

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

An analytic model for redshift-space distortions

### Permalink

<https://escholarship.org/uc/item/0x638983>

### Journal

Monthly Notices of the Royal Astronomical Society, 437(1)

### ISSN

0035-8711

### Authors

Wang, Lile

Reid, Beth

White, Martin

### Publication Date

2014

### DOI

10.1093/mnras/stt1916

Peer reviewed

# An analytic model for redshift-space distortions

Lile Wang<sup>1,2\*</sup>, Beth Reid<sup>1</sup>, and Martin White<sup>1,3</sup>

<sup>1</sup> *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

<sup>2</sup> *Department of Physics and Tsinghua Centre for Astrophysics (THCA), Tsinghua University, Beijing 100084, China*

<sup>3</sup> *Departments of Physics and Astronomy, University of California, Berkeley, CA 94720, USA*

Accepted 2012...; Received 2012...; in original form 2012

## ABSTRACT

Understanding the formation and evolution of large-scale structure is a central problem in cosmology and enables precise tests of General Relativity on cosmological scales and constraints on dark energy. An essential ingredient is an accurate description of the pairwise velocities of biased tracers of the matter field. In this paper we compute the first and second moments of the pairwise velocity distribution by extending the Convolution Lagrangian Perturbation theory (CLPT) formalism of Carlson et al. (2012). Our predictions outperform standard perturbation theory calculations in many cases when compared to statistics measured in N-body simulations. We combine the CLPT predictions of real-space clustering and velocity statistics in the Gaussian streaming model of Reid & White (2011) to obtain predictions for the monopole and quadrupole correlation functions accurate to 2 and 4 per cent respectively down to  $< 25h^{-1}$  Mpc for halos hosting the massive galaxies observed by SDSS-III BOSS. We also discuss contours of the 2D correlation function and clustering “wedges”. We generalize the scheme to cross-correlation functions.

**Key words:** gravitation – galaxies: haloes – galaxies: statistics – cosmological parameters – cosmology: large-scale structure of Universe.

## 1 INTRODUCTION

The large-scale structure (LSS) of the Universe, as traced for example by the distribution of galaxies, is the focus of several ongoing and upcoming observational campaigns. In addition to furthering our understanding of the cosmic web these projects seek to investigate fundamental physics, including the properties of the initial conditions, the imprint of (massive) neutrinos becoming non-relativistic, and the behavior of the mysterious dark energy. Improving our theoretical understanding of the LSS will enhance the scientific return of these projects. In particular, a more detailed understanding of the anisotropy in the observed clustering is of great interest, as the imprint of peculiar velocities in redshift survey maps (known as redshift-space distortions: RSD) allows a consistency test in general relativity between the expansion history and growth of perturbations; such tests could provide support for modified gravity theories as an explanation for the observed cosmic expansion. Moreover, a precise understanding of the peculiar velocity induced anisotropy in galaxy clustering would improve our ability to measure

the geometrically induced anisotropy known as the Alcock-Paczynski effect (Alcock & Paczynski 1979), and thus constrain the expansion rate  $H(z)$  directly (for further details, see e.g. Samushia et al. 2011).

In this work we shall investigate an analytic model to predict the two-point function of biased tracers of large-scale structure based on perturbation theory. There is a large literature using perturbative techniques to study RSD (see for example reviews in Hamilton 1998; Bernardeau et al. 2002; Carlson et al. 2009). Standard perturbation theory (SPT) adopts an Eulerian description of fluids, focusing on the velocity field and density contrast (e.g. Peebles 1980 for linear theory and Juszkiewicz 1981; Vishniac 1983; Goroff et al. 1986; Makino et al. 1992; Jain & Bertschinger 1994 for higher orders). On the other hand, Lagrangian perturbation theory performs an expansion in the Lagrangian displacement field (Buchert 1992, 1994; Bouchet et al. 1995). Lagrangian perturbation theory (LPT) and SPT give identical results for the matter power spectrum in real space when expanded to the same order (Matsubara 2008b). However, it is easier to include redshift space distortions in LPT: a time derivative of the original displacement field is simply added in the line-of-sight direction. Furthermore, current theories

\* Email: wll9004@gmail.com

of galaxy formation rely on the cooling of gas within dark matter potential wells to form galaxies. Therefore, like dark matter halos, galaxies are biased tracers of the underlying matter distribution. A local Lagrangian bias model seems to provide a better description of dark matter halo clustering than a local Eulerian bias (e.g. Roth & Porciani 2011; Baldauf et al. 2012; Chan et al. 2012; Wang & Szalay 2012), although additional terms involving the tidal tensor may also become important for high mass halos (Sheth et al. 2012).

Recently Lagrangian perturbation theory was extended by a resummation scheme known as “integrated perturbation theory” (iPT; Matsubara 2008b,a). A key success of iPT is a very accurate description of the redshift-space two-point correlation function of dark matter halos on scales of interest for studying baryon acoustic oscillations (BAO). Unfortunately, the iPT predictions are inaccurate on scales  $20 - 70h^{-1}$  Mpc, even though deviations from linear theory are still only  $\sim 10$  per cent.

A recent paper by Carlson et al. (2012) introduced convolution Lagrangian perturbation theory (CLPT) which improves the iPT method by resumming more terms in the perturbative expansion. CLPT gives dramatically better results on small scales when compared to N-body simulations, particularly for the description of the redshift-space clustering of dark matter. The methodology is easily extendable to compute properties of the pairwise halo velocity distributions that generate redshift-space distortions. The primary purpose of this paper is to examine CLPT’s accuracy in predicting these statistics, in comparison with N-body simulations. We will see that CLPT provides an accurate description of the velocity distributions. Unfortunately, the CLPT predictions for the anisotropy in the two-point correlation function measured by the quadrupole are still inaccurate on the quasi-linear scales of interest (see figure 5 of Carlson et al. 2012). Therefore, in this paper we combine the real-space correlation function and the velocity statistics predicted by CLPT with the non-perturbative approach advocated in Reid & White (2011, the scale-dependent Gaussian streaming model). This model convolves the real-space two-point correlation function with an approximation to the (scale-dependent) velocity distribution functions to predict redshift space clustering.

