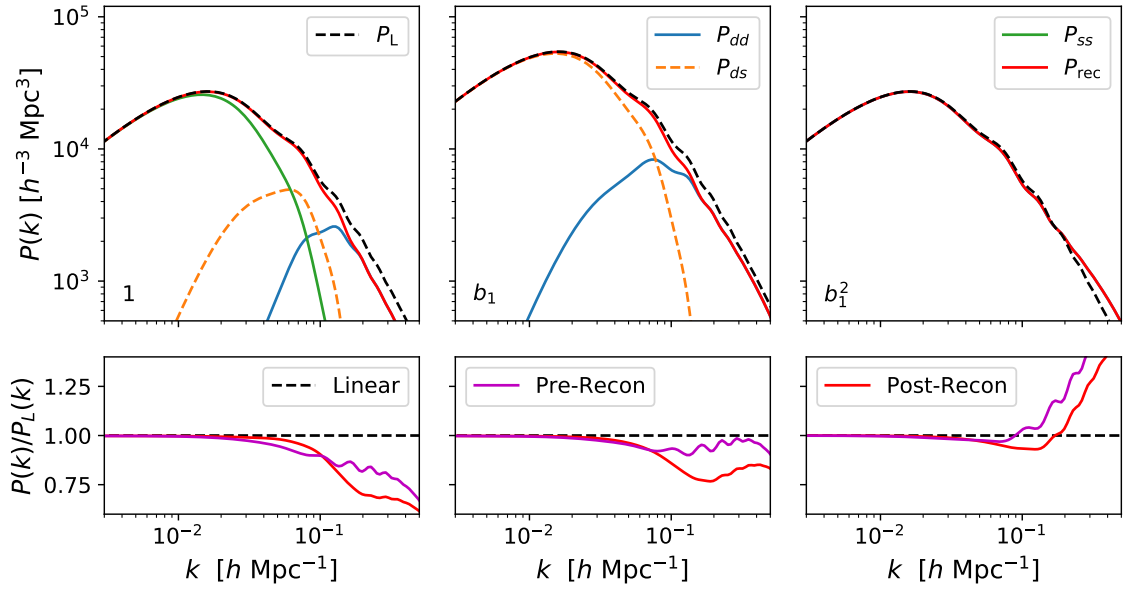
Figure 2. (Top) Real-space power spectra contributions, displaced-displaced, displaced-shifted and shifted-shifted, for the lowest order bias terms 1, b_1, b_1^2, and their sum, compared to linear theory at $z = 0$. The pure matter piece is the only term that receives contributions from all three combinations of d and s, and the b_1^2 term consists only of the dd contribution. All three bias terms tend to linear theory on large scales but exhibit somewhat different broadband behavior at high k. (Bottom) The ratio of the above bias terms with the linear theory power spectrum, compared with the pre-reconstruction Zeldovich power spectrum. While both the pre- and post-reconstruction Zeldovich spectra differ with the linear spectrum in the broadband at small scales, the Zeldovich approximation predicts that the oscillatory features in the reconstructed spectrum are almost identical to those in the linear spectrum, such that the wiggles are almost completely normalized out for the reconstructed spectrum.

(right column). In the quadratic bias contributions (middle panels), reconstruction can be seen to dampen the amplitude of the BAO feature in the b_2 and $b_1 b_2$ contributions, which "wiggle" in Fourier space, while leaving the spectrally smooth b_2^2 contribution essentially intact. Since the BAO feature in the quadratic bias contributions will tend to smear and shift the observed BAO peak from its linear theory position, reconstruction serves to remove these confounding nonlinearities as expected. The shear terms have less pronounced (i.e. smoother) features at the BAO scale — which we will show in section 8 as being essentially in-phase with the linear theory oscillations — that are less affected by reconstruction.

Finally, as noted in the discussion below equation (2.4), the exponentiated A_{ij} in Zeldovich power spectra are assumed to be long wavelength, IR modes which can be resummed while contributions from the rest of the shorter modes are perturbatively expanded. These expanded modes thus carry also a UV (small-scale) sensitivity that should be renormalized by adding the appropriate counterterms, quadratic in wavenumber and proportional to the Zeldovich power spectrum: $\alpha_{ab} k^2 P_{\text{Zel}}^{ab}$. In principle, we expect such counterterms in all three pieces of our reconstructed power spectrum, however given that the P^{ss} consists of mostly IR modes we expect its counterterm contribution to be suppressed relative to similar terms in P^{dd} and P^{ds}, though it could still be non-vanishing due to contributions we neglected in approximating equation (3.1). While the counterterms α_{dd} and α_{ds} are highly nondegenerate
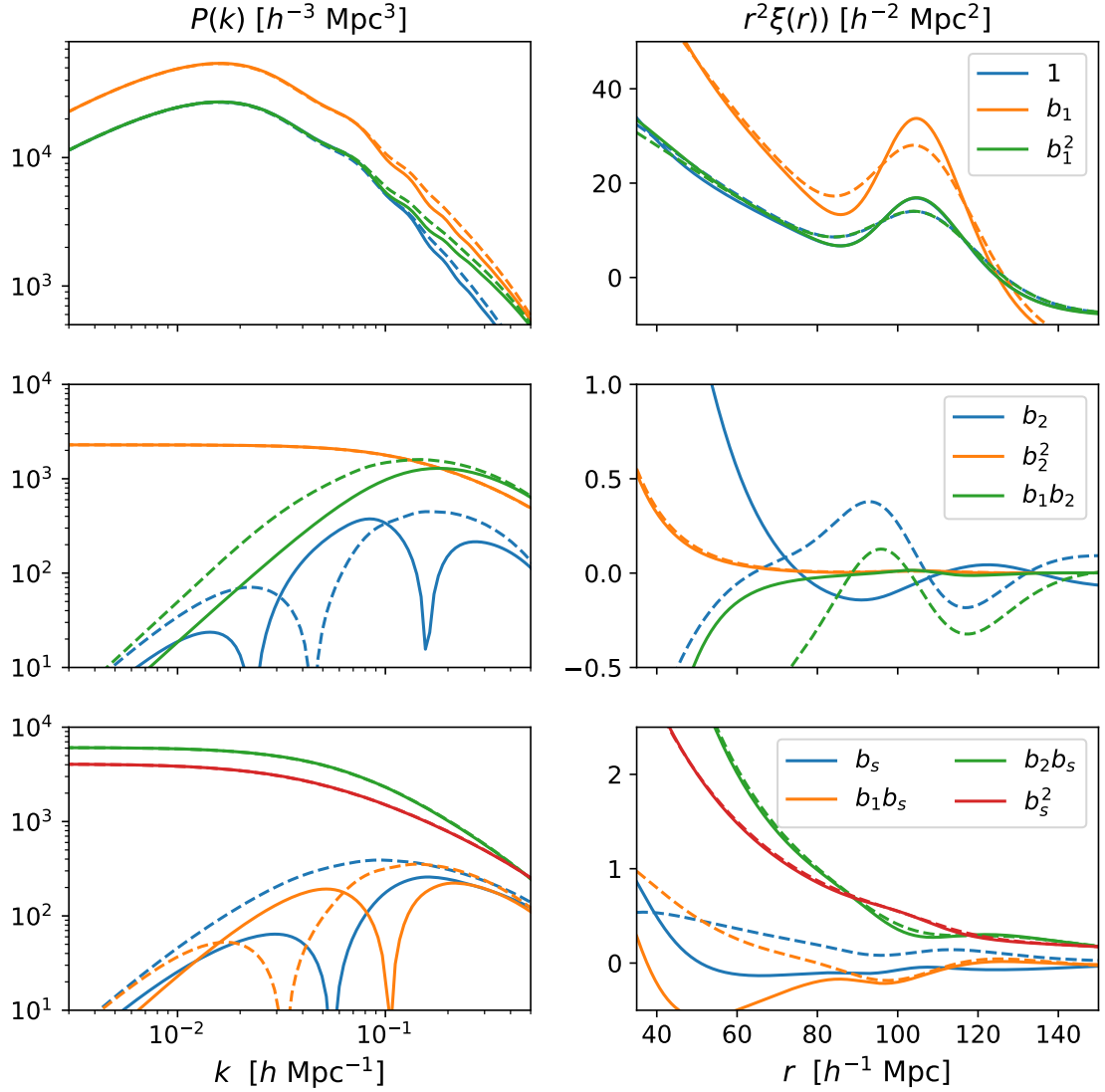
Figure 3. Contributions to the pre- and post-reconstruction (dashed and solid) power spectra and correlations functions (left and right columns) in real space from linear through quadratic bias terms at $z = 0$. Note that the matter (blue) and b_1^2 (green) curves in the top right panel are essentially degenerate, especially at the large scales shown.

due to the different supports of P^{dd} and P^{ds} (see figure 2) in k-space, in this work we will also explore modelling the reconstructed power spectrum using only one counterterm, $\propto k^2 P_{\mathrm{Zel}}$, for both P^{dd} and P^{ds} contributions, since such a contribution would also be degenerate with any potential derivative biases (see e.g. ref. [36]). We will return to the difference between these options in section 7.

4.2 Redshift space

In this section we develop analytic expressions for the redshift-space reconstructed power spectrum in both **Rec-Sym** and **Rec-Iso**. Methods recently developed in ref. [50] allow us to

extend the LPT redshift-space power spectrum calculation to include bias and the specifics of reconstruction, which we summarise here and present in detail in appendix C. As we will show shortly, **Rec-Sym** and **Rec-Iso** are not equivalent even to linear order. Specifically, we have

$$P_{\text{sym}}(\mathbf{k}) = (b + f\nu^2)^2 P_{\text{L}}(k) + \mathcal{O}(P_{\text{L}}^2) \tag{4.10}$$

$$P_{\text{iso}}(\mathbf{k}) = \left[(b + f\nu^2)(1 - \mathcal{S}) + \mathcal{S}\right]^2 P_{\text{L}}(k) + \mathcal{O}(P_{\text{L}}^2), \tag{4.11}$$

i.e. while **Rec-Sym** restores supercluster infall at linear order, **Rec-Iso** removes redshift-space distortions at large scales while keeping them at small scales. As we will see, this produces a smooth modulation in the broadband power nondegenerate with the BAO wiggles.

Since both the smoothed and displaced fields are uniformly multiplied by R_{ij} in **Rec-Sym**, it is straightforward to calculate the reconstructed power spectrum using equation (2.4) with

$$\boldsymbol{\Psi}^d(\mathbf{k}) = [1 - \mathcal{S}(k)]\,\mathbf{R}\boldsymbol{\Psi}(\mathbf{k})\,, \qquad \boldsymbol{\Psi}^s(\mathbf{k}) = -\mathcal{S}(k)\mathbf{R}\boldsymbol{\Psi}(\mathbf{k})\,. \tag{4.12}$$

In particular the angular structure of the $\mathbf{q}$ integral follows as in the calculation of the galaxy power spectrum without further modifications, and the set of bias terms in the dd, ds and ss spectra are identical to the real space case. The reconstructed power spectrum can then be calculated as one would the unreconstructed redshift space power spectrum. We develop the formalism to do the latter in appendix C.3 and comment on the changes required to go to the reconstructed case therein.

The cross spectrum in **Rec-Iso** is slightly different since only the displaced field is multiplied by the redshift space transformation, R_{ij}. The displaced and shift fields in this case are thus instead

$$\boldsymbol{\Psi}^d(\mathbf{k}) = [1 - \mathcal{S}(k)]\,\boldsymbol{\Psi}^R(\mathbf{k}) = [1 - \mathcal{S}(k)]\,\mathbf{R}\boldsymbol{\Psi}(\mathbf{k})\,, \qquad \boldsymbol{\Psi}^s(\mathbf{k}) = -\mathcal{S}(k)\boldsymbol{\Psi}(\mathbf{k}). \tag{4.13}$$

Since the displaced and shift moves thus lie in redshift and real space, respectively, their auto spectra can also respectively be calculated as in **Rec-Sym** and real space reconstruction; however, the cross spectrum is only "half transformed" into redshift space and thus requires special attention. The exponentiated two-points displacements are given by

$$A_{ij}^{ds,\text{Iso}} = \langle \boldsymbol{\Psi}_i^d \boldsymbol{\Psi}_j^d \rangle + \langle \boldsymbol{\Psi}_i^s \boldsymbol{\Psi}_j^s \rangle - 2\langle \boldsymbol{\Psi}_i^d(\mathbf{q}_2)\boldsymbol{\Psi}_j^s(\mathbf{q}_1) \rangle \tag{4.14}$$
$$= R_{in}R_{jm}\langle \boldsymbol{\Psi}_n^d \boldsymbol{\Psi}_m^d \rangle_{\text{RealSpace}} + \langle \boldsymbol{\Psi}_i^s \boldsymbol{\Psi}_j^s \rangle_{\text{RealSpace}} - 2R_{in}\langle \boldsymbol{\Psi}_n^d(\mathbf{q}_2)\boldsymbol{\Psi}_j^s(\mathbf{q}_1) \rangle_{\text{RealSpace}},$$

such that the zero-lag piece due to the displaced-displaced correlation is fully transformed into redshift space, the zero-lag piece due to the shifted-shifted correlation is untransformed, and the coordinate dependent displaced-shifted correlation is "half transformed." In particular, defining as usual $\mathbf{q} = q\,\hat{q}$ and $\hat{k} \cdot \hat{q} = \mu$, the last piece is

$$k_i k_j \langle \boldsymbol{\Psi}_i^d(\mathbf{q}_2)\boldsymbol{\Psi}_j^s(\mathbf{q}_1) \rangle = k_i k_j (\delta_{ik} + f\hat{n}_i\hat{n}_k)(\tilde{X}^{ds}\delta_{kj} + \tilde{Y}^{ds}\hat{q}_k\hat{q}_j)$$
$$= k^2(1 + f\nu^2)\tilde{X}^{ds} + k^2(\mu^2 + f\mu\nu(\hat{q} \cdot \hat{n}))\tilde{Y}^{ds}, \tag{4.15}$$

where we have defined the tilded quantities without the usual zero lag piece[6]

$$\langle \boldsymbol{\Psi}^d(\mathbf{q}_2)\boldsymbol{\Psi}^s(\mathbf{q}_1) \rangle_{\text{Real Space}} = \tilde{X}^{ds}(q)\,\delta_{ij} + \tilde{Y}^{ds}(q)\,\hat{q}_i\hat{q}_j.$$

---

[6]For notational simplificity, the functions X and Y are always defined in real space.

Note that $2\tilde{Y} = -Y$ since Y does not possess a zero-lag piece. The azimuthal-angle dependence in $\hat{q} \cdot \hat{n}$ will require us to do the integral (appendix C.2)

$$\int \frac{d\phi}{2\pi} e^{A\mu\sqrt{1-\mu^2}\cos\phi} = \sum_{\ell=0}^{\infty} H_\ell^{(0)}(A) \ (A\mu^2)^\ell,$$

where we have defined

$$H_\ell^{(0)}(A) = \sum_{m=0}^{\ell} \frac{(-1)^{\ell-m} A^{2m-\ell} \Gamma(m+\frac{1}{2})}{\sqrt{\pi}\Gamma(2m+1)\Gamma(2m-\ell+1)\Gamma(\ell-m+1)}.$$

Note that the Γ functions in the denominator will kill any terms in the sum for which $2m-\ell$ is negative, such that the sum really only contains $\ell/2$ terms and is always convergent in A. The full cross spectrum is then given by

$$P^{(ds)}(\mathbf{k}) = e^{-\frac{1}{2}k^2(\alpha_0\Sigma^{(dd)2}+\Sigma^{(ss)2})}\int d^3\mathbf{q} \ e^{ikq\mu+k^2(1+f\nu^2)(\tilde{X}^{(ds)}+\mu^2\tilde{Y}^{(ds)})}$$

$$\times \sum_{\ell=0}^{\infty} H_\ell^{(0)}(A)A^\ell\mu^{2\ell}\left(1 + ib_1 k\mu U^{(d)}(q) - \frac{1}{2}b_2 k^2\mu^2 U^{(d)}(q)^2 + \dots\right) \qquad (4.16)$$

where $A = k^2 f\nu\sqrt{1-\nu^2}\tilde{Y}^{(ds)}$ and we have defined $\Sigma^2 = \tilde{X}(0)$ and $\alpha_0 = 1 + f(f+2)\nu^2$. The remaining integrals can then be performed using the usual tricks for powers of μ using the series described in appendix C.1, and are explicitly given at the end of appendix C.2.

Figures 4 and 5 show the various bias contributions to the reconstructed redshift space power spectrum monopoles and quadrupoles within **Rec-Sym** and **Rec-Iso**, respectively. A significant difference between the two methods can be seen by comparing the matter (i.e. "1") pieces in the top panels of the two figures. While all three linear bias contributions to the reconstructed power spectrum monopole ($\propto 1$, b_1, b_1^2) approach the Zeldovich monopole in the large scale limit in **Rec-Sym**, the matter contribution to the **Rec-Iso** monopole instead approaches the b_1^2 contribution, which does not receive redshift space distortions in the linear theory limit. This is because the power spectrum at the largest scales is dominated by the autospectrum of the un-redshifted shift field, P^{ss}. While the matter and b_1 contributions to the reconstructed quadrupole approach linear theory in **Rec-Sym**, they vanish on large scales in **Rec-Iso**. On the other hand, the majority of the higher bias contributions (excluding b_2 and b_s) are sourced only by P^{dd} and are thus identical between the two methods, as can be seen by comparing the lower two rows of figures 4 and 5. This corresponds to our intuition that redshift-space distortions are less prominent for highly biased tracers, and that the differences between Rec-Iso and Rec-Sym disappear if we remove RSD. In addition, the contributions enumerated above are supplemented by counterterms ($\alpha_{dd}^\ell, \alpha_{ds}^\ell, \alpha_{ss}^\ell$), where we need a separate counterterm for each pair and multipole as discussed below equation (2.7), though as in the real space case we also explore the possibility of only fitting one counterterm each for the net reconstructed monopole and quadrupole.

5 Reconstructed correlation function

The configuration space two-point function (the correlation function) can be obtained from our Fourier-space results by Hankel transform. It is also possible to rewrite the q-dependent
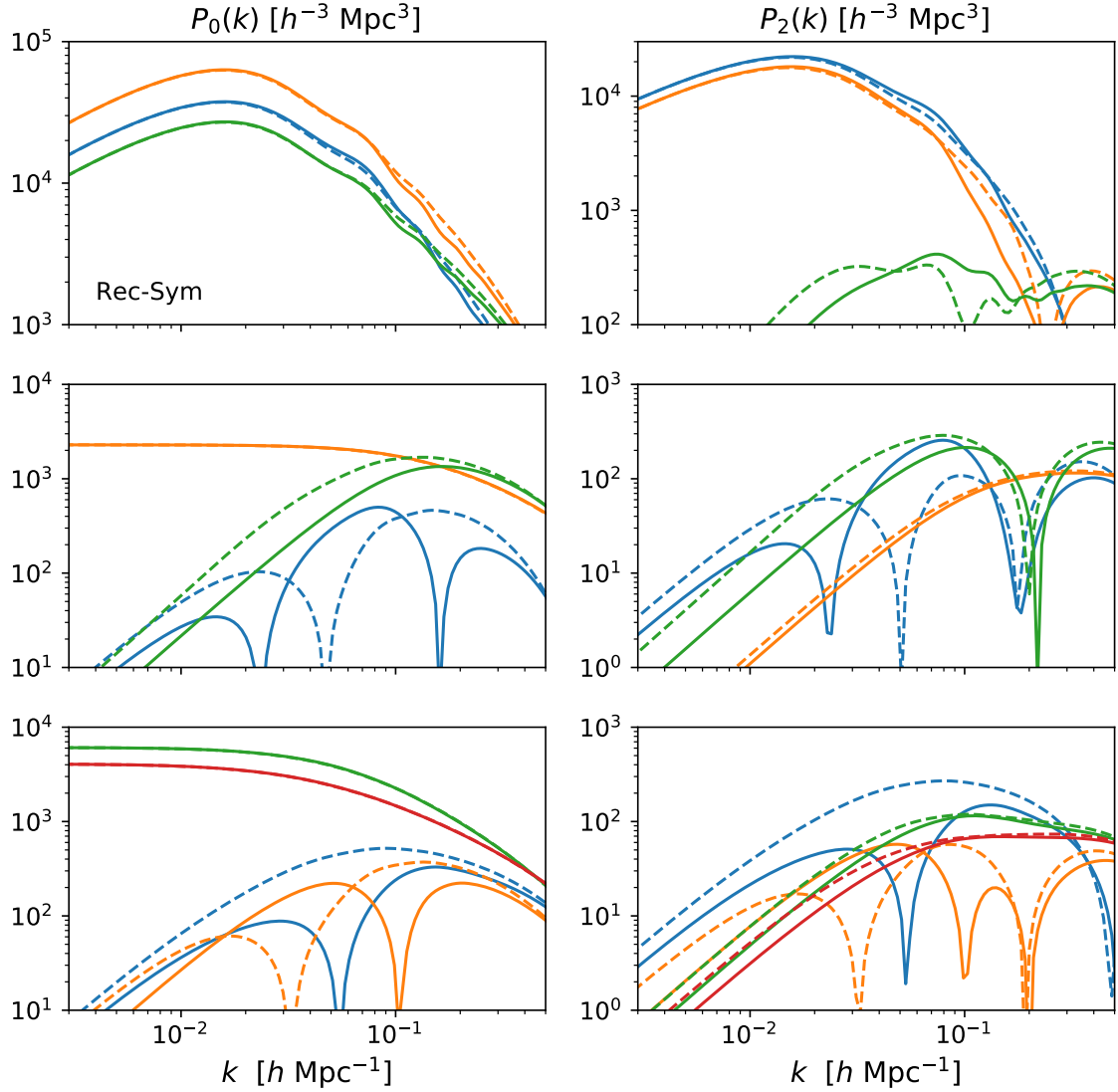
Figure 4. Bias contributions to the pre- and post-reconstruction (dashed and solid) $z = 0$ redshift space power spectra monopole and quadrupoles in the **Rec-Sym** scheme. The color scheme and line styles follow those in figure 3. The lowest-order contributions to the reconstructed monopole and quadrupole due to the linear bias b_1 tend to the Kaiser approximation at large scales. Note the different y-axis ranges on different panels.

integrals to compute $\xi(r, \nu_r)$ directly, where $\nu_r = \hat{n} \cdot \hat{r}$. Here we reprise the calculation of ref. [16], extending it to include the additional bias terms and commenting explicitly on several numerical issues which arise. We have checked that our Fourier and configuration space results agree numerically to significantly sub-percent levels in both real and redshift space for both **Rec-Sym** and **Rec-Iso**.
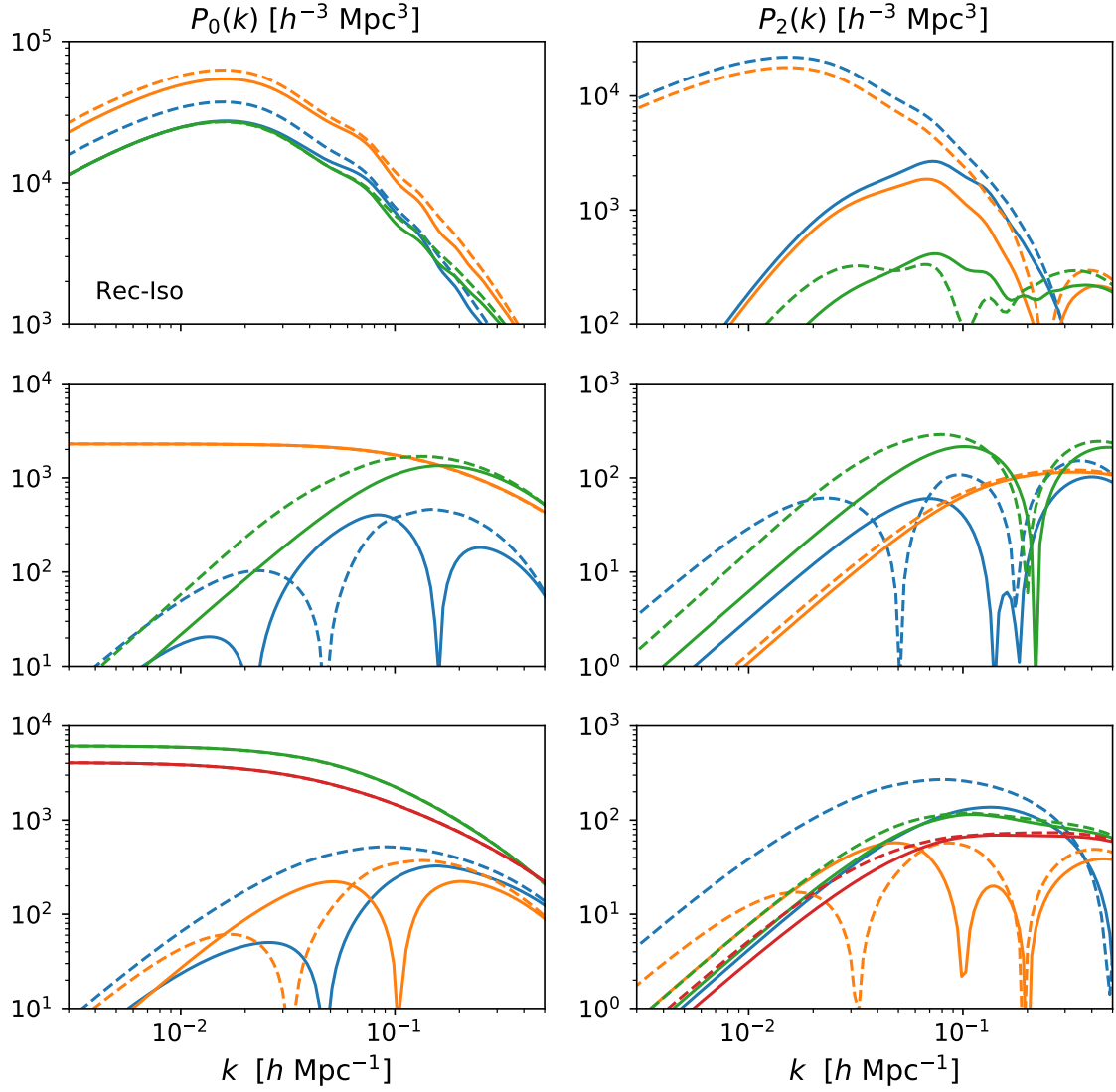
Figure 5. Same as figure 4, but for **Rec-Iso** at $z = 0$. Unlike in **Rec-Sym**, the linear bias contributions to the monopole and quadrupole do not tend to the Kaiser limit on large scales but to the real space linear power spectrum, as evidenced by reduced power in the monopole compared to the pre-reconstruction Zeldovich power spectrum, and contributions to the quadrupole vanishing on large scales. However, many of the higher bias contributions are identical to those in **Rec-Sym** (figure 4).

The general formula for the cross spectrum of two tracers a and b given in equation (2.4) can be Fourier transformed to give [36–38, 42]

$$1 + \xi^{ab}(r) = \int \frac{d^3q}{(2\pi)^{3/2}|A^{ab}|^{1/2}} e^{-(1/2)(q_i - r_i)(A_{ab}^{-1})_{ij}(q_j - r_j)}$$

$$\times \left\{ 1 - (b_1^b U_i^a + b_1^a U_i^b) g_i + b_1^a b_1^b \xi_L + \frac{1}{2} b_2^a b_2^b \xi_L^2 \right.$$

– 14 –

$$-\frac{1}{2}\left[b_2^b U_i^a U_j^a + b_2^a U_i^b U_j^b + 2b_1^a b_1^b U_i^a U_j^b\right] G_{ij} - \left[b_1^a b_2^b U_i^a + b_1^b b_2^a U_i^b\right]\xi_L g_i$$

$$-\frac{1}{2}\left[b_s^a \Upsilon_{ij}^b + b_s^b \Upsilon_{ij}^a\right] G_{ij} - \left[b_1^a b_s^b V_i^{ab} + b_1^b b_s^a V_i^{ba}\right] g_i$$

$$+\frac{1}{2}(b_2^a b_s^b + b_2^b b_s^a)\chi^{12} + b_s^a b_s^b \zeta + \alpha_{ab}\mathrm{tr}G + \dots \Bigg\},\tag{5.1}$$

where we have defined

$$g_i = (A_{ab}^{-1})_{ij}(q_j - r_j),\ G_{ij} = (A_{ab}^{-1})_{ij} - g_i g_j \tag{5.2}$$

and placed the superscript ab in A^{ab} into the subscript for notational convenience. In the configuration space calculation above, the Lagrangian two-point functions (e.g. A_{ij}, U_i, Υ_{ij}) can be computed using the formulae provided in appendix A. The above formula can be translated into redshift space by multiplying the Lagrangian two-point functions with vector indices by the appropriate factors of $R_{ij} = \delta_{ij} + f\hat{n}_i\hat{n}_j$. Taking the line-of-sight to be in the z direction without any loss of generality, this is equivalent to multiplying by the matrix $\mathrm{diag}(1,1,1+f)$. When calculating the un-reconstructed redshift space correlation function, this multiplication is equivalent to multiplying each z component index of vector and tensor quantities (e.g. U_z^a or A_{yz}^{ab}) by $1+f$, and dividing the corresponding components in the matrix inverse, A_{ab}^{-1}, by the same factor. The redshift-space counterterm $\alpha_2 k^2 \nu^2$ can be included in the correlation function by adding $\alpha_2 \hat{n}_i \hat{n}_j G_{ij}$, which similarly is equivalent to $\alpha_2 G_{zz}$ when picking z as the line-of-sight direction.

The reconstructed correlation function in real and redshift space can be calculated using equation (5.1) by defining "displaced" and "shifted" tracers as in the case of the power spectrum (sections 4.1 and 4.2) and calculating the combined quantity $\xi_{\mathrm{recon}} = \xi_{dd} + \xi_{ss} - 2\xi_{ds}$. For reconstruction using **Rec-Sym**, the same shortcuts of multiplying by factors of $1+f$ in lieu of matrix multiplication and inversion apply, since all vector and tensor quantities undergo the same transformation by R_{ij}. The calculation for **Rec-Iso** is more complicated. As was the case in Fourier space, the displaced-displaced and shifted-shifted auto-correlation functions are equal to their counterparts in **Rec-Sym** and real-space reconstruction, while the displaced-shifted cross-correlation function contains a mix of real and redshift space factors. In particular, from equation (4.14) we see that the two zero-lag pieces and one q-dependent piece of A_{ij}^{ds} in **Rec-Iso** are independently transformed by different numbers of R_{ij}'s. For this reason, when calculating the correlation function in **Rec-Iso**, the matrix inverse of A^{ds} in redshift space cannot be simply obtained by dividing the real space inverse by factors of $1+f$; rather, the uninverted matrix must be redshifted piece by piece as in equation (4.14) and then inverted numerically (we use Cholesky decomposition).

6 Other statistics

While the correlation function and power spectrum are the most frequently considered 2-point functions, there are other variants that have some advantages. Since these can all be written in terms of the correlation function or power spectrum, our model provides a consistent prediction for them as well. Of particular interest for BAO is the ω_ℓ statistic of ref. [51], which combines the scale-localization of the Fourier-space methods with the compactness and easy treatment of masks of the configuration-space methods.

In principle ω_ℓ can be calculated from either the configuration-space or Fourier-space expressions given above, but we have found it more convenient to start from the Fourier

$\lg M$	Redshift	$\bar{n}$	b
$12.0 - 12.5$	0.0	3.45	0.87
$12.5 - 13.0$	0.0	1.18	1.05
$13.0 - 13.5$	0.0	0.38	1.30

Table 1. Number densities and bias values for the halo samples we use. Halo masses are $\log_{10}$ of the mass in $h^{-1}M_\odot$, number densities are times $10^{-3}\,h^3\mathrm{Mpc}^{-3}$.

expressions. Since these are computed using FFTlog they naturally cover a very wide range of k, making the transforms to ω_ℓ easy to implement. For example

$$\omega_0(r_s) = \int \frac{k^2\,dk}{2\pi^2} P_0(k)\widetilde{W}_0(k\,r_s) \tag{6.1}$$

with $\widetilde{W}_0$ given in ref. [51] (see their figure 1 and appendix A). At large scales $\widetilde{W}_0 \propto k^2$ while at small scales $\widetilde{W}_0 \propto k^{-4}$. Our formalism naturally provides predictions for ω_ℓ using the same set of bias and nuisance parameters as for ξ_ℓ and P_ℓ.

7 Comparison to N-body

To look at the domain of validity of our analytic results we compare to the `DarkSky` N-body simulation suite,[7] specifically simulation `ds14_a` [1]. This simulation used the `2HOT` code [52] to evolve 10240^3 particles in an $(8\,h^{-1}\mathrm{Gpc})^3$ volume to model the growth of structure in a ΛCDM cosmology with $\Omega_M = 1 - \Omega_\Lambda = 0.295$, $h = 0.688$, $n_s = 0.968$ and $\sigma_8 = 0.835$. Initial conditions were generated from a glass using 2nd order Lagrangian perturbation theory at $z = 93$. Halos were found using the `Rockstar` code [53]. We extracted the positions, velocities and masses of halos more massive than $M_{200b} = 10^{12}\,h^{-1}M_\odot$ from the publicly available data at $z = 0$ (data at higher z, which would have been a more relevant comparison, were not available). We computed the halo correlation functions and power spectra, in real and redshift space. For the redshift-space quantities we assumed the plane-parallel approximation with the line-of-sight being the z-axis. We also obtained the linear theory power spectrum used to generate the initial conditions, which we take as the input to our model.

We implemented the algorithm described in section 3 using the periodicity of the box and FFTs to perform the smoothing and computation of the shifts. As for the power spectrum and correlation function, the plane-parallel approximation with line-of-sight the z-axis was assumed for the redshift-space quantities. The code takes as input an assumed large-scale bias, b, and growth parameter, f, in addition to a Gaussian smoothing length, R. We used the b obtained from the ratio of the linear theory and real-space halo power spectra at low k (see table 1), and $f \simeq 0.508$ appropriate to the simulation cosmology at $z = 0$, and note in passing that the goodness-of-fit of our results did not seem to be greatly improved by substituting the linear bias thus obtained with the value of $1 + b_1$ obtained by fitting the pre-reconstruction data with our model up to quasi-nonlinear scales.

We computed the reconstructed field in both real and redshift space. In each case the shifted and displaced positions were computed using a 2048^3 FFT, which resolves the (Gaussian) smoothing length by $2.5 - 5$ grid cells for $R \simeq 10 - 20\,h^{-1}\mathrm{Mpc}$. We used as many

---

[7]http://darksky.slac.stanford.edu.

"random" positions as halos in each case, for simplicity, and computed the power spectra and correlation functions for dd, ds and ss assuming periodic boundary conditions. The reconstructed power spectrum or correlation function can then be computed as $dd - 2ds + ss$, and we can look at each of the contributions separately. Note that our choice of equal numbers of randoms and data points means the shot noise on the reconstructed power spectrum is twice that of the pre-reconstructed field.

We compare the N-body results to our model with b_1 and b_2 and include the minimal set of counterterms as described in the preceding sections (one and three pre- and post-reconstruction, respectively, in real space) as well as a constant shot noise component fit to the data. For brevity our discussion will focus on halos with masses between $12.5 < \log_{10}(M/h^{-1}M_\odot) < 13.0$., though we obtained qualitatively similar results in the lower and higher mass bin as well, and show fits of the reconstructed redshift space power spectrum in the latter at the end of this section. We have checked that including nonzero shear bias b_s does not visibly improve the goodness-of-fit. The top-left pair of panels of figure 6 compares the unreconstructed real-space power spectrum in our model with $(b_1, b_2) = (0.02, -0.8)$ with that in DarkSky. The quadratic bias, b_2, accounts for a non-negligible fraction of the total power at essentially all scales and significantly reduces the constant shot noise term in the fit. We find that with a counterterm $\alpha \approx 11 \, h^{-2}\mathrm{Mpc}^2$ our model agrees with the data at the percent level out to $k \simeq 0.4 \, h \, \mathrm{Mpc}^{-1}$. The counterterm accounts for roughly a 10% correction at $k = 0.1 \, h \, \mathrm{Mpc}^{-1}$, and it is worth noting that even in its absence our model accurately captures the BAO features in the power spectrum, as evidenced by the lack of oscillatory features in the fit residuals.