This paper is structured as follows. In Sections 2 and 3 we provide some analytic prerequisites for evaluating clustering statistics with CLPT. Section 4 contains the primary new calculation in this work – the prediction of pairwise velocity statistics for biased tracers in CLPT. We evaluate both auto- and cross-correlation statistics. In Section 5 we review the Gaussian streaming model, the basis of our final model for the redshift space halo correlation function. In Section 6 we show the numerical evaluation of those perturbative analytic results. Those CLPT results are input into the Gaussian streaming model in Section 6.1.2 and compared with values obtained directly and indirectly by simulations. Section 7 gives the summary of this article.

## 2 REVIEW

Before we present our calculation of the velocity moments in CLPT, let us review some background material to set our notation and conventions.

### 2.1 Background

Throughout this work we will adopt the “plane-parallel” approximation, so that the line-of-sight (LOS) is chosen along a single Cartesian axis ( $\hat{\mathbf{z}}$ ). While wide-angle effects could potentially be important in modern surveys (Pápai & Szapudi 2008), Samushia et al. (2012) have shown that in practice these effects are small given current errors. The redshift-space position  $\mathbf{s}$  of an object differs from its real-space position  $\mathbf{r}$  due to its peculiar velocity,

$$\mathbf{s} = \mathbf{x} + v_z(\mathbf{x})\hat{\mathbf{z}}, \quad (1)$$

where  $v_z(\mathbf{x}) \equiv u_z(\mathbf{x})/(aH)$  is the LOS component of object’s velocity (assumed non-relativistic) in units of the Hubble velocity. In linear theory, the peculiar velocity field is assumed curl-free, and its divergence is sourced by the underlying matter fluctuations:

$$\nabla \cdot \mathbf{v} = -f\delta_m \quad (2)$$

where  $f \equiv d \ln D / d \ln a$  and  $D(a)$  is the growth rate of fluctuations in linear theory. Measurements of two-point clustering as a function of angle with respect to the LOS direction can directly constrain  $f$  times the normalization of matter fluctuations (e.g. Guzzo et al. 2008; Percival & White 2009; White et al. 2009).

In this paper, we will focus on the prediction of the two-point correlation function:

$$\xi(\mathbf{r}) = \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle. \quad (3)$$

In real-space,  $\xi(\mathbf{r}) \equiv \xi(r)$  is only a function of the separation length, while in redshift-space  $\xi(\mathbf{s})$  depends on the cosine of the angle between the pair separation vector and the LOS,  $\mu_s \equiv \hat{\mathbf{s}} \cdot \hat{\mathbf{z}}$ . It is convenient and common to condense the information in  $\xi(\mathbf{s})$  into Legendre polynomial moments (we use  $L_\ell$  for  $\ell$ th order Legendre polynomial to avoid ambiguity):

$$\xi(s, \mu_s) = \sum_{\ell} \xi_{\ell}(s) L_{\ell}(\mu_s). \quad (4)$$

By symmetry, odd  $\ell$  moments vanish. In linear theory, only  $\ell = 0, 2, 4$  contribute; we will focus our model predictions on those moments. In §6.1.2 we shall also look at clustering “wedges” (e.g. Kazin et al. 2012), but these require no further formalism.

Throughout this paper, we adopt the Einstein summation convention and the following convention of Fourier transform and its inverse ( $n$  is the number of dimensions, here usually 1 or 3):

$$\tilde{F}(\mathbf{k}) = \int d^n x F(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}}, \quad F(\mathbf{x}) = \int \frac{d^n k}{(2\pi)^n} \tilde{F}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}}. \quad (5)$$

### 2.2 Integrated Perturbation Theory (iPT): formalism for biased tracers

Lagrangian perturbation theories perform a perturbative expansion in the displacement field  $\Psi = \Psi^{(1)} + \Psi^{(2)} + \Psi^{(3)} + \dots$ . Here,  $\Psi$  relates the Eulerian (final) coordinates  $\mathbf{x}$  and Lagrangian (initial) coordinates  $\mathbf{q}$  of a mass element or discrete tracer object:

$$\mathbf{x}(\mathbf{q}, t) = \mathbf{q} + \Psi(\mathbf{q}, t). \quad (6)$$

The relation between Eulerian and Lagrangian fields of matter density contrast ( $\delta = \rho/\bar{\rho} - 1$ ) is given by

$$[1 + \delta_m(\mathbf{x}, t)] d^3x = [1 + \delta_m(\mathbf{q}, 0)] d^3q = d^3q. \quad (7)$$

Matsubara (2008a,b) laid out the formalism for including redshift-space distortions and non-linear local Lagrangian biasing within Lagrangian perturbation theory. The density contrast of our tracer field in Lagrangian space,  $\delta(\mathbf{q})$ , is related to the underlying Lagrangian matter density fluctuations smoothed on scale  $R$ :

$$1 + \delta(\mathbf{q}) = F[\delta_{m,R}(\mathbf{q})], \quad (8)$$

where  $F(\delta)$  is the bias function. Note that the smoothing scale  $R$  naturally drops out in the final predictions for all statistics of interest in this paper, but is necessary to keep intermediate quantities well-behaved. Thus

$$1 + \delta(\mathbf{x}, t) = \int d^3q F[\delta_{m,R}(\mathbf{q})] \delta^D[\mathbf{x} - \mathbf{q} - \Psi(\mathbf{q}, t)]. \quad (9)$$

After a coordinate transformation  $\{\mathbf{q}_1, \mathbf{q}_2\} \rightarrow \{\mathbf{q} = \mathbf{q}_2 - \mathbf{q}_1, \mathbf{Q} = (\mathbf{q}_1 + \mathbf{q}_2)/2\}$ , and expressing both  $F$  and  $\delta^D$  in Eq. (9) by their Fourier representations, the two-point correlation function in real-space is given by [see also equations (15) through (20) in Carlson et al. (2012)]:

$$1 + \xi(\mathbf{r}) = \int d^3q \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{q}-\mathbf{r})} \int \frac{d\lambda_1}{2\pi} \frac{d\lambda_2}{2\pi} \times \tilde{F}(\lambda_1) \tilde{F}(\lambda_2) \left\langle e^{i(\lambda_1 \delta_1 + \lambda_2 \delta_2 + \mathbf{k}\cdot\Delta)} \right\rangle, \quad (10)$$

where  $\delta_{1,2} = \delta(\mathbf{q}_{1,2})$ ,  $\Delta = \Psi(\mathbf{q}_2) - \Psi(\mathbf{q}_1)$ , and  $\tilde{F}(\lambda)$  is the Fourier transform of  $F(\delta_R)$  with coordinates pair  $\delta_R$  versus  $\lambda$ .