The remaining panels of figure 6 show the fit for the reconstructed power spectra at three smoothing scales $R = 10, 15, 20 \, h^{-1}\mathrm{Mpc}$. We have tested whether the data could be reproduced using only one counterterm, α (shown in orange), or equivalently from one derivative bias b_{∇^2}, and find that such a choice dramatically reduces the range-of-validity of the model compared to three counter terms. While we adopted a rather conservative approach in fitting these data, prioritizing the accuracy of our predictions at low k rather than producing reasonable-looking fits to smaller scales, our model with three counterterms $(\alpha_{dd}, \alpha_{ds}, \alpha_{ss})$ nonetheless reproduces both the broadband power and oscillatory features of the reconstructed power spectrum out to $k = 0.2 \, h \, \mathrm{Mpc}^{-1}$ at the percent level for $R = 15$ and $20 h^{-1}$ Mpc. That each of the three constituent spectra in $P^{\mathrm{recon}} = P^{dd} + P^{ss} - 2P^{ds}$ has distinct short-wavelength behavior and k-space supports underlies the success of our model with three counterterms — each of which has highly nondegenerate scale dependence — versus the one-counterterm alternative. We have found that setting $\alpha_{ss} = 0$ does not qualitatively alter the degree to which our model fits the data; we have made this choice in all of our fits below, but note that as $\alpha_{ss}k^2$ vanishes quadratically towards low k, the data are also naturally rather insensitive to it. Indeed, since nonlinear corrections are typically of order $k^2\Sigma^2$, and the smoothing scale is chosen such that $(\Sigma/R)^2 \ll 1$, the insensitivity of P^{ss} to these corrections follows almost by construction. However, a bump-like feature around $k = 0.1 \, h \, \mathrm{Mpc}^{-1}$ is persistent across all the fits, peaking at less than half a percent when $R = 20 \, h^{-1}\mathrm{Mpc}$ and growing to a full percent at $R = 10 \, h^{-1}\mathrm{Mpc}$. The appearance of such a feature, growing towards smaller smoothing scales, is consistent with our neglect of nonlinear corrections to the smoothed displacements, which should increase towards smaller smoothing scales roughly as Ψ/R; we discuss one such nonlinearity in appendix E. For sufficiently small smoothing scales, even the assumption that the smoothing of the BAO feature can be essentially captured with resummed *linear* displacements $\Psi^{d,s}$ will break down, and indeed

our fit residuals begin to show noticeable oscillatory behavior at the smallest smoothing scale shown ($R = 10\,h^{-1}\mathrm{Mpc}$). At $R = 15\,h^{-1}\mathrm{Mpc}$ and in the sample variance limit with Gaussian errors, the feature at $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ should be detectable with $\chi^2 = V_{\mathrm{obs}}/(2h^{-3}\ \mathrm{Gpc}^3)$, where V_{obs} is the total observed volume. If we were to instead smooth using the larger $R = 20\,h^{-1}\mathrm{Mpc}$, the χ^2 is roughly halved. For such a smoothing this feature represents a χ^2-penalty of 0.2 for a sample variance limited survey of $14\,000\,\mathrm{deg}^2$ covering $0 \leq z \leq 0.3$, and would be slightly smaller for finite number density.

The pre- and post-reconstruction real-space correlation functions can be directly compared by computing the Fourier transforms of the above fits. However, in comparing our theory with `DarkSky` we found that the $z = 0$, pre-reconstruction halo power spectra all have significant excess power at low k compared to the predictions of linear theory with scale-independent bias. The origin of this excess is unclear, and is not addressed in ref. [1]. It appears to arise from a significant number of low k modes, and so is unlikely to be simply a statistical fluctuation in the initial conditions. It shows up in all of our halo samples, and is highly correlated among mass bins. This excess power is small for modes to the right of the power spectrum peak and probably has only a small impact on the dynamics on BAO scales. In Fourier space we simply confine our fitting and modeling to $k > 0.01\,h\,\mathrm{Mpc}^{-1}$. In configuration space, however, the additional long-wavelength power slightly distorts the shape of the BAO peak, and to enable a fair comparison we have added appropriate long-wavelength modes to our theoretical predictions assuming linear theory; specifically, we find that the fitting form $P_{\mathrm{lw}}(k) = A\,(k/k_0)^n$, where $A = 3.5 \times 10^4 h^{-3}\ \mathrm{Mpc}^3$, $k_0 = 10^{-3}\,h\,\mathrm{Mpc}^{-1}$ and $n = -1.7$, describes well both the long-wavelength excess seen in the power spectrum below $k < 0.01\,h\,\mathrm{Mpc}^{-1}$ and dramatically improves the agreement between the unreconstructed correlation function in theory and `DarkSky`. The contribution to the pre- and post-reconstruction power spectra and correlation function of these long wavelength modes is shown in figure 7. Without the long-wavelength correction, the `DarkSky` results do not agree with theory on the large scales to the right of the BAO peak, which should be well-described within linear theory, nor in the BAO "dip," both pre- and post-reconstruction. Due to the ad-hoc nature of our correction, in the remainder of this section we will focus our comparisons on Fourier space, wherein long-wavelength modes must decouple. However, we caution that small, localized features in Fourier space can cause extended distortions in configuration space where data points are highly correlated. In figure 8, we show the effect of the $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ bump described in the previous section by additively "filling" it with a small, localized Gaussian profile, as shown in the left panel. The effects of this bump, Fourier-transformed, are shown in the right panel: while sub-percent in Fourier space, the $k \simeq 0.1\,h\,\mathrm{Mpc}^{-1}$ feature gives rise to visible distortions to the BAO feature in configuration space.

Finally, fits for the pre- and post-reconstruction power spectra in redshift space are shown in figure 9. We have chosen to summarize the angular dependence of the redshift-space power spectrum in terms of its monopole and quadrupole, though our model predicts the full $P(k, \mu)$ and higher multipoles as well. As in real space, we have fitted for the bias parameters (b_1, b_2) using the unreconstructed data and applied the same set of bias parameters to predict the power spectra in both **Rec-Sym** and **Rec-Iso**. We adopt the full set of six counterterms, three each $\alpha^\ell_{dd}, \alpha^\ell_{ds}, \alpha^\ell_{ss}$ for the monopole ($\ell = 0$) and quadrupole ($\ell = 2$), but also explore the possibility of utilizing only one counterterm α^ℓ per multipole (corresponding to a derivative bias for both the halo density and velocity). In all cases, our base model with six counterterms fits the data at the percent level or below past $k = 0.2\,h\,\mathrm{Mpc}^{-1}$ in both the monopole and quadrupole moments. Notably the Zeldovich approximation produces oscillation-free
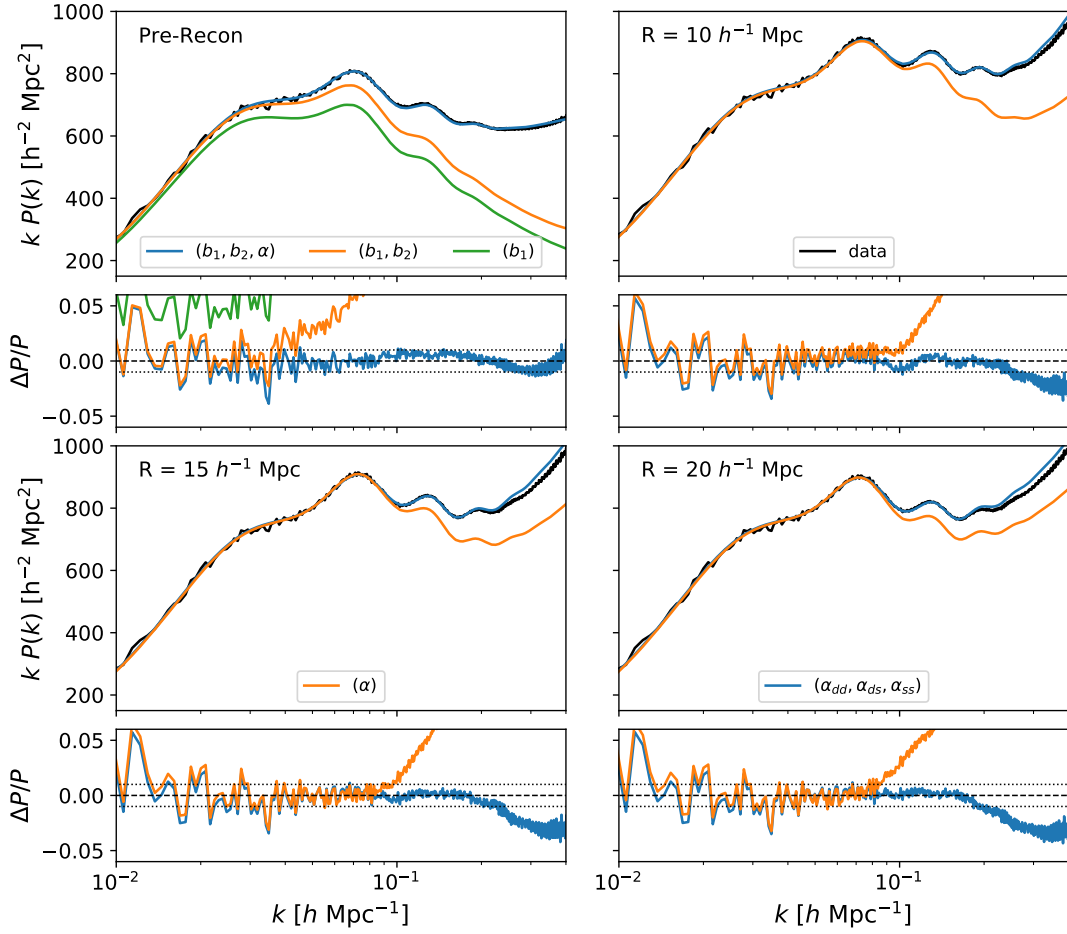
Figure 6. Fits to the pre- and post-reconstruction real-space halo power spectra in `DarkSky` for halos of mass between $12.5 < \log_{10}(M/h^{-1} M_\odot) < 13.0$ at three smoothing scales ($R = 10, 15, 20\, h\, \mathrm{Mpc}^{-1}$), assuming Zeldovich power spectra with biases (b_1, b_2) and one counterterm per spectrum (three total for the reconstructed case). The upper plot of each vertical pair of panels shows the product of the wavevector magnitude and power spectrum $k\, P(k)$ while the lower plot shows the fit residuals as a fraction of measure power $\Delta P/P = (P_{\mathrm{fit}} - P_{\mathrm{nbody}})/P_{\mathrm{nbody}}$. In the top-left pair of panels we show the incremental contributions from b_2 and the counterterm α (which contributes close to 10% of the power at $k = 0.1\, h\, \mathrm{Mpc}^{-1}$) to the fit, which agrees with the simulation at the percent level (dotted line in the lower plots) at all scales shown. In the remaining panels we use the same bias parameters to fit the reconstructed power spectrum, allowing only counterterms to vary. Our model with three counterterms can fit the data at the percent level out to $k = 0.2\, h\, \mathrm{Mpc}^{-1}$, though a bump-like feature at $k = 0.1\, h\, \mathrm{Mpc}^{-1}$ becomes more prominent at smaller smoothing scales, where nonlinear corrections beyond the Zeldovich approximation presumably become more important (see text). Also shown in orange are fits using one counterterm — or equivanlently one derivative bias — which fit less well past $k = 0.1\, h\, \mathrm{Mpc}^{-1}$. We fined that setting the counterterm α_{ss} to zero does not materially affect our fits. Note that there is excess power in the data at the largest scales shown, as discussed in the text.

residuals even in the absence of counterterms (green), with the counterterms providing a physics-based broadband model ($\sim \alpha_{ab}^{\ell} k^2 P_{\ell,ab}$) that reproduces the N-body results at the percent level. Our fits do not explicitly include nonlinear redshift space distortions such as
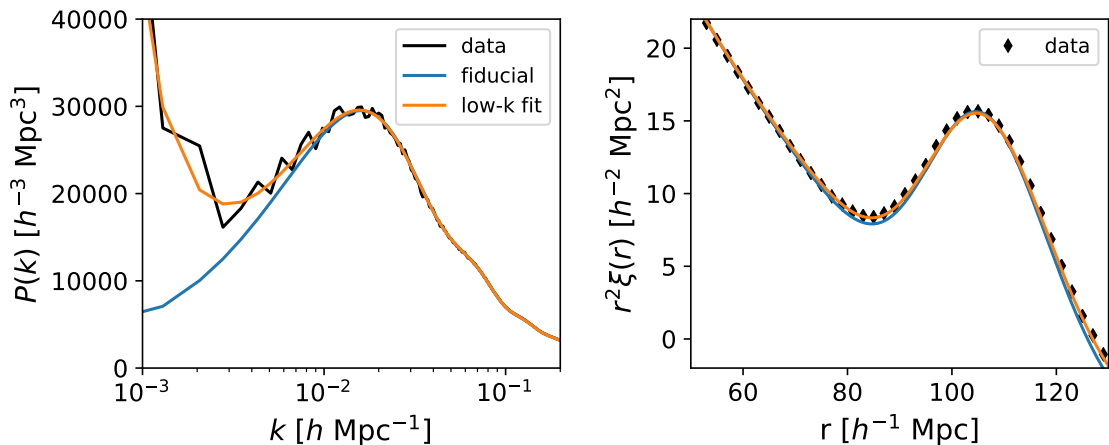
Figure 7. Halos in `DarkSky` exhibit significant excess power compared to theory at large scales in Fourier space which should be well-described by linear theory. (Left) Fits to the real-space power spectrum with and without our ad hoc correction $P_{\rm lw} = A\,(k/k_0)^n$, shown in blue and orange respectively. At the largest scales shown, the excess power is significantly larger than the scatter. The fits prefer slighly different, though qualitatively similar, bias values. (Right) The same fits in configuration space. The uncorrected data systematically trends below the data at separataions above the BAO peak and in the BAO "dip," while the fit with $P_{\rm lw}$ added goes through all the data points.

fingers-of-god, though such effects are perturbatively accounted for by velocity counterterms to lowest order. For completeness, in figure 10 we show the same fits for the mass bin $13.0 < \log(M/h^{-1}M_\odot) < 13.5$, where our model fits the data at percent level over a similar range of scales using the parameters $(b_1, b_2) = (0.23, -1.0)$.

Lastly, let us comment on the comparison fits in pre- and post- reconstructed cases. Given that our shift field, χ, is constructed only from long-wavelength modes explicitly isolated from observed field, δ, by filtering out the nonlinear scales larger than $k \gtrsim 1/R$, we have no reason to suppose that the perturbative structure of our results will significantly change. In other words, by performing the mapping in equation (4.1), we have reconstructed only the long modes, thereby reducing nonlinear smoothing due to large scale (infrared) modes, while the bulk of the small-scale nonlinear modes, as well as FoG effects, should remain unreduced. In addition, Lagrangian perturbation theory (PT) conveniently separates nonlinearities due to long and short modes, exponentially resumming the former while expanding the latter order-by-order [38, 39]. Because of this, we do not expect dramatically different PT behavior in the pre- and post-reconstructed results. These arguments are also supported by figures 9 and 10, which show our model exhibits quantitatively similar degrees of fit pre- and post reconstruction.

8 Comparison to earlier work

There has been significant theoretical activity in modeling post-reconstruction clustering (see references in the introduction). Our framework encompasses most of these previous perturbation theory expressions when appropriate approximations and phenomenological choices are accounted for. To the best of our knowledge, the framework presented here captures
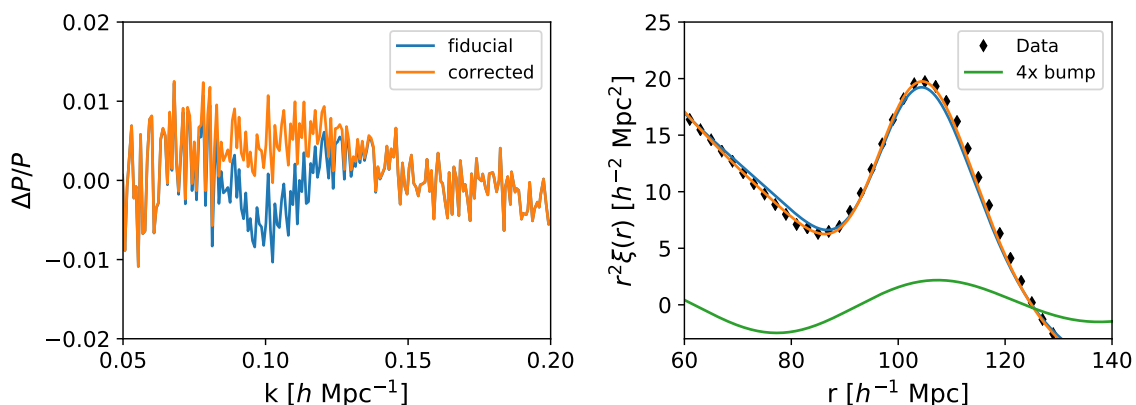
Figure 8. A sub-percent level feature in the power spectrum near $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ can lead to visible distortions in the BAO feature in $\xi(r)$. (Left) Residuals for the fit as a fraction of total measured power in the simulations, as defined in the caption of figure 6. The orange curve shows the residuals when our theory is corrected using a Gaussian profile localized at $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ compared to the fiducial fit (blue), whose residuals exhibit a dip centered at $k = 0.1\,h\,\mathrm{Mpc}^{-1}$. (Right) The fiducial and corrected correlation functions. The bump in the left panel, whose Fourier transform is shown magnified in the green curve, induces distortions in the BAO feature across a range of separations $r \sim 60 - 120\,h^{-1}\mathrm{Mpc}$.

for the first time all of the relevant post-reconstruction effects and is unique in accurately handling both Fourier and configuration space results, in real and redshift space and includes all the bias operators to quadratic order.

Not all models are based on perturbation theory calculations however, and many phenomenological models have been introduced in order to describe the post-reconstruction statistics. Restricting ourselves just to models of the 'standard' reconstruction algorithm [2], section 3.1 of ref. [10] discusses early models (which were of the form $P(k) = B(k)P_{\mathrm{lin}}(k) + A(k)$ with $B(k)$ and $A(k)$ smooth functions). Starting with the first applications to data in ref. [10] the form used to fit reconstructed power spectra is based upon a split between a "smooth" and "wiggle" contribution to $P(k) = P_{\mathrm{nw}}(k) + \Delta P_{\mathrm{w}}(k)$, with a phenomenological damping of the wiggle component motivated by perturbation theory [8]. In ref. [10] the parameters of the model were fit to N-body simulations, and this has become common. This approach has dominated the modeling of observations to date (e.g. refs. [54–56] for recent examples) though ref. [57] is an example of an analysis that did not take this approach. However, we note that the choice of the wiggle/no-wiggle split exhibits a certain amount of freedom in the separation of the wiggle and broadband part. This of course implies that, in order to extract accurate information from the e.g. BAO, either both wiggle and broadband part have to be modeled to the same level of accuracy, or the extracted wiggle part from the data needs to exactly correspond to the model (see also refs. [58, 59] for related discussion). The latter requirement, even though implicitly assumed in most of the current BAO treatments, is rarely subject to performance checks and scrutiny. In this context, it is also worth noting that the common choice of P_{nw} derived in ref. [60] does not fully capture the broadband linear power spectrum at the precision attained by modern Boltzmann codes. Figure 11 shows three possible linear wiggle power spectra, based on no-wiggle spectra computed using the fitting formula from ref. [60], B-splines [61] or a Savitsky-Golay filter;
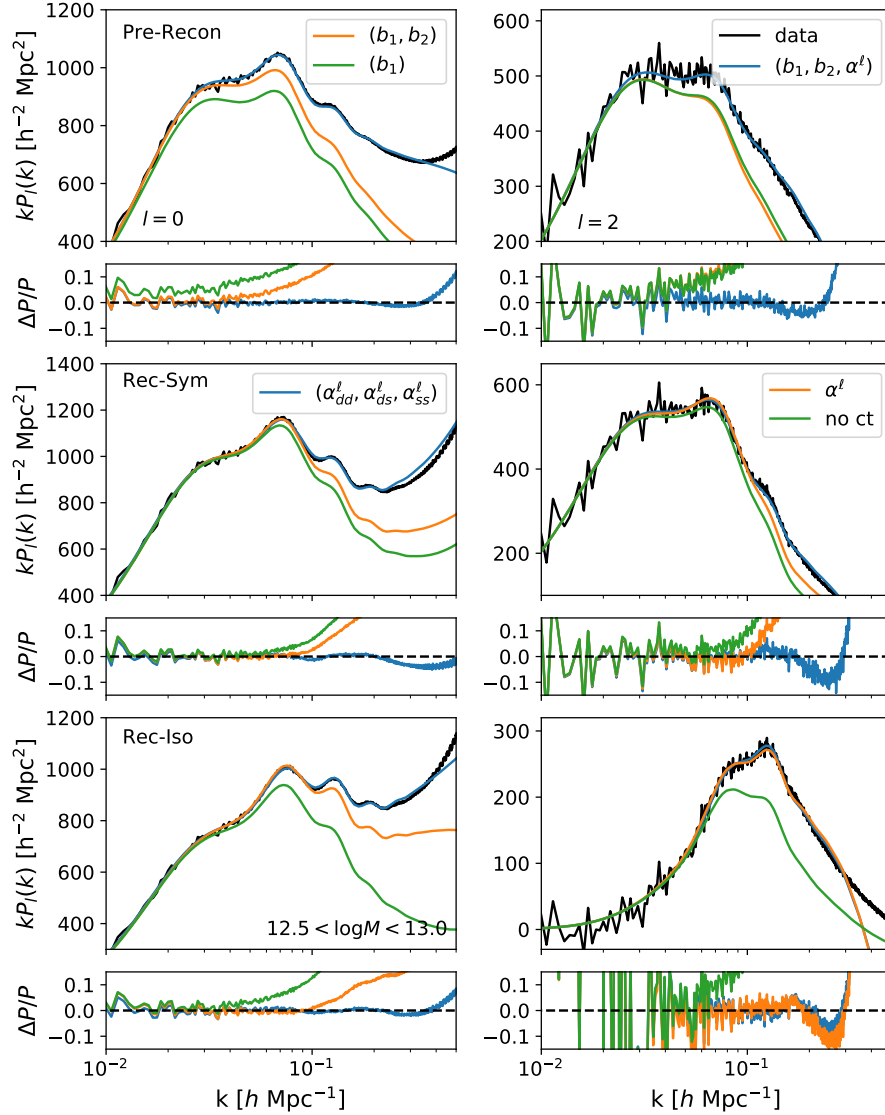
Figure 9. Fits for the pre- and post-reconstruction redshift-space power spectrum monopole (left) and quadrupole (right) for halos in the mass range $12.5 < \log_{10}(M/h^{-1}M_\odot) < 13.0$. The fractional residuals $\Delta P/P$ are defined in figure 6. All spectra were fit using a consistent set of bias parameters $(b_1, b_2) = (0.02, -0.8)$, whose independent contributions are shown in the top row, determined by fitting the pre-reconstruction data, such that only the counterterms were fitted in constructing the curves in the bottom two rows. Our model with the full set of six counterterms — three each for the monopole and quadrupole respectively — fits both the reconstructed monopole and quadrupole in both schemes out to $k = 0.2\,h\,\mathrm{Mpc}^{-1}$ to a few percent and reproduce the phase and amplitude of the oscillatory BAO wiggles.

even the latter two, which agree asymptotically with the full linear theory power spectrum, exhibit noticeably different oscillatory behavior. This indicates that extracting the corresponding wiggle spectra from the data is a challenging and sensitive step which can, on the other hand, be avoided if the broadband is included in the theoretical framework. Models
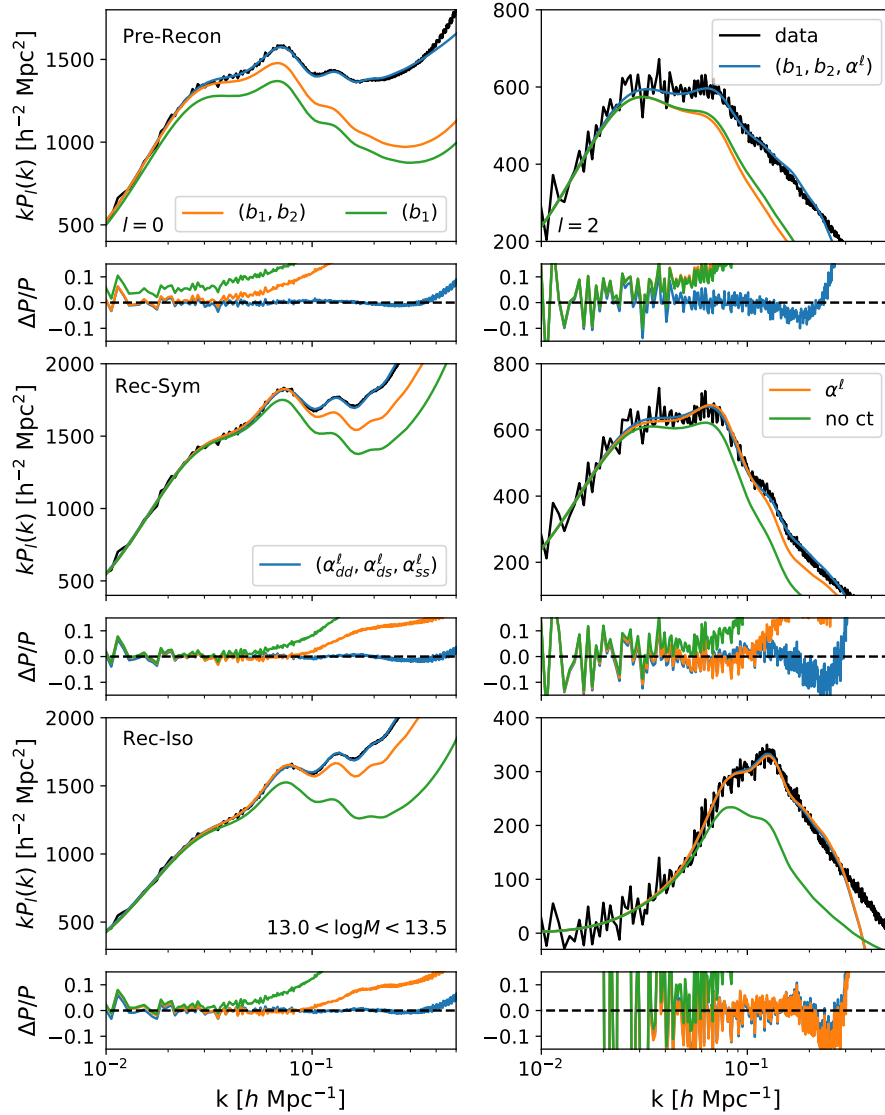
Figure 10. Like figure 9, but for halos in the mass bin $13.0 < \log(M/h^{-1}M_\odot) < 13.5$. Here, our model prefers the bias parameters $(b_1, b_2) = (0.23, -1.0)$ and accurately fits the data over a similar range of scales.

phenomenologically relying on a wide separation of scale, assuming scale-independent bias or sufficient smoothness that could be accounted for by nuisance parameters such as $A(k)$ above, might suffer from overall systematic offsets. Finally, it is also often the case that the nuisance parameters and BAO scaling parameters are not consistent between the configuration-space and Fourier-space analyses (i.e. the two do not form a Fourier transform pair) which could prove problematic if fits in both spaces are combined.

By contrast the Zeldovich calculation above gives a consistent framework for understanding the nonlinear smoothing of the BAO feature, both pre- and post-reconstruction, in both configuration and Fourier space. Roughly speaking, the Gaussian smoothing kernel in the empirical model is replaced by a Lagrangian coordinate-dependent kernel $\exp[-k_i k_j A_{ij}(\mathbf{q})/2]$.
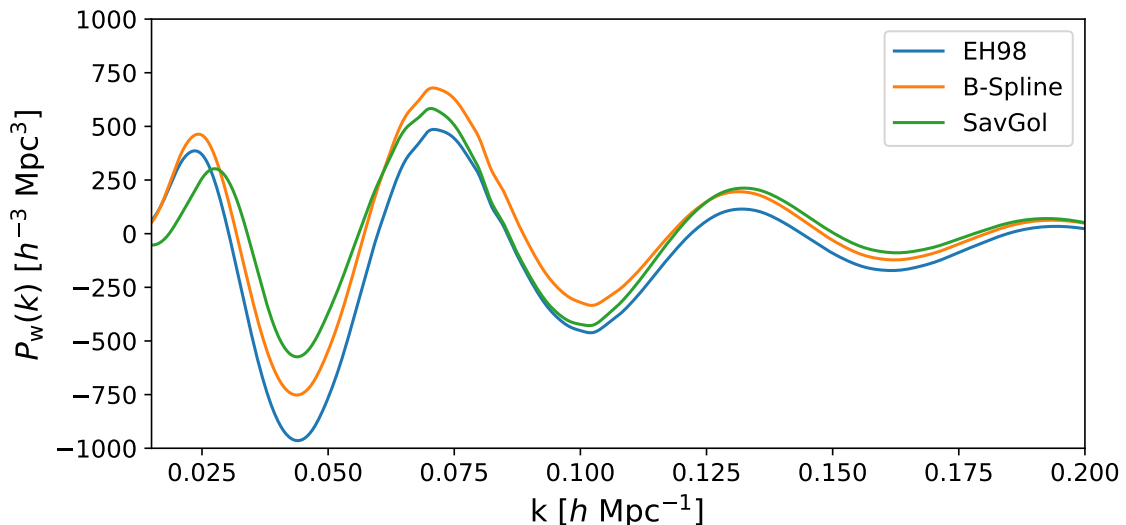
Figure 11. The linear wiggle power spectrum for three choices of P_{nw}. The conventional choice (EH98 [60]) does not accurately capture the large scale power, and we have investigated two possible methods to mitigate this discrepancy: one based on B-splines, described in ref. [61] and another based on a Savitsky-Golay filter in $\ln(k)$. The wiggle power spectra isolated using these three methods exhibit visibly different oscillatory behavior.

One might thus hope to formally extract the model for wiggle-only part as an approximation to the calculation presented in the main body of this paper; indeed, such a calculation was performed in ref. [61] and extended to terms involving linear bias, redshift space distortions and reconstruction in ref. [26].[8] Figure 12 compares the results of our full Zeldovich calculation in the Rec-Sym scheme, with the broadband subtracted out by calculating the corresponding Zeldovich power spectrum using the no-wiggle power spectrum, versus the resummed linear wiggle power spectrum (RWiggle; using the proper exponential damping dependencies given in appendix D), for the same linear bias values and with all higher bias terms set to zero. The two are in excellent agreement, especially in the case of the reconstructed power spectrum, with RWiggle slighly underdamping the BAO wiggles towards small scales compared with the full Zeldovich calculation for the unreconstructed power spectrum.

However, even though RWiggle and the full Zeldovich calculation exhibit a high level of agreement on the shape of the wiggle component, the RWiggle method depends on the separation procedure of the wiggle and broadband components while the full Zeldovich calculation requires no such steps. Specifically, the Zeldovich calculation deals only with the combination $P_w^{Zel} + P_{nw}^{Zel}$, which is obviously invariant under the split, while RWiggle models only the split-dependent P_w^{Zel}. This implies that in order to use RWiggle in practical analyses either the broadband part needs to be modeled to equally high accuracy or a highly accurate wiggle extraction procedure is needed in order to guarantee feasible comparison of theoretical model and the data. The latter seems to be a challenging task, potentially subject to systematic offsets and bias. On the other hand, the resulting differences in the wiggle spectrum

---

[8]We note that redshift-space reconstruction model presented in ref. [26] contains phenomenological damping factors that do not capture the exact behaviour of X_{ds} term given by in equation (4.9). We repeat this calculation and derive the proper damping factors for the wiggle component in appendix D.

should still be broadband and could be fit away using nuisance parameters using sufficiently general broadband models.

Finally, our model differs from most in the literature in taking into account higher bias terms such as b_2 and b_s, allowing us to assess systematic effects introduced by assuming scale-independent bias. These higher biases can contribute both significant broadband power (e.g. the top-left panel of figure 9) and modulate the phase and amplitude of BAO oscillations through mode-coupling effects [13]. However, explicit calculation shows that the latter effect is only noticeable at very high values of bias. Figure 13 shows the effects on the wiggle component of adding nonzero quadratic density and shear biases b_2, b_s, for bias values $(b_1, b_2) \approx (5, 20)$ chosen according to the peak-background split (PBS) on a Press-Schechter mass function [62], and assuming $b_s \approx b_2$, as compared to RWiggle. The quadratic density bias, b_2, induces an apparent phase shift towards large k, and can be seen to be essentially out-of-phase with the linear BAO wiggles; however, these out-of-phase contributions are dramatically reduced by reconstruction. By contrast the shear bias, b_s, produces oscillatory features roughly in-phase with the linear theory contributions and is largely unaffected by reconstruction. For completeness, we have also plotted the potential oscillatory contribution of a derivative bias, b_{∇^2}, which modulates the overall amplitude of the power spectrum and is degenerate with the various counterterms, α_{ab}.

To investigate the extent to which the broadband and oscillatory contributions of higher bias terms can be mitigated by a suitable broadband model, we conducted an exploratory "fit" of the redshift-space monopole and quadrupole pre- and post-reconstruction in the case where the truth is given by the Zeldovich approximation including nonzero b_2 and b_s but fit by an empirical model with only b_1, an isotropic BAO scale paramter α_{BAO} and polynomial broadband contributions of the form employed in ref. [63] before reconstruction. Specifically, we assume an empirical model of the form

$$P_{l,\mathrm{fit}}(k) = \alpha_{\mathrm{BAO}}^{-3} P_{l,b_1}\left(\frac{k}{\alpha_{\mathrm{BAO}}}\right) + \frac{a_{1,l}}{k^3} + \frac{a_{2,l}}{k^2} + \frac{a_{3,l}}{k} + a_{4,l} + a_{5,l}k, \tag{8.1}$$

both pre- and post-reconstruction, where P_{l,b_1} denotes redshift-space multipoles in the Zeldovich approximation with all higher biases set to zero. For this exercise we assumed a sample variance limited survey at $z = 0$ and $z = 1.2$ with Gaussian covariances between the monopole and quadrupole, fit up to $k_{\mathrm{max}} = 0.25\,h\,\mathrm{Mpc}^{-1}$ and note that the results are independent of survey volume. In figure 14 we have plotted the resulting shifts in the BAO scale assuming PBS values for b_1 and b_2, taking values for b_s as a function of b_1 from ref. [64]. At $z = 0$, we find that neglecting higher biases in favor of the empirical model induces shifts of less than half a percent in the BAO scale over a wide range of halo masses both pre- and post-reconstruction, though reconstruction more than halves the forecasted shift for essentially all values of bias surveyed (figure 14). At $z = 1.2$ the shifts are further reduced, amounting to less than a tenth of a percent across a wide range of bias values prior to reconstruction and essentially vanishing post reconstruction. These shifts would be well-within the margin of error of both current and next-generation surveys like DESI [65], especially post-reconstruction, suggesting that nonlinearities (e.g. higher bias) in the power spectrum should not hinder accurate recovery of the BAO signal. On the other hand, the value of the linear bias, b_1, was significantly affected by the choice of broadband model, with fits from the empirical model deviating from the true value by more than five percent in many cases.
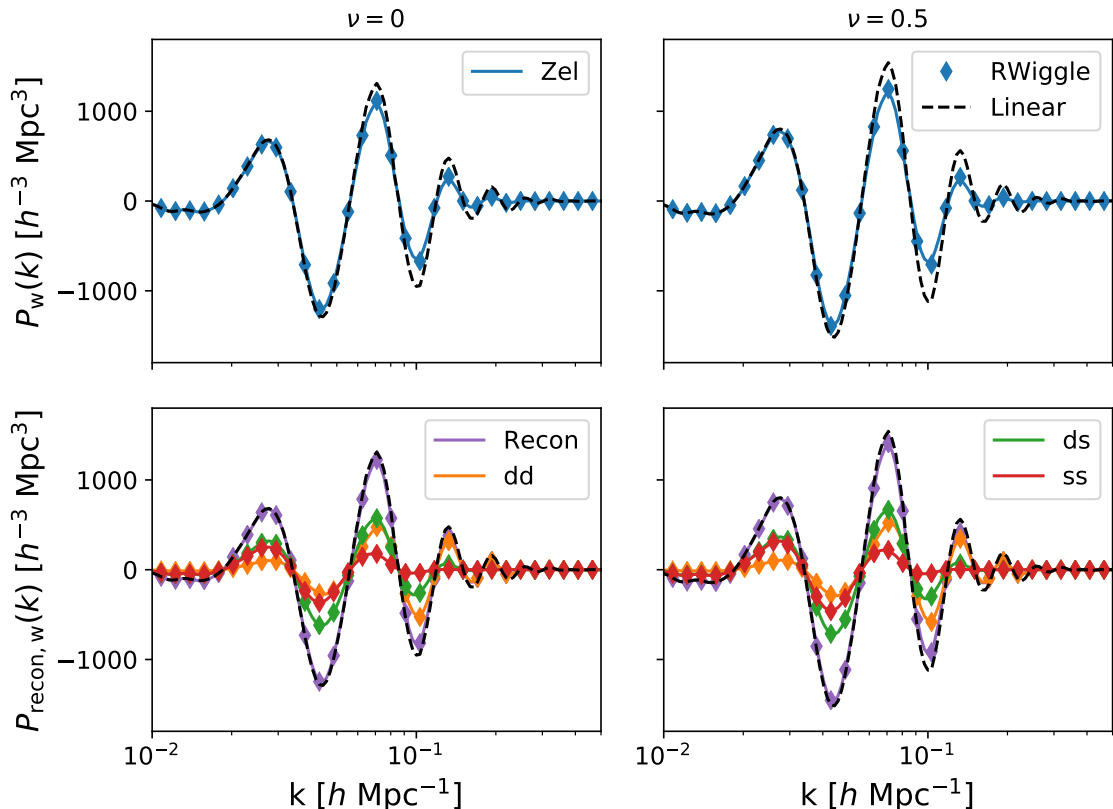
Figure 12. Comparison of Zeldovich with IR-resummed linear theory (RWiggle) for reconstructed and unreconstructed spectra at $z = 0$ and $\nu = 0$ and 0.5 with $b_1 = 0.5$ using Rec-Sym with higher biases set to zero. RWiggle slightly under-predicts damping at high k (but see footnote 14), especially for the unreconstructed power spectra.

9 Conclusions

Baryon acoustic oscillations (BAO) are an important probe of fundamental physics and a prime focus of upcoming surveys such as DESI [65] and EUCLID [66]. The BAO features act as a "standard ruler" whose cosmological evolution is largely immune to astrophysical effects but whose signal-to-noise ratio is lowered by nonlinear structure formation. BAO reconstruction attempts to sharpen the BAO signal by removing some of the nonlinear smearing due to large scale displacements [2]. In this paper we develop an analytical model, within the Lagrangian perturbation theory framework, to study the algorithm for density-field reconstruction proposed in ref. [2]. Linear Lagrangian perturbation theory (the Zeldovich approximation) provides an excellent description of these nearly linear displacements and BAO smoothing pre-reconstruction [28, 37], making LPT a promising arena within which to model the effects of reconstruction.