### 3 CLUSTERING AND VELOCITY STATISTICS IN CLPT

In this section we extend the work of Matsubara (2008a) and Carlson et al. (2012) to enable the calculation of moments of the pairwise velocity distribution for tracers that are biased in a local Lagrangian sense. Once the real-space two-point correlation function and pairwise velocity distributions are known, they determine the observed two-point clustering in redshift-space. These are also the key ingredients of the Gaussian streaming model, as we shall discuss further in §5.

#### 3.1 Velocity moments in CLPT: formalism

The relative peculiar velocity between two tracers at Eulerian coordinates  $\mathbf{x}_1$  and  $\mathbf{x}_2$  can be simply expressed in terms of the time derivative of the displacement field  $\Psi$ :

$$\mathbf{u}(\mathbf{x}_2) - \mathbf{u}(\mathbf{x}_1) = a(\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1) = a\dot{\Delta}. \quad (11)$$

In the time-independent approximation to the perturbative kernels where  $\Psi^{(k)} \propto D^k$  for  $D$  the linear growth function (equation (46) of Matsubara 2008b),

$$\dot{\Psi}^{(k)} = kHf\Psi^{(k)}. \quad (12)$$

Thus we have a perturbative expansion for the Cartesian components of  $\mathbf{v}_n$  (adopting the units of Eq. 1) in terms of

the components of  $\Delta_n$

$$\mathbf{v}_n(\mathbf{x}_2) - \mathbf{v}_n(\mathbf{x}_1) = \sum_k kf\Delta_n^{(k)} = \frac{\dot{\Delta}_n}{H}. \quad (13)$$

We follow common practice and define the velocity generating function  $Z(\mathbf{r}, \mathbf{J})$  by

$$Z(\mathbf{r}, \mathbf{J}) = \int d^3q \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{q}-\mathbf{r})} \int \frac{d\lambda_1}{2\pi} \frac{d\lambda_2}{2\pi} \times \tilde{F}(\lambda_1) \tilde{F}(\lambda_2) \left\langle e^{i(\lambda_1 \delta_1 + \lambda_2 \delta_2 + \mathbf{k}\cdot\Delta + \mathbf{J}\cdot\dot{\Delta}/H)} \right\rangle. \quad (14)$$

Note that  $\xi(\mathbf{r}) = Z(\mathbf{r}, 0) - 1$  (Eq. (10)), and derivatives of  $Z$  give the pairwise velocity moments of interest (i.e., the numerators in our Eqs. 45 and 46):

$$\begin{aligned} & \left\langle [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})] \left\{ \prod_{k=1}^p [\mathbf{v}_{i_k}(\mathbf{x} + \mathbf{r}) - \mathbf{v}_{i_k}(\mathbf{x})] \right\} \right\rangle \\ &= \int d^3q \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{q}-\mathbf{r})} \int \frac{d\lambda_1}{2\pi} \frac{d\lambda_2}{2\pi} \\ & \times \tilde{F}(\lambda_1) \tilde{F}(\lambda_2) \left\langle \left( \prod_{k=1}^p \left( \frac{\dot{\Delta}_{i_k}}{H} \right) \right) e^{i(\lambda_1 \delta_1 + \lambda_2 \delta_2 + \mathbf{k}\cdot\Delta)} \right\rangle \\ &= \prod_{k=1}^p \left( -i \frac{\partial}{\partial \mathbf{J}_{i_k}} \right) Z(\mathbf{r}, \mathbf{J}) \Big|_{\mathbf{J}=0} \end{aligned} \quad (15)$$

Here the set  $\{i_k\}$  specifies the Cartesian coordinate direction for each derivative with respect to  $J_{i_k}$ . Before proceeding to evaluate Eq. (15), we generalize the definitions of the functions  $K$ ,  $L$ , and  $M$  in Carlson et al. (2012) to include  $\mathbf{J}$ . These functions are convenient shorthand for intermediate results.

$$X = \lambda_1 \delta_1 + \lambda_2 \delta_2 + \mathbf{k}\cdot\Delta + \mathbf{J}\cdot\dot{\Delta}/H,$$

$$K_{p,\{i_1\dots i_p\}}(\mathbf{q}, \mathbf{k}, \lambda_1, \lambda_2) = \left\langle \left( -i \frac{\partial}{\partial \mathbf{J}_{i_k}} \right)^p e^{iX} \right\rangle \Big|_{\mathbf{J}=0},$$

$$L_{p,\{i_1\dots i_p\}}(\mathbf{q}, \mathbf{k}) = \int \frac{d\lambda_1}{2\pi} \frac{d\lambda_2}{2\pi} K_{p,\{i_1\dots i_p\}}(\mathbf{q}, \mathbf{k}, \lambda_1, \lambda_2),$$

$$M_{p,\{i_1\dots i_p\}}(\mathbf{r}, \mathbf{q}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{q}-\mathbf{r})} L_{p,\{i_1\dots i_p\}}(\mathbf{q}, \mathbf{k}),$$

$$\begin{aligned} & \left\langle [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})] \left\{ \prod_{k=1}^p [\mathbf{v}_{i_k}(\mathbf{x} + \mathbf{r}) - \mathbf{v}_{i_k}(\mathbf{x})] \right\} \right\rangle \\ &= \int d^3q M_{p,\{i_1,\dots,i_p\}}(\mathbf{r}, \mathbf{q}). \end{aligned} \quad (16)$$

The first subscript of these functions indicates the number of derivative terms  $p$ , and the second is a list containing the Cartesian indices of the  $p$  derivatives.