We develop a self-consistent framework with which to calculate the two-point statistics of galaxies, employing a consistent set of parameters to fit the power spectrum and correlation functions, pre- and post-reconstruction in real and redshift space. The broad validity of such LPT models allows for joint fits to the pre- and post-reconstruction two-point statistics
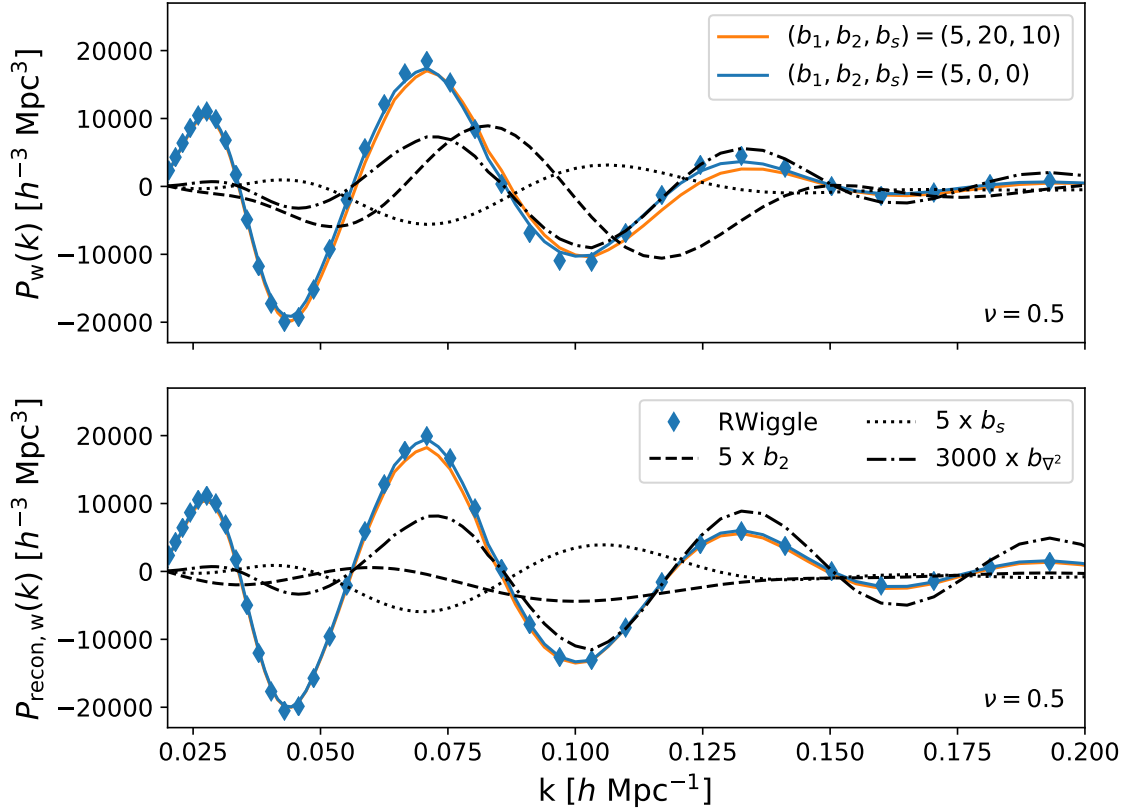
Figure 13. The $z = 0$ Zeldovich power spectrum at $\nu = 0.5$, before and after reconstruction using Rec-Sym, shown with and without contributions from the quadratic bias and shear biases when $(b_1, b_2, b_s) = (5, 20, 10)$. For comparison, the RWiggle prediction is shown in the diamond points, and the isolated b_2 contributions are shown as a black dot-dashed line multiplied by a factor of five. For the unreconstructed spectrum, the b_2 contributions (with shear bias set to zero) can be seen to be essentially out-of-phase with the linear theory wiggles and induce a phase shift in the power spectrum. These contributions are greatly reduced in the reconstructed spectrum. The shear contributions, on the other hand, are more-or-less in phase with linear theory and unchanged by reconstruction. For completeness, we have also plotted contributions from a possible derivative bias b_{∇^2}, which modulate the amplitude of the wiggles in a manner growing with wave number.

enabling e.g. a fit for redshift-space distortions and the linear growth rate, $f\sigma_8$, simultaneously with Alcock-Paczynski distortions constrained by BAO analyses [16]. Based on ref. [50], we derive explicit formulae, to calculate the redshift-space power spectrum within the Zeldovich approximation, both pre- and post-reconstruction, as an infinite series of spherical Bessel transforms. Our model updates the developments for the reconstructed correlation function in ref. [16], and is — as far as we are aware — the first model of reconstruction to include a consistent set of bias terms up to quadratic order, including shear and derivative biases. We show that the oscillatory behavior induced by the quadratic density bias, b_2, are out of phase with the linear BAO feature and greatly reduced post-reconstruction, while those due to the quadratic shear bias, b_s, are in-phase and essentially unchanged. In addition, we show that each multipole moment of the reconstructed power spectrum should be, to lowest
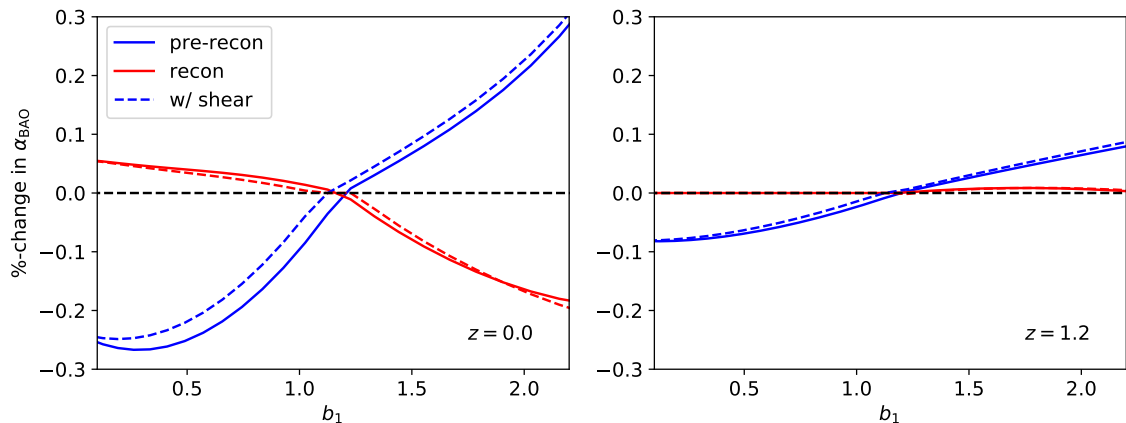
Figure 14. Shifts in the recovered isotropic BAO scale, α_{BAO}, in redshift space fit using a model with only b_1 nonzero and polynomial broadband contributions in both the monopole and quadrpole, when truth is given by the Zeldovich approximation with nonzero quadratic bias. Values of b_1 and b_2 were chosen according to the peak-background split, while values for b_s were taken from ref. [64]. (Left) Shifts in the BAO scale at $z = 0$. Fitting with the empirical model results in only sub-percent shifts across a wide range of halo masses, which are further more than halved after reconstruction. The solid and dashed lines show the shift with and without the quadratic shear bias b_s, whose effect is subdominant to b_2. (Right) The same shifts calculated at $z = 1.2$. Even prior to reconstruction, fitting with the empirical model results in less than a tenth of a percent shift in the BAO scale over a wide range of biases; after reconstruction the shift due to nonlinear bias becomes essentially zero.

order, corrected for by a set of three counterterms each, which perturbatively correct both nonlinear smoothing and broadband power.

We compare our analytic predictions with N-body data from the `DarkSky` simulation [1] at $z = 0$, focusing on halos between $12.5 < \log_{10}(M/M_\odot) < 13.0$. Our base model, involving only b_1 and b_2 and appropriate counterterms, jointly fits the pre-reconstruction real-space power spectrum and redshift-space monopole out to $k = 0.4\,h\,\mathrm{Mpc}^{-1}$, and the quadrupole out to $k = 0.2\,h\,\mathrm{Mpc}^{-1}$, reproducing the oscillatory BAO wiggles in the data with high fidelity. Our model with the same bias parameters performs equally well in configuration space around the BAO scale, though we found it necessary to correct for a large excess in large-scale power encountered in the `DarkSky` data. Utilizing the same values for the bias parameters but allowing counterterms to vary, we find that our model performs similarly in real space post-reconstruction for smoothing scales $R = 15$ and $20\,h^{-1}\mathrm{Mpc}$, reproducing both the oscillatory features and broadband past $k = 0.2\,h\,\mathrm{Mpc}^{-1}$, but fails to reproduce the oscillatory features when $R = 10\,h^{-1}\mathrm{Mpc}$, likely due to the fact that we have worked to lowest order and at $z = 0$ displacements on that scale are significantly nonlinear. We point out a less severe feature in the residuals at $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ that diminishes with larger smoothing scales which we believe arise from higher order terms and caution that neither our calculation nor the standard reconstruction algorithm take these into account. A more complete, iterative reconstruction scheme (e.g. ref. [20]) may reduce these features. The modeling of these nonlinearities, and possible remedies, are beyond the scope of this paper, but as an exploratory example we calculate the effects of one possible nonlinearity due to the mapping between Eulerian and Lagrangian coordinates in appendix E.

Our model also predicts the multipole moments of the redshift-space power spectrum and correlation functions in both of the redshift-space schemes (**Rec-Sym** and **Rec-Iso**) we consider. This is critical in order for it to be applied to data, since the most constraining BAO measurements are performed in redshift space. The model provides a good fit to the monopole and quadruople moments of $P(k)$ measured in `DarkSky` in both the **Rec-Sym** and **Rec-Iso** schemes for smoothing scales of $R = 15\,h^{-1}\mathrm{Mpc}$ or larger (at $z = 0$). Again, for smaller smoothing scales the Zeldovich model differs from the N-body results (as expected). These effects would be smaller at higher redshift, where the theory is more likely to be applied.

Finally, there exists an extensive literature studying the modeling of reconstruction and the BAO signal, and we compare our model to several existing alternatives. One popular technique, based on ref. [8], is to separate the power spectrum into a smooth "no wiggle" component and an oscillatory "wiggle component," and to damp the latter by an exponential factor fit to simulations while supplementing the former with a polynomial in wavenumber to fit the broadband power. This technique can be more rigorously derived as a particular resummation of the nonlinear contribution of long-wavelength modes much like our Zeldovich calculation itself [26, 61], in which case the damping parameters can be derived theoretically. When the "wiggle" components are isolated we find that the latter is in excellent agreement with our Zeldovich calculation, particularly after reconstruction. In appendix D we re-derive the IR-resummed "wiggle" power spectrum (RWiggle) directly within our Zeldovich framework, updating the exponential damping for the cross term P^{ds}. We highlight that our Zeldovich framework naturally encompasses broadband effects, while methods depending on wiggle/no-wiggle splitting might be subject to additional systematic offsets and biases. These could originate from the fact that the wiggle/no-wiggle splitting is not unique, and thus relies on correctly predicting the broadband or extracting the corresponding wiggle part from the data to high accuracy. On the other hand, the Zeldovich framework correctly captures broadband power over a large range of scales in addition to reproducing the oscillatory features in the reconstructed power spectrum. In fits to N-body data, we show how counterterms correct the sharpness of the BAO feature and broadband power simultaneously and consistently. Moreover, our model goes beyond linear bias to include quadratic density and shear bias, which we show contribute oscillatory terms to $P(k)$ that vary independently in amplitude and phase.

We close by noting a few avenues for future work. An obvious extension of our model is to include nonlinearities arising both from gravitational clustering and the reconstruction itself (e.g. appendix E). The former may be most easily included in the context streaming models [36, 50], wherein the real-space modifications due to reconstruction and those proportional to the growth rate f can be separately treated as modifications to the statistics of the galaxy density and galaxy density-weighted velocities, respectively, and which in addition have the advantage of resumming biased contributions to redshift-space distortions as well as nonlinear redshift-space phenomena like fingers-of-god. It is, however, not a-priori obvious which type of nonlinearity will present the most significant corrections. Other fruitful avenues would be to investigate the impact of wrong parameters on reconstruction (e.g. refs. [27, 46]) or to update the present treatment to newer reconstruction techniques. Finally, one could investigate the utility of our model for upcoming surveys like DESI [65] or Euclid [66]. These surveys will operate at higher redshifts where our calculations should perform even better, and our model will be a natural arena in which to understand the effects of highly biased tracers and the effects of cosmic evolution (e.g. evolving b and σ_8) on the BAO feature measured in broad redshift bins.

We have publicly released our codes for configuration[9] and Fourier[10] space reconstruction, with the hope that they will be useful to other researchers. We have checked that the Hankel transform of the Fourier space code agrees, term by term, with the configuration space code to better than 1%, except very close to zero crossings.

Acknowledgments

A Cross-spectra correlators

In this appendix we give analytic expressions for the two-point functions required to calculate cross-spectra, which are slightly different from those required to calculate the auto-spectra more commonly seen in the literature.

The two-point function for the Lagrangian displacement between two species separated by Lagrangian distance $\mathbf{q}$ is given by

$$A_{ij}^{ab}(\mathbf{q}) = \langle \boldsymbol{\Psi}_i^a \boldsymbol{\Psi}_j^a \rangle + \langle \boldsymbol{\Psi}_i^b \boldsymbol{\Psi}_j^b \rangle - 2\langle \boldsymbol{\Psi}_i^a(\mathbf{q}_2)\boldsymbol{\Psi}_j^a(\mathbf{q}_1) \rangle \equiv X^{ab}(q)\ \delta_{ij} + Y^{ab}(q)\ \hat{q}_i\hat{q}_j \qquad (A.1)$$

where

$$X^{ab}(q) = \frac{2}{3}\int \frac{dk}{2\pi^2}\left[\ \frac{1}{2}\Big(P_L^{aa}(k) + P_L^{bb}(k)\Big) - \Big(j_0(kq) + j_2(kq)\Big)P_L^{ab}(k)\right]$$

$$Y^{ab}(q) = 2\int \frac{dk}{2\pi^2}\ j_2(kq)P_L^{ab}(k). \qquad (A.2)$$

Note that for cross spectra $X^{ab}(q)$ does not in general vanish as $q \to 0$. Similarly we have

$$U_i^b = \langle \Delta_i^{ab}\delta_0(\mathbf{q}_2)\rangle \equiv U^b(q)\hat{q}_i, \quad U_i^a = \langle \Delta_i^{ab}\delta_0(\mathbf{q}_1)\rangle \equiv U^a(q)\hat{q}_i \qquad (A.3)$$

where

$$U^a(q) = -\int \frac{dk\ k}{2\pi^2}\ j_1(kq)P_L^{am}(k) \qquad (A.4)$$

and P^{am} is the linear theory cross spectrum between tracer a and matter, and the corresponding expression for U^b follows by direct substitution.

Finally, the non-scalar shear correlators are given by

$$V_i^{ab} = V^a(q)\hat{q}_i, \quad \Upsilon_{ij}^a = X_{s^2}^a(q)\delta_{ij} + Y_{s^2}^a(q)\hat{q}_i\hat{q}_j \qquad (A.5)$$

---

[9]https://github.com/martinjameswhite/ZeldovichRecon.
[10]https://github.com/sfschen/ZeldovichReconPk.

where the functions of q are given by

$$V^a(q) = 2 \int \frac{dk\,k}{2\pi^2} P_L^{am}(k) \left[\frac{4}{15} j_1(kq) - \frac{2}{5} j_3(kq) \right] \int \frac{dk\,k^2}{2\pi^2} P_L^{mm}(k)\,j_2(kq) \qquad \text{(A.6)}$$

and

$$X_{s^2}(q) = 4(\mathcal{J}_3^a)^2, \quad Y_{s^2}(q) = 6(\mathcal{J}_2^a)^2 + 8\mathcal{J}_2^a\mathcal{J}_3^a + 4\mathcal{J}_2^a\mathcal{J}_4^a + 4(\mathcal{J}_3^a)^2 + 8\mathcal{J}_3^a\mathcal{J}_4^a + 2(\mathcal{J}_4^a)^2 \text{ (A.7)}$$

where following refs. [36, 37] we have defined

$$\mathcal{J}_2^a = \int \frac{dk\,k}{2\pi^2} P_L^{am}(k) \left[\frac{2}{15} j_1(kq) - \frac{1}{5} j_3(kq) \right] \qquad \text{(A.8)}$$

$$\mathcal{J}_3^a = \int \frac{dk\,k}{2\pi^2} P_L^{am}(k) \left[-\frac{1}{5} j_1(kq) - \frac{1}{5} j_3(kq) \right] \qquad \text{(A.9)}$$

$$\mathcal{J}_4^a = \int \frac{dk\,k}{2\pi^2} P_L^{am}(k)\,j_3(kq). \qquad \text{(A.10)}$$

The remaining scalar shear correlators, ζ and χ^{12}, are identical to those found in evaluating the auto-spectrum, and we refer readers to refs. [36, 37].

B The pre- and post-reconstruction Zeldovich propagator

In this appendix we give expressions for the normalized cross-spectrum between the initial and final or reconstructed field. This is essentially a correlation coefficient, though it is also referred to as the propagator [67]. Specifically we define $G_a(\mathbf{k}) = \langle \delta_0(-\mathbf{k})\delta_a(\mathbf{k}) \rangle / \langle \delta_0(-\mathbf{k})\delta_0(\mathbf{k}) \rangle$, within the Zeldovich approximation, which quantifies the extent to which a tracer field a is (de)correlated with the initial density δ_0, and apply our results to derive the reconstructed field. Our results generalize those in ref. [14] to include halo bias.

As defined, the propagator G_a is a special case of the cross spectrum and can be evaluated using equation (2.4) by assuming that the linear field δ_0 is a tracer b with displacement $\mathbf{\Psi}^b = 0$ and bias functional $F^b = \delta_0$, such that any Lagrangian two-point functions involving the displacement $\mathbf{\Psi}^b$ (e.g. U^b) or higher biases (e.g. b_2^b) vanish identically. Unlike in the conventional case, however, F^b does not have a zero order piece equal to unity — we can thus compute our result directly by taking the derivative of equation (2.4) with respect to b_1^b with the above assumptions. This gives

$$P_L(k)G_a(\mathbf{k}) = \int d^3\mathbf{q}\,e^{i\mathbf{k}\cdot\mathbf{q}}\,e^{-k^2\Sigma_{aa}^2/4}\Big[ik_i U_i^a + b_1^a \xi_L\Big] = e^{-k^2\Sigma_{aa}^2/4}\Big(P^{am}(k) + b_1^a P^{mm}(k)\Big), \tag{B.1}$$

where we have used that

$$A_{ij}^{a0}(\mathbf{q}) = \langle \mathbf{\Psi}_i^a \mathbf{\Psi}_j^a \rangle \equiv \frac{1}{2}\Sigma_{aa}^2 \delta_{ij}, \tag{B.2}$$

only receives "half" of the zero-point contribution cf. the power spectrum (where $\langle \mathbf{\Psi}^b \mathbf{\Psi}^b \rangle \neq 0$). Note that all higher bias contributions vanish. The generalization to the redshift space field can be straightforwardly accomplished by multiplying by appropriate factors of R_{ij} in the numerator, though we will focus on real space in this appendix as RSD introduce an equally important but parallel form of decorrelation into the problem.

From the above results, the reconstructed-field ($\delta_{\mathrm{rec}} = \delta_d - \delta_s$) propagator can be written as $G_{\mathrm{rec}} = G_d - G_s$, where

$$G_d(\mathbf{k}) = \frac{e^{-k^2 \Sigma_{dd}^2/4}\left(P^{\mathrm{dm}}(k) + b_1 P^{mm}(k)\right)}{P_L(k)}, \qquad G_s(\mathbf{k}) = \frac{e^{-k^2 \Sigma_{ss}^2/4} P^{\mathrm{sm}}(k)}{P_L(k)}, \qquad \text{(B.3)}$$

where the various linear spectra are defined as in equation (4.6). The real-space post-reconstruction propagator is then

$$G_{\mathrm{rec}}(\mathbf{k}) = e^{-k^2 \Sigma_{dd}^2/4}\left[(1 - \mathcal{S}(k) + b_1\right] + \mathcal{S}(k) e^{-k^2 \Sigma_{ss}^2/4}. \qquad \text{(B.4)}$$

The expression for G_{rec} helps to quantify how much of the decorrelation between the initial conditions and the final field arises due to bulk motions, and the manner in which this can be restored by the standard reconstruction algorithm. Roughly speaking, reconstruction reduces the decorrelation from the full matter Σ^2 to Σ_{dd}^2 past the smoothing scale for the matter piece, with the correlation at low k close to unity assuming that the damping due to Σ_{ss}^2 there is negligible.

C Integrals for redshift space distortions via direct Lagrangian expansion

In this section we describe how to perform the three-dimensional integrals that occur when calculating redshift space power spectra in the Lagrangian formalism.

C.1 Angular integrals

Calculations in Lagrangian perturbation theory frequently require evaluating integrals of the form

$$I_\ell^{(n)}(A, B) = \frac{1}{2} \int_{-1}^{1} d\mu \; \mu^{2\ell+n} e^{iA\mu + B\mu^2}. \qquad \text{(C.1)}$$

These integrals can be conveniently expressed as infinite sums of spherical Bessel functions [50, 68], e.g.

$$I_\ell^{(0)} = \frac{(-1)^\ell e^B}{B^\ell} \sum_{n=0}^{\infty} U(-\ell, n-\ell+1, -B) \left(\frac{-2B}{A}\right)^n j_n(A)$$

$$I_\ell^{(1)} = i\frac{(-1)^\ell e^B}{B^\ell} \sum_{n=0}^{\infty} U(-\ell, n-\ell+1, -B) \left(\frac{-2B}{A}\right)^n j_{n+1}(A)$$

$$I_\ell^{(2)} = \frac{(-1)^\ell e^B}{B^\ell} \sum_{n=0}^{\infty} \left[U(-\ell, n-\ell+1, -B) + \frac{n}{B} U(-\ell, n-\ell, -B)\right] \left(\frac{-2B}{A}\right)^n j_n(A), \quad \text{(C.2)}$$

where $U(a, b, z)$ denotes the confluent hypergeometric function of the second kind.

C.2 Direct Lagrangian expansion: MI

Calculating the power spectrum in redshift space within the Zeldovich approximation requires a few extra steps when compared to the calculation in real space due to the line-of-sight dependence of RSD. In the following two sections we extend the two methods presented in ref. [50], MI and MII, to include bias terms up to quadratic order. These two methods correspond, roughly speaking, to active and passive transformations in Fourier space via

$R_{ij} = \delta_{ij} + f\hat{n}_i\hat{n}_j$, respectively. For the general case, we found MII to be somewhat more convenient, and will therefore offer only a cursory description of MI, except for a special case involving the displaced-shifted field cross spectrum in **Rec-Iso**.

In MI, the (halo auto-) power spectrum in redshift space is given by

$$P_s(k) = \int d^3q \; e^{i\mathbf{k}\cdot\mathbf{q} - \frac{1}{2}k_i k_j A_{ij}^s}\left[1 + 2ib_1 k_i U_i^s + b_1^2 \xi_L + \cdots\right], \tag{C.3}$$

where the superscript s denotes tensor quantities transformed by R_{ij}, e.g. $U_i^s = R_{ij}U_j$. Each quantity in the above integral can be written in terms of scalar functions and dot products between three unit vectors ($\hat{q}$, $\hat{k}$ and $\hat{n}$) whose angular structure underlies redshift space distortions; these are, in particular,

$$\hat{n}\cdot\hat{k} = \nu, \quad \hat{q}\cdot\hat{k} = \mu, \quad \hat{q}\cdot\hat{n} = \mu\nu + \sqrt{1-\mu^2}\sqrt{1-\nu^2}\cos\phi, \tag{C.4}$$

where ϕ is the azimuthal angle in a polar coordinate system where the zenith is given by $\hat{k}$ and the plane $\phi = 0$ is spanned by $\hat{k}$ and $\hat{q}$. The effect of the transformation $\mathbf{R}$ can then be captured by how the usual tensor basis $\hat{q}_i, \delta_{ij}$, etc. is affected:

$$k_i k_j \delta_{ij} \rightarrow k_i k_j R_{\text{in}} R_{jn} = k^2\left[1 + f(2+f)\nu^2\right] \tag{C.5}$$

$$k_i \hat{q}_i \rightarrow k_i R_{ij}\hat{q}_j = k\mu\left[1 + f\nu^2 + f\nu^2\gamma(\mu,\nu)\cos\phi\right], \tag{C.6}$$

where we have defined $\gamma(\mu,\nu) = \sqrt{1-\mu^2}\sqrt{1-\nu^2}/\mu\nu$. The azimuthal dependence of A_{ij}^s requires us to calculate polar-coordinate integrals of the form [50]

$$I_n(f,\mu,\nu) = \int_0^{2\pi}\frac{d\phi}{2\pi}\; e^{-\frac{1}{2}k^2 Y(\alpha_2\gamma\cos\phi + \alpha_3\gamma^2\cos^2\phi)\mu^2}(\gamma\cos\phi)^n \tag{C.7}$$

where $\alpha_2 = 2f\nu^2(1 + f\nu^2)$ and $\alpha_3 = f^2\nu^4$. These can be calculated as analytic power series in μ by taking derivatives of the identity

$$I_\phi(\alpha,\beta,\mu) = \int_0^{2\pi}\frac{d\phi}{2\pi}\; e^{\alpha\mu\sqrt{1-\mu^2}\cos\phi + \beta(1-\mu^2)\cos^2\phi} = \sum_{\ell=0}^\infty F_\ell(\alpha,\beta)\left(\alpha^2\mu^2/\beta\right)^\ell \tag{C.8}$$

where

$$F_\ell(\alpha,\beta) = \sum_{m=0}^\ell \frac{\Gamma(m + \frac{1}{2})}{\pi^{1/2}\Gamma(m+1)\Gamma(1+2m-\ell)\Gamma(2\ell-2m+1)}\left(-\frac{\beta^2}{\alpha^2}\right)^m$$
$$\times M\left(\ell - 2m; \ell - m + \frac{1}{2}; \frac{\alpha^2}{4\beta}\right) M\left(m + \frac{1}{2}; m + 1; \beta\right)$$

and $M(a,b,z)$ are hypergeometric functions of the first kind. Each term proportional to μ^n can in turn be evaluated as a spherical Bessel transform as in equation (C.2).

A simplified but demonstrative example of this calculation can be found in the calculation of the displaced-shifted cross spectrum in redshift space reconstruction via **Rec-Iso**. The H_ℓ^0 expansion in equation (4.16) is essentially equation (C.8) in the limit where $\beta \rightarrow 0$. To proceed from equation (4.16), we can use the identities in C.1 and refactor the resulting double sum over n and ℓ to get

$$P^{(ds)}(\mathbf{k}) = e^{-\frac{1}{2}k^2(\alpha_0\Sigma^{(dd)^2} + \Sigma^{(ss)^2})}\; 4\pi\sum_{n=0}^\infty \int dq\; q^2 e^{k^2(1+f\nu^2)(\tilde{X}^{ds} + \tilde{Y}^{ds})}\left(\frac{-2k\tilde{Y}^{ds}}{q}\right)^n (1+f\nu^2)^n$$

$$\times\left[K_n^{(0)}(q)j_n(kq) - b_1 k U^d(q)K_n^{(0)}(q)j_{n+1}(kq) - \frac{1}{2}b_2 k^2 U^d(q)^2 K_n^{(2)}(q)j_n(kq) + \cdots\right] \tag{C.9}$$

where the redshift-space kernels are given by

$$K_n^{(0)}(q) = \sum_{\ell=0}^{\infty} \left(-\frac{f\nu\sqrt{1-\nu^2}}{1+f\nu^2} \right)^\ell H_\ell(A) \, U(-\ell, n-\ell+1, -B)$$

$$K_n^{(2)}(q) = \sum_{\ell=0}^{\infty} \left(-\frac{f\nu\sqrt{1-\nu^2}}{1+f\nu^2} \right)^\ell H_\ell(A) \left[U(-\ell, n-\ell+1, -B) + \frac{n}{B} \, U(-\ell, n-\ell, -B) \right]$$

and, as before, $A = k^2 f\nu\sqrt{1-\nu^2}\,\tilde{Y}^{ds}$ and $B = k^2(1+f\nu^2)\,\tilde{Y}^{ds}$. Deriving these kernels for the other terms is entirely analagous.[11]

C.3 Direct Lagrangian expansion: M$_{\rm II}$

An equivalent approach to rotating each Lagrangian displacement $\boldsymbol{\Psi}$ by $\mathbf{R}$ is to instead passively transform the Fourier basis by $\mathbf{R}^{\rm T}$, such that the wavenumber is given by $K_i \equiv R_{ij}k_j$. Defining $\mu = \hat{K} \cdot \hat{q}$ as the angle between the transformed wave vector and Lagrangian separation $\mathbf{q}$, we have the dot products

$$K \cdot K = k^2 \left(1 + f(2+f)\nu^2 \right)$$

$$\hat{K} \cdot \hat{n} = k\nu(1+f)$$

$$\hat{n} \cdot \hat{q} = \frac{\mu\nu(1+f)}{\sqrt{1+f(2+f)\nu^2}} + \sqrt{1-\mu^2}\sqrt{1 - \left(\frac{\nu^2(1+f)^2}{1+f(2+f)\nu^2} \right)} \cos\phi$$

$$k \cdot q = kq\Big(c\mu - s\sqrt{1-\mu^2}\cos\phi\Big), \ c = \frac{1+f\nu^2}{\sqrt{1+f(2+f)\nu^2}}, \ s = \frac{f\nu\sqrt{1-\nu^2}}{\sqrt{1+f(2+f)\nu^2}}. \quad \text{(C.10)}$$

The power spectrum in this frame is then simply given by substituting K for k in the real space expression; e.g. the term in the exponential can be written in terms of the vector $\mathbf{K}$ as

$$-\frac{1}{2}k_i k_j A_{ij}^s = -\frac{1}{2}k_i k_j R_{il}R_{jm}A_{lm} = -\frac{1}{2}K_\ell K_m A_{\ell m}. \quad \text{(C.11)}$$

In redshift space, unlike in real space, the Fourier product $\mathbf{k} \cdot \mathbf{q}$ requires azimuthal angle integrals due to the appearance of $\cos\phi$ in the final line of (C.10); however, in M$_{\rm II}$ the azimuthal dependence is factored entirely into the Fourier factor $i\mathbf{k} \cdot \mathbf{q}$ such that the ϕ integral can be performed analytically:

$$P_s(k) = \int dq \, d\mu \, q^2 \, e^{ikqc\mu - \frac{1}{2}K^2(X+Y\mu^2)} \left(\int d\phi \, e^{-ikqs\sqrt{1-\mu^2}\cos\phi} \right) \left[1 + 2ib_1 K_i U_i(\mathbf{q}) + \dots \right]$$

$$= 2\pi \int dq \, d\mu \, q^2 \, e^{ikqc\mu - \frac{1}{2}K^2(X+Y\mu^2)} \, J_0(kqs\sqrt{1-\mu^2}) \left[1 + 2i\mu \, b_1 K U(q) + \dots \right]. \quad \text{(C.12)}$$

The remaining q and μ integrals can then be calculated using the usual combination of spherical-Bessel decompositions and Hankel transforms with the help of the identity [50]:

$$\int d\mu \, e^{i\mu A + \mu^2 B} J_0(C\sqrt{1-\mu^2}) = 2e^B \sum_{\ell=0}^{\infty} \left(\frac{-2}{\rho} \right)^m \tilde{G}_\ell^{(0)}(A,B,\rho) j_\ell(\rho), \quad \text{(C.13)}$$

---

[11]The mater contribution was given in ref. [50].

where $\rho = \sqrt{A^2 + C^2}$ and the function $\tilde{G}_m(A, B, \rho)$ is given by

$$\tilde{G}_m^{(0)}(A, B, \rho) = \sum_{n=m}^{\infty} f_{nm} \left(\frac{BA^2}{\rho^2}\right)^n {}_2F_1\left(\frac{1}{2} - n, -n; \frac{1}{2} - m - n; \frac{\rho^2}{A^2}\right),$$

$_2F_1$ is the ordinary hypergeometric function, and we have defined

$$f_{nm} = \frac{\Gamma(m + n + \frac{1}{2})}{\Gamma(m + 1)\Gamma(n + \frac{1}{2})\Gamma(1 - m + n)}. \tag{C.14}$$

In our specific case, $\rho = kq$, $B = -K^2 Y/2$ and $A = kqc$. Defining

$$I_n = \int d\mu \; (i\mu)^n e^{iA\mu + \mu^2 B} J_0(C\sqrt{1 - \mu^2}) \equiv 2e^B \sum_{l=0}^{\infty} \left(\frac{-2}{\rho}\right)^l \tilde{G}_l^{(n)}(A, B, \rho) j_\ell(\rho) \tag{C.15}$$

such that $I_n = I_0^{(n)}(A)$ is the n$^{\text{th}}$ full derivative of I_0 with respect to A, we have recursively

$$\tilde{G}_l^{(n)} = \frac{d\tilde{G}_l^{(n-1)}}{dA} + \frac{A}{2} \; \tilde{G}_{l-1}^{(n-1)}. \tag{C.16}$$

The first two derivatives of $\tilde{G}^{(0)}$ are given by

$$\frac{d\tilde{G}_m^{(0)}}{dA} = \sum_{n=m}^{\infty} \left(\frac{BA^2}{\rho^2}\right)^n f_{nm} \left[\left(\frac{2n}{A} - \frac{2nA}{\rho^2}\right) {}_2F_1\left(\frac{1}{2} - n, -n; \frac{1}{2} - m - n; \frac{\rho^2}{A^2}\right)\right. \tag{C.17}$$

$$\left. + \left(-\frac{2\rho^2}{A^3} + \frac{2}{A}\right) \frac{(\frac{1}{2} - n)(-n)}{(\frac{1}{2} - m - n)} {}_2F_1\left(\frac{3}{2} - n, 1 - n; \frac{3}{2} - m - n; \frac{\rho^2}{A^2}\right)\right]$$

$$\frac{d^2\tilde{G}_m^{(0)}}{dA^2} = \sum_{n=m}^{\infty} \left(\frac{BA^2}{\rho^2}\right)^n f_{nm} \left(\frac{\rho^2 - A^2}{\rho^4}\right) \left[(2m - 1 - 4n(m+1)) \; {}_2F_1\left(\frac{1}{2} - n, -n; \frac{1}{2} - m - n; \frac{\rho^2}{A^2}\right)\right.$$

$$\left. + (1 - 4n^2 + m(4n - 2)) \; {}_2F_1\left(\frac{3}{2} - n, -n; \frac{1}{2} - m - n; \frac{\rho^2}{A^2}\right)\right]. \tag{C.18}$$

These are sufficient to calculate all terms up to quadratic order in the bias expansion. In our fiducial cosmology at $z = 0$ and along the line of sight ($\nu = 1$) we find that the sums in $\tilde{G}_m$ converge at the sub-percent level within thirty summands.

From the above, contributions to the redshift-space Zeldovich power spectrum can be calculated using spherical Bessel transforms of the specific form

$$P_s(\mathbf{k}) \ni 4\pi \sum_{\ell=0}^{\infty} \int dq \; q^2 \; e^{-\frac{1}{2}K^2(X+Y)} \left(\frac{-2}{kq}\right)^\ell \tilde{G}_\ell^{(n)}(kqc, -\frac{1}{2}K^2 Y, kq) \; \mathcal{A}_n(q) \; j_\ell(kq), \tag{C.19}$$

where the scalar function $\mathcal{A}_n$ are tabulated in table 2.

D Wiggle/no-wiggle split

Most analyses of BAO data to date have employed empirical models for the post-reconstruction power spectrum or correlation function often motivated by theoretical calculations and calibrated to N-body simulations. Refs. [26, 61] showed that the analytic form

$\mathcal{A}_n$	$n=0$	$n=1$	$n=2$	...
1	1	0	0	
b_1	0	$2KU(q)$	0	
b_1^2	$\xi_L(q)$	0	$K^2 U(q)^2$	
b_2	0	0	$K^2 U(q)^2$	
$b_1 b_2$	0	$KU(q)\xi_L(q)$	0	
b_2^2	$\frac{1}{2}\xi_L(q)^2$	0	0	
b_s	$-K^2 X_{s^2}(q)$	0	$K^2 Y_{s^2}(q)$	
$b_1 b_s$	0	$2KV(q)$	0	
$b_2 b_s$	$\chi^{12}(q)$	0	0	
b_s^2	$\zeta(q)$	0	0	

Table 2. Table of power spectrum contributions in $\mathrm{M_{II}}$.

of these empirical models can be interpreted within perturbation theory as a resummation of bulk displacements at the BAO scale. In this appendix we re-derive their results within our Zeldovich calculation, updating the scale dependences and redshift-space factors where appropriate.