## 4 EVALUATING THE CLPT PREDICTIONS

### 4.1 Evaluating $\xi(r)$ in CLPT

To begin, we review the calculation of the real-space two-point correlation function  $\xi(r)$  in CLPT, first presented in Carlson et al. (2012), building upon the work of Matsubara

(2008a,b). The cumulant expansion theorem,

$$\langle e^{iX} \rangle = \exp \left[ \sum_{N=1}^{\infty} \frac{i^N}{N!} \langle X^N \rangle_c \right], \quad (17)$$

makes the evaluation of  $K$  tractable; here  $\langle X^N \rangle_c$  is the  $N$ th cumulant of the random variable  $X$ . Taylor expanding the exponential on the right hand side of Eq. (17) only for terms that vanish in the limit  $|\mathbf{q}| \rightarrow \infty$  and keeping only terms up to  $O(P_L^2)$ , Carlson et al. (2012), Sec. 4 obtain

$$\begin{aligned} K_0 &= \left\langle e^{iX} \right\rangle \Big|_{\mathbf{J}=0} \\ &= e^{-(1/2)A_{ij}k_i k_j} e^{-(1/2)(\lambda_1^2 + \lambda_2^2)\sigma_R^2} \left\{ 1 - \lambda_1 \lambda_2 \xi_L + \frac{1}{2} \lambda_1^2 \lambda_2^2 \xi_L^2 \right. \\ &\quad - (\lambda_1 + \lambda_2) U_i k_i + \frac{1}{2} (\lambda_1 + \lambda_2)^2 U_i U_j k_i k_j - \frac{i}{6} W_{ijk} k_i k_j k_k \\ &\quad + \lambda_1 \lambda_2 (\lambda_1 + \lambda_2) \xi_L U_i k_i - \frac{i}{2} (\lambda_1 + \lambda_2) A_{ij}^{10} k_i k_j \\ &\quad \left. + \frac{i}{2} (\lambda_1^2 + \lambda_2^2) U_i^{20} k_i - i \lambda_1 \lambda_2 U_i^{11} k_i + O(P_L^3) \right\}. \end{aligned} \quad (18)$$

In Eq. (18), we adopt the following short-hand definitions:

$$\begin{aligned} \langle \delta_1^2 \rangle_c &= \langle \delta_2^2 \rangle_c = \sigma_R^2, \quad \langle \delta_1 \delta_2 \rangle_c = \xi_L(\mathbf{q}), \\ U_i^{mn} &= \langle \delta_1^m \delta_2^n \Delta_i \rangle_c, \quad A_{ij}^{mn} = \langle \delta_1^m \delta_2^n \Delta_i \Delta_j \rangle_c, \\ W_{ijk}^{mn} &= \langle \delta_1^m \delta_2^n \Delta_i \Delta_j \Delta_k \rangle_c, \\ U_i &= U_i^{10}, \quad A_{ij} = A_{ij}^{00}, \quad W_{ijk} = W_{ijk}^{00}. \end{aligned} \quad (19)$$

Subsequently we apply (see Matsubara 2008a; Carlson et al. 2012)

$$\int \frac{d\lambda}{2\pi} \tilde{F}(\lambda) e^{\lambda^2 \sigma_R^2 / 2} (i\lambda)^n = \langle F^{(n)} \rangle, \quad (20)$$

where  $\langle F^{(n)} \rangle$  is the expectation value of the  $n$ th derivative of  $F(\delta_R)$ . This relation enables us to conduct such transformations from integration with respect to  $\lambda$  to bias parameters:

$$\begin{aligned} (\lambda_1 + \lambda_2) &\rightarrow 2\langle F' \rangle, \quad (\lambda_1^2 + \lambda_2^2) \rightarrow 2\langle F'' \rangle, \\ \lambda_1 \lambda_2 &\rightarrow \langle F'^2 \rangle, \quad \lambda_1^2 \lambda_2^2 \rightarrow \langle F''^2 \rangle, \\ \lambda_1 \lambda_2 (\lambda_1 + \lambda_2) &\rightarrow 2\langle F' \rangle \langle F'' \rangle. \end{aligned} \quad (21)$$

We can hence evaluate  $L_0$  analytically:

$$\begin{aligned} L_0 &= e^{-(1/2)A_{ij}k_i k_j} \left\{ 1 + \langle F' \rangle^2 \xi_L + \frac{1}{2} \langle F'' \rangle^2 \xi_L^2 + 2i \langle F' \rangle U_i k_i \right. \\ &\quad - [\langle F'' \rangle + \langle F' \rangle^2] U_i U_j k_i k_j - \frac{i}{6} W_{ijk} k_i k_j k_k - \langle F' \rangle A_{ij}^{10} k_i k_j \\ &\quad + 2i \langle F' \rangle \langle F'' \rangle \xi_L U_i k_i + i \langle F'' \rangle U_i^{20} k_i \\ &\quad \left. - i \langle F' \rangle^2 U_i^{11} k_i + O(P_L^3) \right\}. \end{aligned} \quad (22)$$

The integration with respect to  $\mathbf{k}$  is then conducted, using some basic relations for Gaussian integration (see appendix

C of Carlson et al. 2012), which gives

$$\begin{aligned} M_0 &= \int \frac{d^3 \mathbf{k}}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{q} - \mathbf{r})} L_0(\mathbf{q}, \mathbf{k}) \\ &= \frac{1}{(2\pi)^{3/2} |A|^{1/2}} e^{-(1/2)(A^{-1})_{ij}(q_i - r_i)(q_j - r_j)} \\ &\quad \times \left\{ 1 + \langle F' \rangle^2 \xi_L + \frac{1}{2} \langle F'' \rangle^2 \xi_L^2 - 2\langle F' \rangle U_i g_i + \frac{1}{6} W_{ijk} \Gamma_{ijk} \right. \\ &\quad - [\langle F'' \rangle + \langle F' \rangle^2] U_i U_j G_{ij} - \langle F' \rangle^2 U_i^{11} g_i - \langle F'' \rangle U_i^{20} g_i \\ &\quad \left. - 2\langle F' \rangle \langle F'' \rangle \xi_L U_i g_i - \langle F' \rangle A_{ij}^{10} G_{ij} + O(P_L^3) \right\}. \end{aligned} \quad (23)$$

Here we define

$$\begin{aligned} g_i &= (A^{-1})_{ij}(q_j - r_j), \quad G_{ij} = (A^{-1})_{ij} - g_i g_j, \\ \Gamma_{ijk} &= (A^{-1})_{ij} g_k + (A^{-1})_{ki} g_j + (A^{-1})_{jk} g_i - g_i g_j g_k. \end{aligned} \quad (24)$$