Let us first examine the displaced-displaced cross spectrum in redshift space. Following refs. [26, 61] we split the displacement two-point function into $A_{ij}^{dd} = A_{ij}^{dd,\mathrm{nw}} + \Delta A_{ij}^{dd,\mathrm{w}}$, where the no-wiggle and wiggle pieces are calculated by substituting P_{nw} and ΔP_{w} into equation (A.2). Making the assumption that the latter, $\Delta A_{ij}^{dd,\mathrm{w}}$, is small enough as to be perturbative,[12] we can Taylor expand the exponential in the Zeldovich integrand to get

$$P^{dd}(\mathbf{k}) = \int d^3\mathbf{q}\, e^{-i\mathbf{k}\cdot\mathbf{q} - \frac{1}{2}K_i K_j A_{ij}^{dd,\mathrm{nw}}} \left(1 - \frac{1}{2}K_i K_j \Delta A_{ij}^{dd,\mathrm{w}} + \mathcal{O}(k^4\Sigma^4)\right)\left(1 + 2ib_1 K_i U_i^d(q) + \dots\right),$$

where we have used the transformed $K_i = R_{ij}k_j$ to encode redshift-space effects. Given that the no-wiggle spectrum reproduces the broadband scale dependence of the linear theory power spectrum, we can think of the no-wiggle exponential as resumming the non-BAO component of large scale bulk flows. Since the wiggle component contributes negligibly to the displacement power in the perturbative limit, keeping only one power of the wiggle power spectrum in our calculations serves to distinguish the effect of the IR bulk flows from BAO phenomena. The two-point functions entering the bias terms can likewise be split into no-wiggle and wiggle pieces, e.g. $U(q) = U^{\mathrm{nw}} + \Delta U^{\mathrm{w}}$, where again, roughly speaking, the former will contribute only to the broadband power while the latter will give rise to oscillatory behavior. Keeping the above expression to order[13] $\mathcal{O}(k^2\Sigma^2)$, and discarding terms that don't

---

[12]Taking the nonlinear scale to be given by $k_{\mathrm{nl}}^2\Sigma^2(z) \sim 1$, we have $k_{\mathrm{nl}} \propto D^{-1}(z)$, such that the peak magnitude of $k_i k_j \Delta A_{ij}^{\mathrm{w}}$ is roughly in the few tenths of a percent range for our reference cosmology independent of redshift.

[13]Note, however, that terms involving more powers of the wiggle displacement will be more suppressed than those involving no-wiggle displacements.

contain any no-wiggle pieces, we then have

$$P^{dd}(\mathbf{k}) \ni \int d^3\mathbf{q} \; e^{-i\mathbf{k}\cdot\mathbf{q}-\frac{1}{2}K_iK_jA_{ij}^{dd,\mathrm{nw}}} \left(-\frac{1}{2}K_iK_j\Delta A_{ij}^{dd,\mathrm{w}} + 2ib_1K_i\Delta U_i^{d,\mathrm{w}}(q) + b_1^2\xi_L^{\mathrm{w}}(q) + \dots \right)$$

$$\approx e^{-\frac{1}{2}K_iK_j\bar{A}_{ij}^{dd,\mathrm{nw}}} \int d^3\mathbf{q} \; e^{-i\mathbf{k}\cdot\mathbf{q}} \left(-\frac{1}{2}K_iK_j\Delta A_{ij}^{dd,\mathrm{w}} + 2ib_1K_i\Delta U_i^{d,\mathrm{w}}(q) + b_1^2\xi_L^{\mathrm{w}}(q) + \dots \right)$$

where in the final line we have used the fact that the wiggle contributions will be confined in support around the BAO scale ($q \sim 100\,\mathrm{Mpc}$) and the non-wiggle pieces vary smoothly at this Lagrangian separation, so we can pull the exponentiated no-wiggle contribution out of the integral as an average. Following ref. [61], we have defined the quantity $\bar{A}_{ij}^{dd,\mathrm{nw}}$ as the "average" of the un-barred quantity over the support of the wiggle component; to zeroth order in the approximation this is equivalent to evaluating A_{ij} at the peak $q_{\max}$ of the support of the wiggle feature. Neglecting any angular effects in $\nu = \hat{q}\cdot\hat{k}$, which will enter at higher order in the wave number, we further have that $\bar{A}_{ij}^{dd,\mathrm{nw}} \simeq (X^{dd,\mathrm{nw}} + \frac{1}{3}Y^{dd,\mathrm{nw}})\delta_{ij}$.[14] Plugging in for the expression of $\mathbf{K} = \mathbf{R}^{\mathbf{T}}\mathbf{k}$, with $K^2 = (1 + f(f+2)\nu^2)k^2$, the wiggle contribution to the power spectrum is then approximately

$$P^{dd}(\mathbf{k})_{\mathrm{wiggle}} \approx e^{-\frac{1}{2}K^2\Sigma_{dd}^2} \left[(\hat{K}\cdot\hat{k})^2 P^{dd,\mathrm{w}}(k) + 2b_1(\hat{K}\cdot\hat{k})P^{\mathrm{dm,w}}(k) + b_1^2 P^{mm,\mathrm{w}}(k) \right]$$

$$= e^{-\frac{1}{2}K^2\Sigma_{dd}^2} \left[(1+f\nu^2)^2(1-\mathcal{S}(k))^2 + 2b_1(1+f\nu^2)(1-\mathcal{S}(k)) + b_1^2 \right] P_{\mathrm{w}}(k)$$

$$= e^{-\frac{1}{2}K^2\Sigma_{dd}^2} \left[(1+f\nu^2)(1-\mathcal{S}(k)) + b_1 \right]^2 P_{\mathrm{w}}(k) \tag{D.2}$$

where in the penultimate equality we have used the definition of the displaced field and defined $\Sigma_{dd}^2 = (X^{dd,\mathrm{nw}} + \frac{1}{3}Y^{dd,\mathrm{nw}})(q_{\max})$ to be evaluated at the peak of the wiggle displacements. This recovers the form of the empirical model in ref. [26] when we take the Eulerian bias to be $b_1^E = 1 + b_1$, and stick to the damping expansion approximation introduced in ref. [61]. Explicit expressions for X and Y are given in equation (4.9). Taking $\mathcal{S} \to 0$ in the above expression gives the unreconstructed power spectrum within this approximation.

We can now derive the analytical form of the reconstructed power spectrum for **Rec-Sym** in this approximation. Explicitly, we have

$$P_{\mathrm{wiggle}}^{ds}(\mathbf{k}) = -e^{-\frac{1}{2}K^2\Sigma_{ds}^2}\Big((1+f\nu^2)(1-\mathcal{S}(k)) + b_1 \Big)(1+f\nu^2)\mathcal{S}(k)P_{\mathrm{w}}(k) \tag{D.3}$$

$$P_{\mathrm{wiggle}}^{ss}(\mathbf{k}) = e^{-\frac{1}{2}K^2\Sigma_{ds}^2}(1+f\nu^2)^2\mathcal{S}(k)^2 P_{\mathrm{w}}(k). \tag{D.4}$$

The Σ_{ab}^2 are defined as in the dd case. These are the same expressions as derived in ref. [26], though we differ on the expressions for the Σ_{ab}^2 that are involved. Our expressions also agree with those in refs. [19, 27] in the limit that $q_{\max} \to \infty$, though we note that this limit doesn't as accurately capture the damping of the feature since it resums the IR displacements at

---

[14]The factor of a third, included also in ref. [61] but not in ref. [26], comes from the angular average $\langle\hat{q}_i\hat{q}_j\rangle = \delta_{ij}/3$. This can be justified by noting that the integral

$$\frac{1}{2}\int d\mu \; e^{ikq\mu - \frac{1}{2}k^2\mu^2Y/2} = e^{-k^2Y/6}j_0(kq) + \mathcal{O}(k^4\Sigma^4) \tag{D.1}$$

We note, however, that this prescription is only approximate; for example, the same integral with an additional factor of μ in the integrand, relevant for the b_1 contribution, would instead yield $\exp[-3k^2Y/10]\,j_1(kq)$ at leading order. In general, bias contributions with more angular dependence will be damped more. This effect is automatically included in the full Zeldovich calculation.

q beyond the BAO scale. Adding the three spectra together, we recover the Kaiser limit as $k \to 0$, with different damping factors entering at different scales via Σ_{ab}^2. Note that in **Rec-Sym** the angular dependence of the damping is identical in each piece and is encoded within the ν dependence of K^2.

The reconstructed power spectrum with **Rec-Iso** requires a few additional modifications. The displaced-displaced auto spectrum is unchanged, and the shifted-shifted auto spectrum can be calculated by setting $f = 0$ in all formulae, as noted in the main body of the text. However, the ds cross spectrum requires more care, since the zero lag pieces do not transform equally. In particular, direct inspection of the exact expression in equation (4.16) shows that we should instead define

$$-\frac{1}{2}k^2\Sigma_{ds,\mathrm{iso}}^2 = -\frac{1}{2}k^2\left[(1 + f(f+2)\nu^2)\Sigma^{(dd)} + \Sigma^{(ss)} - 2(1 + f\nu^2)\left(\tilde{X}^{ds} + \frac{1}{3}\tilde{Y}^{ds}\right)\right]_{q=q_{\max}}. \tag{D.5}$$

Note this expression differs in detail from that in ref. [26]. The cross spectrum is then instead

$$P_{\mathrm{wiggle}}^{ds,\mathrm{iso}}(\mathbf{k}) = -e^{-\frac{1}{2}k^2\Sigma_{ds,\mathrm{iso}}^2}\left[(1 + f\nu^2)(1 - \mathcal{S}(k)) + b_1\right]\mathcal{S}(k)P_{\mathrm{w}}(k), \tag{D.6}$$

where the angular dependence is subsumed into the defintion of $\Sigma_{ds,\mathrm{iso}}$. Unlike **Rec-Sym**, the damping factor in **Rec-Iso** is not captured by a single angular dependence.

We end this section with a discussion of the inclusion of higher bias terms and other corrections. As seen in the main body of the text, higher bias terms b_2 and b_s, incorporated in our Zeldovich calculation, contribute not only to the broadband but also serve to shift and smear the BAO feature itself. It might thus be of interest to extend the above approximation to include also these higher bias contributions. A potential avenue has been highlighted in ref. [26], although an approach closer to our perturbative bias expansion could also be explored.

Finally, the calculation in ref. [26] included a derivative bias, $b_{\nabla^2}\nabla^2\delta$, as a proxy to estimate the contributions of the higher bias operators. These derivative bias terms can easily be included in the above expressions by substituting $b_1 \to b_1 + k^2 b_{\nabla^2}$. However, there is another context in which such a term might arise in which it would differ across the three pieces dd, ds and ss: if the smoothing due to the Σ_{ab}^2's as defined above do not accurately capture the IR bulk flows — for example if the broadband properies of P_{nw} are slightly off — but differ by some perturbatively small $k^2\delta\Sigma_{ab}^2$, the resulting correction could be corrected for by terms of the form $c_{ab}^2 k^2 P^{ab,\mathrm{w}}(k)$, where c_{ab}^2 would constants fit individually to dd, ds and ss. Such corrections are essentially identical to the EFT corrections described in the text for the full Zeldovich calculation.

E Nonlinearities from the Lagrangian to Eulerian mapping

In standard density field reconstruction, each galaxy is shifted by a smoothed displacement field χ evaluated at the galaxy's current *Eulerian* position $\mathbf{x} = \mathbf{q} + \boldsymbol{\Psi}(\mathbf{q})$ (equation (4.2)). In the main body of the text, we worked in the approximation that $\chi(\mathbf{x}) \approx \chi(\mathbf{q})$, with the understanding that nonlinear corrections would be suppressed by the smoothing scale $\sim \boldsymbol{\Psi}/R$. The goal of this appendix is to flesh out this statement by explicitly computing the leading order corrections to the reconstructed matter power spectrum due to the mapping nonlinearity in real space. For the sake of brevity we will defer the effects of other nonlinearities, such as those arising from dynamics or from translating between displacements and densities, to

future work. Earlier treatments of such effects in Eulerian perturbation theory can be found in refs. [17, 21].

Assuming that the shift field $\chi(\mathbf{q})$ defined in Lagrangian space is Gaussian, the displaced field with mapping nonlinearities unsuppressed is given by

$$\tilde{\boldsymbol{\Psi}}_i^d(\mathbf{q}) = \boldsymbol{\Psi}_i^d(\mathbf{q}) + \boldsymbol{\Psi}_n \partial_n \chi_i(\mathbf{q}) + \frac{1}{2} \boldsymbol{\Psi}_n \boldsymbol{\Psi}_m \partial_n \partial_m \chi_i(\mathbf{q}) + \ldots \equiv \boldsymbol{\Psi}_i^d + \boldsymbol{\Psi}_i^{(d,2)} + \boldsymbol{\Psi}_i^{(d,3)} + \ldots \quad \text{(E.1)}$$

where we have kept the convention used in the main text to refer to the linear piece as $\boldsymbol{\Psi}^d = (1 - \mathcal{S}) * \boldsymbol{\Psi}$, referring to the nonlinear field as $\tilde{\boldsymbol{\Psi}}^d = \boldsymbol{\Psi}(\mathbf{q}) + \chi(\mathbf{q} + \boldsymbol{\Psi})$. For the remainder of this appendix we will focus on corrections due to $\boldsymbol{\Psi}^{(d,2)}$.[15]

From the above, we can write the nonlinear displaced-displaced autospectrum as

$$P^{dd}(k) - P_{\text{Zel}}^{dd}(k) = \int d^3\mathbf{q}\, e^{i\mathbf{k}\cdot\mathbf{q} - \frac{1}{2}k_i k_j A_{ij}^{dd}} \left(\exp\left[-\frac{1}{2} k_i k_j A_{ij}^{dd,1-\text{loop}} - \frac{i}{6} k_i k_j k_k W_{ijk}^{dd} \right] - 1 \right) + \mathcal{O}(P_L^3)$$

where we have defined

$$A_{ij}^{dd,1-\text{loop}} = \left\langle \Delta_i^{(dd,2)} \Delta_j^{(dd,2)} \right\rangle_c + 2 \left\langle \Delta_i^{(dd,1)} \Delta_j^{(dd,3)} \right\rangle_c$$
$$W_{ijk}^{dd} = \left\langle \Delta_i^{(dd,1)} \Delta_j^{(dd,1)} \Delta_k^{(dd,2)} \right\rangle_c + (121) + (211). \quad \text{(E.3)}$$

as in the case of the nonlinear matter power spectrum (e.g. [38]). To calculate these we note that

$$\langle \boldsymbol{\Psi}_i^{(d,2)}(\mathbf{q}_2) \boldsymbol{\Psi}_j^{(d,2)}(\mathbf{q}_1) \rangle_c = \left\langle \left(\boldsymbol{\Psi}_n \partial_n \chi_i \right)(\mathbf{q}_2) \left(\boldsymbol{\Psi}_m \partial_m \chi_j \right)(\mathbf{q}_1) \right\rangle_c$$
$$= \langle \boldsymbol{\Psi}_n(\mathbf{q}_2) \boldsymbol{\Psi}_m(\mathbf{q}_1) \rangle \langle \partial_n \chi_i(\mathbf{q}_2) \partial_m \chi_j(\mathbf{q}_1) \rangle + \langle \boldsymbol{\Psi}_n(\mathbf{q}_2) \partial_m \chi_j(\mathbf{q}_1) \rangle \langle \boldsymbol{\Psi}_m(\mathbf{q}_1) \partial_n \chi_i(\mathbf{q}_2) \rangle$$

and

$$W_{ijk}^{dd,112} = 2 \left\langle \boldsymbol{\Psi}_{1,i}^d \boldsymbol{\Psi}_{2,n} - \boldsymbol{\Psi}_i^d \boldsymbol{\Psi}_n \right\rangle \left\langle \boldsymbol{\Psi}_{1,j}^d \partial_n \chi_{2,k} \right\rangle + \left(i \leftrightarrow j \right) \quad \text{(E.4)}$$

where numerical subscripts refer to coordinates $\mathbf{q}_{1,2}$ as usual.

The mapping corrections to the cross spectrum P^{ds} can be similarly calculated. In this case we need the displacement correlators

$$A_{ij}^{ds,22} = \left\langle \boldsymbol{\Psi}_i^{(d,2)} \boldsymbol{\Psi}_j^{(d,2)} \right\rangle = \langle \boldsymbol{\Psi}_n \boldsymbol{\Psi}_m \rangle \langle \partial_n \chi_i \partial_m \chi_j \rangle \quad \text{(E.5)}$$

where all expectation values are evaluated at a point since $\boldsymbol{\Psi}^s$ receives no nonlinear corrections from the Eulerian-Lagrangian mapping and similarly

$$W_{ijk}^{ds,112} = \left\langle \boldsymbol{\Psi}_{i,1}^s \boldsymbol{\Psi}_{j,1}^s \boldsymbol{\Psi}_{k,2}^{(d,2)} \right\rangle - \left\langle \boldsymbol{\Psi}_{i,2}^d \boldsymbol{\Psi}_{j,1}^s \boldsymbol{\Psi}_{k,2}^{(d,2)} \right\rangle - \left\langle \boldsymbol{\Psi}_{i,1}^s \boldsymbol{\Psi}_{j,2}^d \boldsymbol{\Psi}_{k,2}^{(d,2)} \right\rangle$$
$$= \left(\left\langle \boldsymbol{\Psi}_{1,i}^s \boldsymbol{\Psi}_{2,n} \right\rangle - \left\langle \boldsymbol{\Psi}_i^d \boldsymbol{\Psi}_n \right\rangle \right) \left\langle \boldsymbol{\Psi}_{1,j}^s \partial_n \chi_{2,k} \right\rangle + \left(i \leftrightarrow j \right) \quad \text{(E.6)}$$

to one loop order.

<hr/>

[15] At one loop order all corrections due to $\boldsymbol{\Psi}^{(d,3)}$ are degenerate with the counterterms in our model. To see this, note that such corrections contractions with linear displacements, e.g.

$$\langle \boldsymbol{\Psi}_{1,i}^a \boldsymbol{\Psi}_{2,j}^{(d,3)} \rangle = \langle \boldsymbol{\Psi}_m \partial_m \partial_n \chi_j \rangle \langle \boldsymbol{\Psi}_{1,i}^a \boldsymbol{\Psi}_{2,n} \rangle + \frac{1}{2} \langle \boldsymbol{\Psi}_n \boldsymbol{\Psi}_m \rangle \langle \boldsymbol{\Psi}_{1,i}^a \partial_n \partial_m \chi_{2,j} \rangle, \ a = d, s. \quad \text{(E.2)}$$

Multiplied by the appropriate factor of $-\frac{1}{2} k_i k_j$, the two pieces on the right hand side Fourier transform into $\sim k^2 P_L^{am}(k)$ and $\sim k^2 P_L^{as}(k)$, respectively, thus taking the form of our counterterms $\sim k^2 P^{ab}(k)$.
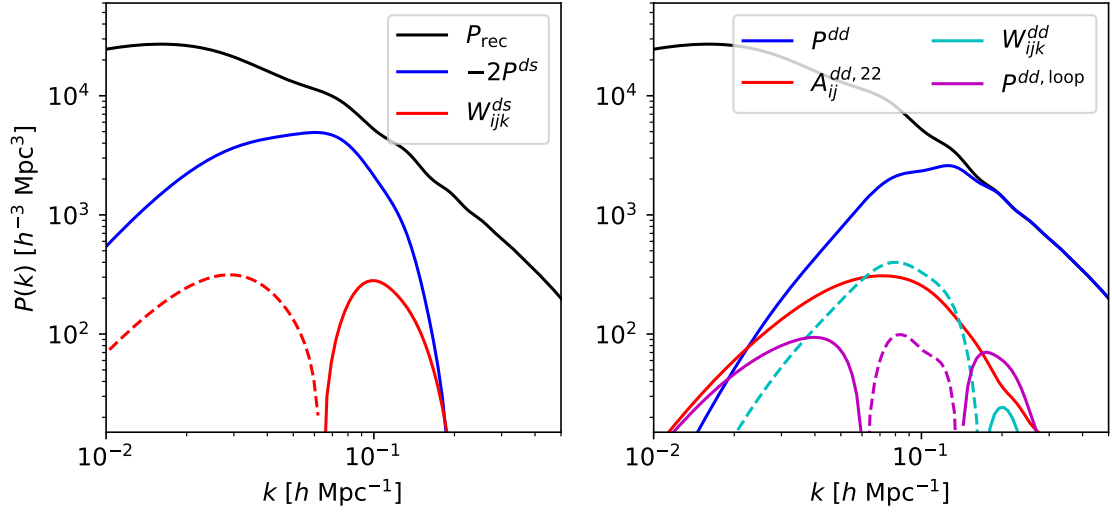
Figure 15. Nonlinear corrections to the reconstructed matter power spectrum due to the Lagrangian-to-Eulerian mapping at one loop order, for $z = 0$ and $R = 15 \ h^{-1}$ Mpc. The left and right panels show contributions to the ds and dd power spectra, respectively. Even for the worst case of $z = 0$, the corrections are never more than a few percent of the total reconstructed power spectrum, though they can become larger than the constituent dd, ds spectra at large or small scales.

As might have been expected, the mapping corrections above all take the form of products of displacement two point functions and their derivatives. Roughly speaking these corrections each have amplitudes given by powers of the Zeldovich displacement Σ^2 and wavenumber k capped at R^{-1} by the smoothing filter; we can thus expect the corrections to enter at order $(\Sigma/R)^4$. For a $15h$ Mpc^{-1} filter at $z = 0$ this amounts to a percent-level effect, with smaller effects at higher z. Concretely, these two point functions can be calculated using

$$\langle \partial_n \chi_i(\mathbf{q}_2) \partial_m \chi_j(\mathbf{q}_1) \rangle = \mathcal{A}^{ss}(q)(\delta_{ij}\delta_{nm} + \dots) + \mathcal{B}^{ss}(q)(\hat{q}_i\hat{q}_j\delta_{nm} + \dots) + \mathcal{C}^{ss}(q)\hat{q}_i\hat{q}_j\hat{q}_n\hat{q}_m$$

$$\langle \boldsymbol{\Psi}_n(\mathbf{q}_2) \partial_m \chi_i(\mathbf{q}_1) \rangle = \mathcal{D}^{\rm sm}(q)(\hat{q}_i\delta_{nm} + \dots) + \mathcal{E}^{\rm sm}(q)\hat{q}_i\hat{q}_n\hat{q}_m$$

$$\langle \boldsymbol{\Psi}_n^a(\mathbf{q}_2) \partial_m \chi_i(\mathbf{q}_1) \rangle = \mathcal{D}^{as}(q)(\hat{q}_i\delta_{nm} + \dots) + \mathcal{E}^{as}(q)\hat{q}_i\hat{q}_n\hat{q}_m$$

where the ellipses denote all distinct permutations and the scalar functions are given by

$$\mathcal{A}^{ss}(q) = \frac{1}{105}\int \frac{dk \ k^2}{2\pi^2} \left(7j_0(kq) + 10j_2(kq) + 3j_4(kq)\right)P_L^{ss}(k)$$

$$\mathcal{B}^{ss}(q) = -\frac{1}{7}\int \frac{dk \ k^2}{2\pi^2} \left(j_2(kq) + j_4(kq)\right)P_L^{ss}(k)$$

$$\mathcal{C}^{ss}(q) = \int \frac{dk \ k^2}{2\pi^2} j_4(kq)P_L^{ss}(k)$$

$$\mathcal{D}^{ab}(q) = \frac{1}{5}\int \frac{dk \ k}{2\pi^2} \left(j_1(kq) + j_3(kq)\right)P_L^{ab}(k)$$

$$\mathcal{E}^{ab}(q) = -\int \frac{dk \ k}{2\pi^2} \ j_3(kq)P_L^{ab}(k) \tag{E.7}$$

where we have used the identification $\chi = \boldsymbol{\Psi}^s$. The remaining correlator $\langle \boldsymbol{\Psi}_{2,i}\boldsymbol{\Psi}_{1,j} \rangle$ is simply minus the non-zero lag piece of A_{ij}. Finally, when some or all of the displacement correlators

in each product are contracted at the same point, as for example in the first term in equation (E.4) and equation (E.5), the resulting contribution becomes proportional to $k^2 P_{ab,\mathrm{Zel}}$ and degenerate with the counterterms included in our model. The above corrections from the nonlinear Eulerian-Lagrangian mapping are plotted at $z = 0$ for the smoothing scale $R = 15 h^{-1}$ Mpc in figure 15. As expected, even at $z = 0$ they are never more than a few percent of the total reconstructed power, though interestingly they can become comparable or larger than the Zeldovich P^{dd} and P^{ds} individually on scales where the Zeldovich spectra lose support. We caution that these curves do not include comparable corrections due to bias or dynamical nonlinearities.

We close with some general comments about nonlinearities in reconstruction. Firstly, the mapping corrections enumerated above are not the only ones at one-loop order; by focusing only on corrections due to $\boldsymbol{\Psi}^{(d,2)}$ we have explicitly avoided the (13) contributions due to third-order mapping corrections. Moreover, as this was an exploratory exercise with which to evaluate the magnitude of mapping nonlinearities, we chose not to include the effects of bias, which would require the inclusion of terms such as $\langle \delta_1^2 \boldsymbol{\Psi}_i^{(d,2)} \rangle$, though these will be in general decomposable into components much like those in equation (E.7). Finally, in addition to mapping nonlinearities, by approximating the shift vector χ with the smoothed Zeldovich displacement we have ignored nonlinearities induced by translating between the density field and displacements. We expect these will be of similar importance to the mapping corrections but defer their evaluation for future work, noting that only that both nonlinearities can be trivially reduced by pushing the smoothing scale R deeper into the linear regime. That these effects are expansions in Σ/R distinguishes them from nonlinear bias or dynamics.

References

[1] S.W. Skillman, M.S. Warren, M.J. Turk, R.H. Wechsler, D.E. Holz and P.M. Sutter, *Dark sky simulations: early data release*, arXiv:1407.2600 [INSPIRE].

[2] 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].

[3] 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].

[4] A. Meiksin, M.J. White and J.A. Peacock, *Baryonic signatures in large scale structure*, *Mon. Not. Roy. Astron. Soc.* **304** (1999) 851 [astro-ph/9812214] [INSPIRE].

[5] H.-J. Seo and D.J. Eisenstein, *Baryonic acoustic oscillations in simulated galaxy redshift surveys*, *Astrophys. J.* **633** (2005) 575 [astro-ph/0507338] [INSPIRE].

[6] M.J. White, *Baryon oscillations*, *Astropart. Phys.* **24** (2005) 334 [astro-ph/0507307] [INSPIRE].

[7] H.-J. Seo and D.J. Eisenstein, *Improved forecasts for the baryon acoustic oscillations and cosmological distance scale*, *Astrophys. J.* **665** (2007) 14 [astro-ph/0701079] [INSPIRE].

[8] D.J. Eisenstein, H.-J. Seo and M.J. White, *On the robustness of the acoustic scale in the low-redshift clustering of matter*, *Astrophys. J.* **664** (2007) 660 [astro-ph/0604361] [INSPIRE].

[9] BOSS collaboration, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2617 [arXiv:1607.03155] [INSPIRE].

[10] X. Xu, N. Padmanabhan, D.J. Eisenstein, K.T. Mehta and A.J. Cuesta, *A 2% distance to z = 0.35 by reconstructing Baryon Acoustic Oscillations — II: fitting techniques*, *Mon. Not. Roy. Astron. Soc.* **427** (2012) 2146 [arXiv:1202.0091] [INSPIRE].

[11] M. Vargas-Magaña et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: potential systematics in fitting of baryon acoustic feature*, *Mon. Not. Roy. Astron. Soc.* **445** (2014) 2.

[12] M. Vargas-Magaña et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: theoretical systematics and Baryon Acoustic Oscillations in the galaxy correlation function*, *Mon. Not. Roy. Astron. Soc.* **477** (2018) 1153 [arXiv:1610.03506] [INSPIRE].

[13] 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].

[14] Y. Noh, M. White and N. Padmanabhan, *Reconstructing baryon oscillations*, *Phys. Rev.* **D 80** (2009) 123501 [arXiv:0909.1802] [INSPIRE].

[15] S. Tassev and M. Zaldarriaga, *Towards an optimal reconstruction of baryon oscillations*, *JCAP* **10** (2012) 006 [arXiv:1203.6066] [INSPIRE].

[16] M. White, *Reconstruction within the Zeldovich approximation*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 3822 [arXiv:1504.03677] [INSPIRE].

[17] M. Schmittfull, Y. Feng, F. Beutler, B. Sherwin and M.Y. Chu, *Eulerian BAO reconstructions and N-point statistics*, *Phys. Rev.* **D 92** (2015) 123522 [arXiv:1508.06972] [INSPIRE].

[18] H.-J. Seo, F. Beutler, A.J. Ross and S. Saito, *Modeling the reconstructed BAO in Fourier space*, *Mon. Not. Roy. Astron. Soc.* **460** (2016) 2453 [arXiv:1511.00663] [INSPIRE].

[19] J.D. Cohn, M. White, T.-C. Chang, G. Holder, N. Padmanabhan and O. Doré, *Combining galaxy and 21 cm surveys*, *Mon. Not. Roy. Astron. Soc.* **457** (2016) 2068 [arXiv:1511.07377] [INSPIRE].

[20] M. Schmittfull, T. Baldauf and M. Zaldarriaga, *Iterative initial condition reconstruction*, *Phys. Rev.* **D 96** (2017) 023505 [arXiv:1704.06634] [INSPIRE].

[21] C. Hikage, K. Koyama and A. Heavens, *Perturbation theory for BAO reconstructed fields: one-loop results in the real-space matter density field*, *Phys. Rev.* **D 96** (2017) 043513 [arXiv:1703.07878] [INSPIRE].

[22] Y. Yu, H.-M. Zhu and U.-L. Pen, *Halo nonlinear reconstruction*, *Astrophys. J.* **847** (2017) 110 [arXiv:1703.08301] [INSPIRE].

[23] Y. Feng, U. Seljak and M. Zaldarriaga, *Exploring the posterior surface of the large scale structure reconstruction*, *JCAP* **07** (2018) 043 [arXiv:1804.09687] [INSPIRE].

[24] C. Modi, Y. Feng and U. Seljak, *Cosmological reconstruction from galaxy light: neural network based light-matter connection*, *JCAP* **10** (2018) 028 [arXiv:1805.02247] [INSPIRE].

[25] R. Hada and D.J. Eisenstein, *An iterative reconstruction of cosmological initial density fields*, *Mon. Not. Roy. Astron. Soc.* **478** (2018) 1866 [arXiv:1804.04738] [INSPIRE].

[26] Z. Ding, H.-J. Seo, Z. Vlah, Y. Feng, M. Schmittfull and F. Beutler, *Theoretical systematics of future Baryon Acoustic Oscillation surveys*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 1021 [arXiv:1708.01297] [INSPIRE].

[27] B.D. Sherwin and M. White, *The impact of wrong assumptions in BAO reconstruction*, *JCAP* **02** (2019) 027 [arXiv:1808.04384] [INSPIRE].

[28] Ya. B. Zeldovich, *Gravitational instability: an approximate theory for large density perturbations*, *Astron. Astrophys.* **5** (1970) 84 [INSPIRE].

[29] T. Buchert, *A class of solutions in Newtonian cosmology and the pancake theory*, *Astron. Astrophys.* **223** (1989) 9 [ɪɴSPIRE].

[30] F. Moutarde, J.-M. Alimi, F.R. Bouchet, R. Pellat and A. Ramani, *Precollapse scale invariance in gravitational instability*, *Astrophys. J.* **382** (1991) 377.

[31] E. Hivon, F.R. Bouchet, S. Colombi and R. Juszkiewicz, *Redshift distortions of clustering: a Lagrangian approach*, *Astron. Astrophys.* **298** (1995) 643 [astro-ph/9407049] [ɪɴSPIRE].

[32] 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] [ɪɴSPIRE].

[33] 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] [ɪɴSPIRE].

[34] 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] [ɪɴSPIRE].

[35] 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] [ɪɴSPIRE].

[36] Z. Vlah, E. Castorina and M. White, *The Gaussian streaming model and convolution Lagrangian effective field theory*, *JCAP* **12** (2016) 007 [arXiv:1609.02908] [ɪɴSPIRE].

[37] M. White, *The Zel'dovich approximation*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 3630 [arXiv:1401.5466] [ɪɴSPIRE].

[38] Z. Vlah, M. White and A. Aviles, *A Lagrangian effective field theory*, *JCAP* **09** (2015) 014 [arXiv:1506.05264] [ɪɴSPIRE].

[39] 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] [ɪɴSPIRE].

[40] E. Castorina and M. White, *Beyond the plane-parallel approximation for redshift surveys*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 4403 [arXiv:1709.09730] [ɪɴSPIRE].

[41] E. Castorina and M. White, *The Zeldovich approximation and wide-angle redshift-space distortions*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 741 [arXiv:1803.08185] [ɪɴSPIRE].

[42] L. Wang, B. Reid and M. White, *An analytic model for redshift-space distortions*, *Mon. Not. Roy. Astron. Soc.* **437** (2014) 588 [arXiv:1306.1804] [ɪɴSPIRE].

[43] N. Padmanabhan et al., *A 2 per cent distance to $z = 0.35$ by reconstructing Baryon Acoustic Oscillations — I. Methods and application to the Sloan Digital Sky Survey*, *Mon. Not. Roy. Astron. Soc.* **427** (2012) 2132 [arXiv:1202.0090] [ɪɴSPIRE].

[44] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1 [ɪɴSPIRE].

[45] A. Burden, W.J. Percival and C. Howlett, *Reconstruction in Fourier space*, *Mon. Not. Roy. Astron. Soc.* **453** (2015) 456 [arXiv:1504.02591] [ɪɴSPIRE].

[46] P. Carter, F. Beutler, W.J. Percival, J. DeRose, R.H. Wechsler and C. Zhao, *The impact of the fiducial cosmology assumption on BAO cosmological parameter inference*, arXiv:1906.03035 [ɪɴSPIRE].

[47] N.E. Chisari and A. Pontzen, *Unequal time correlators and the Zel'dovich approximation*, *Phys. Rev.* **D 100** (2019) 023543 [arXiv:1905.02078] [ɪɴSPIRE].

[48] 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] [ɪɴSPIRE].

[49] S.-F. Chen, E. Castorina and M. White, *Biased tracers of two fluids in the Lagrangian picture*, *JCAP* **06** (2019) 006 [arXiv:1903.00437] [ɪɴSPIRE].

[50] Z. Vlah and M. White, *Exploring redshift-space distortions in large-scale structure*, *JCAP* **03** (2019) 007 [arXiv:1812.02775] [ɪɴSPIRE].

[51] X. Xu et al., *A new statistic for analyzing Baryon Acoustic Oscillations*, *Astrophys. J.* **718** (2010) 1224 [arXiv:1001.2324] [ɪɴSPIRE].

[52] M.S. Warren, *2HOT: an improved parallel hashed oct-tree N-body algorithm for cosmological simulation*, ACM, U.S.A. (2013) [arXiv:1310.4502] [ɪɴSPIRE].

[53] P.S. Behroozi, R.H. Wechsler and H.-Y. Wu, *The rockstar phase-space temporal halo finder and the velocity offsets of cluster cores*, *Astrophys. J.* **762** (2013) 109 [arXiv:1110.4372] [ɪɴSPIRE].

[54] BOSS collaboration, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2617 [arXiv:1607.03155] [ɪɴSPIRE].

[55] P. Carter, F. Beutler, W.J. Percival, C. Blake, J. Koda and A.J. Ross, *Low redshift Baryon Acoustic Oscillation measurement from the reconstructed 6-degree field galaxy survey*, *Mon. Not. Roy. Astron. Soc.* **481** (2018) 2371 [arXiv:1803.01746] [ɪɴSPIRE].

[56] J.E. Bautista et al., *The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations at redshift of 0.72 with the DR14 luminous red galaxy sample*, *Astrophys. J.* **863** (2018) 110 [arXiv:1712.08064] [ɪɴSPIRE].

[57] E.A. Kazin et al., *The WiggleZ dark energy survey: improved distance measurements to z = 1 with reconstruction of the baryonic acoustic feature*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 3524 [arXiv:1401.0358] [ɪɴSPIRE].

[58] E. Noda, M. Peloso and M. Pietroni, *A robust BAO extractor*, *JCAP* **08** (2017) 007 [arXiv:1705.01475] [ɪɴSPIRE].

[59] T. Nishimichi, E. Noda, M. Peloso and M. Pietroni, *BAO extractor: bias and redshift space effects*, *JCAP* **01** (2018) 035 [arXiv:1708.00375] [ɪɴSPIRE].

[60] D.J. Eisenstein and W. Hu, *Baryonic features in the matter transfer function*, *Astrophys. J.* **496** (1998) 605 [astro-ph/9709112] [ɪɴSPIRE].

[61] 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] [ɪɴSPIRE].

[62] W.H. Press and P. Schechter, *Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation*, *Astrophys. J.* **187** (1974) 425 [ɪɴSPIRE].

[63] 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] [ɪɴSPIRE].

[64] M.M. Abidi and T. Baldauf, *Cubic halo bias in Eulerian and Lagrangian space*, *JCAP* **07** (2018) 029 [arXiv:1802.07622] [ɪɴSPIRE].

[65] DESI collaboration, *The DESI experiment part I: science, targeting and survey design*, arXiv:1611.00036 [ɪɴSPIRE].

[66] EUCLID collaboration, *Euclid definition study report*, arXiv:1110.3193 [ɪɴSPIRE].

[67] M. Crocce and R. Scoccimarro, *Nonlinear evolution of Baryon Acoustic Oscillations*, *Phys. Rev.* **D 77** (2008) 023533 [arXiv:0704.2783] [ɪɴSPIRE].

[68] Z. Vlah, U. Seljak and T. Baldauf, *Lagrangian perturbation theory at one loop order: successes, failures and improvements*, *Phys. Rev.* **D 91** (2015) 023508 [arXiv:1410.1617] [ɪɴSPIRE].