Finally, the desired correlation function is given by the  $\mathbf{q}$  integration of  $M_0$ :

$$1 + \xi(\mathbf{r}) = \int d^3 q M_0(\mathbf{r}, \mathbf{q}). \quad (25)$$

In order to evaluate  $M_0$ , we expand  $U_i^{mn}$ ,  $A_{ij}^{mn}$  and  $W_{ijk}^{mn}$  in (19) with respect to  $\Delta^{(n)}$ , i.e.

$$\begin{aligned} U_i^{mn(p)} &= \langle \delta_1^m \delta_2^n \Delta_i^{(p)} \rangle_c, \quad A_{ij}^{mn(pq)} = \langle \delta_1^m \delta_2^n \Delta_i^{(p)} \Delta_j^{(q)} \rangle_c, \\ W_{ijk}^{mn(pqr)} &= \langle \delta_1^m \delta_2^n \Delta_i^{(p)} \Delta_j^{(q)} \Delta_k^{(r)} \rangle_c, \end{aligned} \quad (26)$$

and up to desired order, we have

$$\begin{aligned} U_i &= U_i^{(1)} + U_i^{(3)} + \dots, \\ U_i^{20} &= U_i^{20(2)} + \dots, \quad U_i^{11} = U_i^{11(2)} + \dots, \\ A_{ij} &= A_{ij}^{(11)} + A_{ij}^{(22)} + A_{ij}^{(13)} + A_{ij}^{(31)} + \dots, \\ A_{ij}^{10} &= A_{ij}^{10(12)} + A_{ij}^{10(21)} + \dots, \\ W_{ijk} &= W_{ijk}^{(112)} + W_{ijk}^{(121)} + W_{ijk}^{(211)} + \dots. \end{aligned} \quad (27)$$

We refer the readers to appendices B through C in Carlson et al. (2012) for details of evaluating those correlators.

## 4.2 The mean pairwise velocity in CLPT

To compute the mean pairwise velocity in CLPT, we first evaluate  $K_{1,n}$ , again making use of the cumulant expansion theorem:

$$\begin{aligned} K_{1,n}(\lambda_1, \lambda_2, \mathbf{k}, \mathbf{q}) &= \\ &= \exp \left[ \sum_{N=0}^{\infty} \frac{i^N}{N!} \langle X^N \rangle_c \right] \left[ \sum_{N=0}^{\infty} \frac{i^N}{N!} \left\langle \frac{\dot{\Delta}_n}{H} X^N \right\rangle_c \right] \Big|_{\mathbf{J}=0}. \end{aligned} \quad (28)$$

Up to the second order of the linear power spectrum [i.e.  $O(P_L^2)$ ], Eq. (28) is recast as

$$\begin{aligned}
 & K_{1,n}(\lambda_1, \lambda_2, \mathbf{k}, \mathbf{q}) \\
 &= f e^{-(1/2)A_{ij}k_i k_j} e^{-(1/2)(\lambda_1^2 + \lambda_2^2)\sigma_R^2} \\
 &\times \left\{ i(\lambda_1 + \lambda_2)\dot{U}_n + ik_i \dot{A}_{in} - \frac{1}{2}(\lambda_1^2 + \lambda_2^2)\dot{U}_n^{20} \right. \\
 &- \lambda_1 \lambda_2 \dot{U}_n^{11} - \frac{1}{2}k_i k_j \dot{W}_{ijn} - (\lambda_1 + \lambda_2)k_i \dot{A}_{in}^{10} \\
 &- i\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)\xi_L \dot{U}_n - i(\lambda_1 + \lambda_2)^2 k_i U_i \dot{U}_n \\
 &\left. - i\lambda_1 \lambda_2 \xi_L k_i \dot{A}_{in} - i(\lambda_1 + \lambda_2)k_i k_j U_i \dot{A}_{in} + O(P_L^3) \right\}. \tag{29}
 \end{aligned}$$

where we define (up to desired order)

$$\begin{aligned}
 \dot{U}_n &= \frac{\langle \delta_1 \dot{\Delta}_n \rangle}{f} = U_n^{(1)} + 3U_n^{(3)} + \dots, \\
 \dot{U}_n^{20} &= \frac{\langle \delta_1^2 \dot{\Delta}_n \rangle}{f} = U_n^{20(2)} + \dots, \\
 \dot{U}_n^{11} &= \frac{\langle \delta_1 \delta_2 \dot{\Delta}_n \rangle}{f} = U_n^{11(2)} + \dots, \\
 \dot{A}_{in} &= \frac{\langle \Delta_i \dot{\Delta}_n \rangle}{f} = A_{in}^{(11)} + 3A_{in}^{(13)} + A_{in}^{(31)} + 2A_{in}^{(22)} + \dots, \\
 \dot{A}_{in}^{10} &= \frac{\langle \delta_1 \Delta_i \dot{\Delta}_n \rangle}{f} = 2A_{in}^{10(12)} + A_{in}^{10(21)} + \dots, \\
 \dot{W}_{ijn} &= \frac{\langle \delta_1 \Delta_i \Delta_j \dot{\Delta}_n \rangle}{f} = 2W_{ijn}^{(112)} + W_{ijn}^{(121)} + W_{ijn}^{(211)} + \dots. \tag{30}
 \end{aligned}$$

Integrate with respect to  $\lambda_1$  and  $\lambda_2$ , we have (see the appendices of Carlson et al. 2012)

$$\begin{aligned}
 L_{1,n} &= \int \frac{d\lambda_1}{2\pi} \frac{d\lambda_2}{2\pi} \tilde{F}(\lambda_1) \tilde{F}(\lambda_2) K_{1,n}(\lambda_1, \lambda_2, \mathbf{k}, \mathbf{q}) \\
 &= f e^{-(1/2)A_{ij}k_i k_j} e^{-(1/2)(\lambda_1^2 + \lambda_2^2)\sigma_R^2} \\
 &\times \left\{ 2\langle F' \rangle \dot{U}_n + ik_i \dot{A}_{in} + \langle F'' \rangle \dot{U}_n^{20} + \langle F' \rangle^2 \dot{U}_n^{11} \right. \\
 &- \frac{1}{2}k_i k_j \dot{W}_{ijn} + 2i\langle F' \rangle k_i \dot{A}_{in}^{10} + 2\langle F' \rangle \langle F'' \rangle \xi_L \dot{U}_n \\
 &+ 2i[\langle F'' \rangle + \langle F' \rangle^2]k_i U_i \dot{U}_n + i\langle F' \rangle^2 \xi_L k_i \dot{A}_{in} \\
 &\left. - 2\langle F' \rangle k_i k_j U_i \dot{A}_{in} + O(P_L^3) \right\}. \tag{31}
 \end{aligned}$$

Then evaluate the integration over  $\mathbf{k}$ , we have

$$\begin{aligned}
 M_{1,n} &= \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{q} - \mathbf{r})} L_{1,n}(\mathbf{q}, \mathbf{k}) \\
 &= \frac{f^2}{(2\pi)^{3/2} |A|^{1/2}} e^{-(1/2)(A^{-1})_{ij}(q_i - r_i)(q_j - r_j)} \\
 &\times \left\{ 2\langle F' \rangle \dot{U}_n - g_i \dot{A}_{in} + \langle F'' \rangle \dot{U}_n^{20} + \langle F' \rangle^2 \dot{U}_n^{11} \right. \\
 &- \frac{1}{2}G_{ij} \dot{W}_{ijn} - 2\langle F' \rangle g_i \dot{A}_{in}^{10} + 2\langle F' \rangle \langle F'' \rangle \xi_L \dot{U}_n \\
 &- 2[\langle F'' \rangle + \langle F' \rangle^2]g_i U_i \dot{U}_n - \langle F' \rangle^2 \xi_L g_i \dot{A}_{in} \\
 &\left. - 2\langle F' \rangle G_{ij} U_i \dot{A}_{in} + O(P_L^3) \right\}, \tag{32}
 \end{aligned}$$

and finally,

$$v_{12,n}(\mathbf{r}) = [1 + \xi(r)]^{-1} \int d^3 q M_{1,n}(\mathbf{r}, \mathbf{q}). \tag{33}$$

Typically  $v_{12,n}$  is projected along the direction of pair separation vector, i.e.  $v_{12} = v_{12,n} \hat{r}_n$ .

### 4.3 The pairwise velocity dispersion in CLPT

The integration kernel for the velocity dispersion tensor is

$$\begin{aligned}
 K_{2,nm}(\lambda_1, \lambda_2, \mathbf{k}, \mathbf{q}) &= \left( -i \frac{\partial}{\partial J_m} \right) \left( -i \frac{\partial}{\partial J_n} \right) \left\langle e^{iX(\mathbf{J})} \right\rangle \Big|_{\mathbf{J} \rightarrow 0} \\
 &= \exp \left[ \sum_{N=0}^{\infty} \frac{i^N}{N!} \langle X^N \rangle_c \right] \left\{ \left[ \sum_{N=0}^{\infty} \frac{i^N}{N!} \langle \dot{\Delta}_n \dot{\Delta}_m X^N \rangle_c \right] \right. \\
 &\left. + \left[ \sum_{N=0}^{\infty} \frac{i^N}{N!} \langle \dot{\Delta}_n X^N \rangle_c \right] \left[ \sum_{M=0}^{\infty} \frac{i^M}{M!} \langle \dot{\Delta}_m X^M \rangle_c \right] \right\}, \tag{34}
 \end{aligned}$$

Expanding Eq. (34) to second order in the linear power spectrum (i.e. to the order of  $O(P_L^2)$ ):

$$\begin{aligned}
 & K_{2,nm}(\lambda_1, \lambda_2, \mathbf{k}, \mathbf{q}) \\
 &= f^2 e^{-(1/2)A_{ij}k_i k_j} e^{-(1/2)(\lambda_1^2 + \lambda_2^2)\sigma_R^2} \\
 &\times \left\{ (\lambda_1 + \lambda_2)^2 \dot{U}_n \dot{U}_m - (\lambda_1 + \lambda_2)(\dot{A}_{in} k_i \dot{U}_m + \dot{A}_{im} k_i \dot{U}_n) \right. \\
 &- \dot{A}_{im} k_i \dot{A}_{jn} k_j + [1 - \lambda_1 \lambda_2 \xi_L - (\lambda_1 + \lambda_2)U_i k_i] \ddot{A}_{nm} \\
 &\left. + i(\lambda_1 + \lambda_2) \ddot{A}_{nm}^{10} + i\ddot{W}_{inm} k_i + O(P_L^3) \right\}, \tag{35}
 \end{aligned}$$

where we define (up to the desired order)

$$\begin{aligned}
 \ddot{A}_{nm} &= \frac{\langle \dot{\Delta}_n \dot{\Delta}_m \rangle}{f^2} = A_{nm}^{(11)} + 3A_{nm}^{(13)} + 3A_{nm}^{(31)} + 4A_{nm}^{(22)}, \\
 \ddot{A}_{10,nm} &= \frac{\langle \delta_1 \dot{\Delta}_n \dot{\Delta}_m \rangle}{f^2} = 2A_{nm}^{10(12)} + 2A_{nm}^{10(21)}, \\
 \ddot{W}_{inm} &= \frac{\langle \delta_1 \Delta_i \dot{\Delta}_n \dot{\Delta}_m \rangle}{f^2} = 2W_{inm}^{(112)} + 2W_{inm}^{(121)} + W_{inm}^{(211)}. \tag{36}
 \end{aligned}$$

Then evaluate the integration with respect to  $\lambda_1, \lambda_2$ , we have

$$\begin{aligned}
 L_{2,nm} &= \int \frac{d\lambda_1}{(2\pi)} \frac{d\lambda_2}{(2\pi)} \tilde{F}(\lambda_1) \tilde{F}(\lambda_2) K_{2,nm}(\lambda_1, \lambda_2, \mathbf{k}, \mathbf{q}) \\
 &= \{ 2[\langle F' \rangle^2 + \langle F'' \rangle] \dot{U}_n \dot{U}_m + 2i\langle F' \rangle (\dot{A}_{in} k_i \dot{U}_m + \dot{A}_{im} k_i \dot{U}_n) \\
 &- \dot{A}_{im} k_i \dot{A}_{jn} k_j + [1 + \langle F' \rangle^2 \xi_L + 2i\langle F' \rangle U_i k_i] \ddot{A}_{nm} \\
 &+ 2\langle F' \rangle \ddot{A}_{nm}^{10} + i\ddot{W}_{inm} k_i + O(P_L^3) \} f^2 e^{-(1/2)A_{ij}k_i k_j}, \tag{37}
 \end{aligned}$$

and then with respect to  $\mathbf{k}$ :

$$\begin{aligned}
 M_{2,nm} &= \int \frac{d^3 k}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{q} - \mathbf{r})} L_{2,nm}(\mathbf{q}, \mathbf{k}) \\
 &= \frac{f^2}{(2\pi)^{3/2} |A|^{1/2}} e^{-(1/2)(A^{-1})_{ij}(q_i - r_i)(q_j - r_j)} \\
 &\times \left\{ 2[\langle F' \rangle^2 + \langle F'' \rangle] \dot{U}_n \dot{U}_m - 2\langle F' \rangle (\dot{A}_{in} g_i \dot{U}_m + \dot{A}_{im} g_i \dot{U}_n) \right. \\
 &- \dot{A}_{im} \dot{A}_{jn} G_{ij} + [1 + \langle F' \rangle^2 \xi_L - 2\langle F' \rangle U_i g_i] \ddot{A}_{nm} \\
 &\left. + 2\langle F' \rangle \ddot{A}_{nm}^{10} - \ddot{W}_{inm} g_i + O(P_L^3) \right\}. \tag{38}
 \end{aligned}$$

Finally,  $\sigma_{12, nm}^2$  can be obtained by

$$\sigma_{12, nm}^2(\mathbf{r}) = [1 + \xi(r)]^{-1} \int d^3q M_{2, nm}(\mathbf{r}, \mathbf{q}). \quad (39)$$

Desired component of pairwise velocity dispersion can be obtained by different components or kinds of contractions of the tensor  $\sigma_{12, nm}^2$ . In order to obtain the velocity dispersion components parallel to and perpendicular to the pairwise separation unit vector  $\hat{r}$ , we project  $\sigma_{12, nm}^2$  into different directions:

$$\sigma_{\parallel}^2 = \sigma_{12, nm}^2 \hat{r}_n \hat{r}_m, \quad \sigma_{\perp}^2 = (\sigma_{12, nm}^2 \delta_{nm}^K - \sigma_{\parallel}^2) / 2. \quad (40)$$

#### 4.4 Cross-correlation of halos with different bias parameters

It is a straightforward generalization of the above to handle cross-correlations between two tracers with different biases. We note that the displacement field  $\Psi$  is identical for all species – the difference is only in their bias parameters. Therefore, in this “cross-correlation” scenario we have different  $\tilde{F}_j$  ( $j = 1$  or  $2$ ) for  $\lambda_1$  and  $\lambda_2$ . Equation (20) is hence recast as:

$$\int \frac{d\lambda_j}{2\pi} \tilde{F}_j(\lambda_j) e^{\lambda_j^2 \sigma_R^2 / 2} (i\lambda_j)^n = \langle F_j^{(n)} \rangle, \quad j = 1, 2. \quad (41)$$

Hence we can adopt a list of transformations for bias parameters to obtain cross-correlation between different species, which can be straightforwardly deduced from Eq. (41):

$$\begin{aligned} (\lambda_1 + \lambda_2) &\rightarrow [\langle F_1' \rangle + \langle F_2' \rangle], \\ (\lambda_1^2 + \lambda_2^2) &\rightarrow [\langle F_1'' \rangle + \langle F_2'' \rangle], \\ \lambda_1 \lambda_2 &\rightarrow \langle F_1' \rangle \langle F_2' \rangle, \\ \lambda_1^2 \lambda_2^2 &\rightarrow \langle F_1'' \rangle \langle F_2'' \rangle, \\ \lambda_1 \lambda_2 (\lambda_1 + \lambda_2) &\rightarrow [\langle F_1' \rangle \langle F_2'' \rangle + \langle F_2' \rangle \langle F_1'' \rangle]. \end{aligned} \quad (42)$$

Those transformations can also be derived and verified by using the symmetry in  $\lambda_1$  and  $\lambda_2$  of the relevant expressions. They can be applied to Eqs. (23), (32), and (38), with  $\langle F_j' \rangle$  and  $\langle F_j'' \rangle$  ( $j = 1, 2$ ) being bias parameters for two different species. It is easy to verify that the cross-correlation expressions reduce to the auto-correlation expressions when  $\langle F_1' \rangle = \langle F_2' \rangle$  and  $\langle F_1'' \rangle = \langle F_2'' \rangle$ .

## 5 THE GAUSSIAN STREAMING MODEL

Both iPT and CLPT have difficulties reproducing the redshift space clustering of biased tracers on small scales. An alternative is the “Gaussian streaming model” introduced in Reid & White (2011), which takes as inputs perturbation theory expressions for the real-space correlation function and the velocity statistics.

The clustering of a population of objects in redshift-space can be related to their underlying real-space clustering and the full pairwise velocity distribution by (Fisher 1995; Scoccimarro 2004)

$$1 + \xi(r_p, r_{\parallel}) = \int_{-\infty}^{\infty} dy [1 + \xi(r)] \mathcal{P}(v_z = r_{\parallel} - y, \mathbf{r}). \quad (43)$$

Here  $r_p$  is the transverse separation in both real and redshift-space,  $r_{\parallel}$  is the LOS pair separation in redshift space, and  $y$  is the LOS separation in real-space, so that  $r^2 = r_p^2 + y^2$ .

Reid & White (2011) showed that even though the true  $\mathcal{P}$  is certainly non-Gaussian, approximating it with a Gaussian provides an accurate description of the redshift space correlation function of massive halos:

$$1 + \xi^s(r_p, r_{\parallel}) = \int \frac{dy}{[2\pi\sigma_{12}^2(r, \mu)]^{1/2}} [1 + \xi(r)] \times \exp \left\{ -\frac{[r_{\parallel} - y - \mu v_{12}(r)]^2}{2\sigma_{12}^2(r, \mu)} \right\}, \quad (44)$$

In the scale-dependent Gaussian streaming model, the Gaussian probability distribution function is centered at  $\mu v_{12}(r)$ , the mean LOS velocity between a pair of tracers as a function of their real space separation:

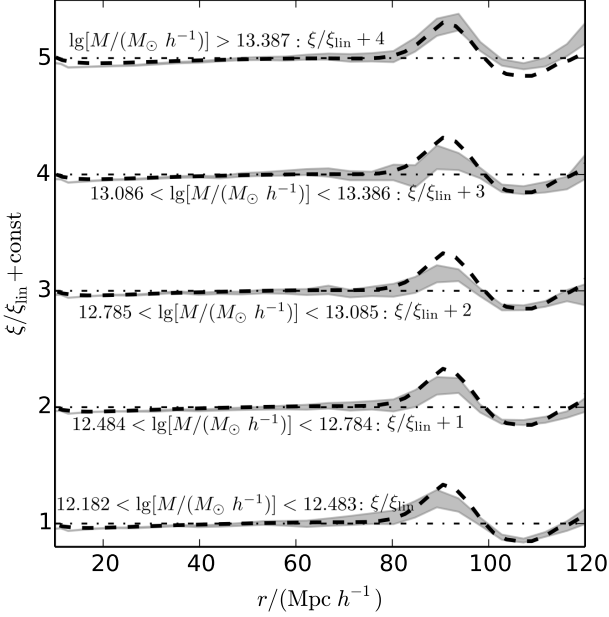
$$v_{12}(r) \hat{\mathbf{r}} = \frac{\langle [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})][\mathbf{v}(\mathbf{x} + \mathbf{r}) - \mathbf{v}(\mathbf{x})] \rangle}{\langle [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})] \rangle} \quad (45)$$

The factor  $[1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})]$  in the numerator and the denominator specifies that we are computing the average relative velocity over pairs of tracers, rather than over randomly chosen points in space. By symmetry, the mean velocity is directed along the pair separation vector; projecting it onto the LOS brings a factor of  $\mu = y/r$  in Eq. (44). Similarly, the width of the velocity PDF is different for components along and perpendicular to the pair separation vector  $\hat{r}$ , so the LOS ( $\hat{z}$ ) velocity dispersion can be decomposed as a sum with contributions from two one-dimensional velocity dispersions.

$$\begin{aligned} \sigma_{12}^2(r, \mu) &= \frac{\langle [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})][v_z(\mathbf{x} + \mathbf{r}) - v_z(\mathbf{x})]^2 \rangle}{\langle [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x} + \mathbf{r})] \rangle} \\ &= \mu^2 \sigma_{\parallel}^2(r) + (1 - \mu^2) \sigma_{\perp}^2(r). \end{aligned} \quad (46)$$

Linear theory expressions for  $v_{12}(r)$  and  $\sigma_{\perp, \parallel}^2(r)$  are given in Fisher (1995); Gorski (1988); Gorski et al. (1989); Reid & White (2011). One finds that the pairwise mean infall velocity,  $v_{12}(r)$ , is proportional to  $bf$ , while  $\sigma_{12}^2(r)$  scales as  $f^2$  with no dependence on the large-scale bias  $b$  at linear order.

Reid & White (2011) evaluated Eqs. (45) and (46) in standard perturbation theory under the assumption of a linear bias  $b$  relating the tracer and matter density fields,  $\delta_t(\mathbf{x}) = b \delta_m(\mathbf{x})$ . There were several shortcomings of this approach however. First, standard perturbation theory does an unsatisfactory job of describing the smoothing of the BAO features in the real-space correlation function. As a result, the analysis of Reid (2012) used iPT to model  $\xi^s(r_p, r_{\parallel})$  above separations of  $70 h^{-1}$  Mpc. Second, the inaccuracy of the streaming model results with standard perturbation theory inputs for the velocity statistics can be traced to inaccuracies in the perturbative calculation of  $v_{12}$  and its derivative,  $dv_{12}/dr$ . This inaccuracy was smallest for halos with second-order bias near zero, which raises the question of whether the source of inaccuracy was the neglect of second order bias terms in the Reid & White (2011) calculation. CLPT naturally includes higher-order bias corrections, and will allow us to quantify the size of the second order contributions to the velocity statistics of interest. For these reasons, we shall consider the combination of CLPT statistics within the Gaussian streaming model ansatz.



**Figure 1.** Real-space correlation function for halos in different mass bins (refer to Table 1). CLPT results are given by heavy dashed curves; simulation results are presented by shaded bands showing the error range. They are divided by linear theory results ( $\xi_{\text{lin}}(r)$ ) to remove the trend, and different mass bins are elevated by different constants (labelled in the figure) to show each more clearly. The bias parameters,  $\langle F' \rangle$  and  $\langle F'' \rangle$ , for the CLPT model are given in Table 1.

## 6 RESULTS

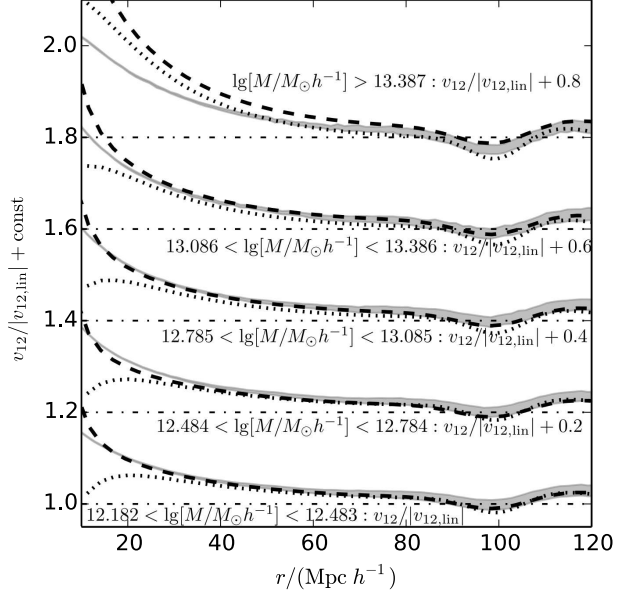
We implemented the formulae above in a C++ code<sup>1</sup>, which numerically evaluates the integrations in Eqs. (25), (33) and (39) for CLPT statistics, and in Eq. (44) for the Gaussian streaming model. In this section we present the results, and compare them with pertinent simulation statistics. The  $N$ -body simulation set used in this work is described in more detail in (Reid & White 2011; White & et al. 2011), in which the halo catalogues are constructed by FoF method. Table 1 lists the halo mass bins we use to compare with our analytic predictions.

### 6.1 Auto-correlation of halos

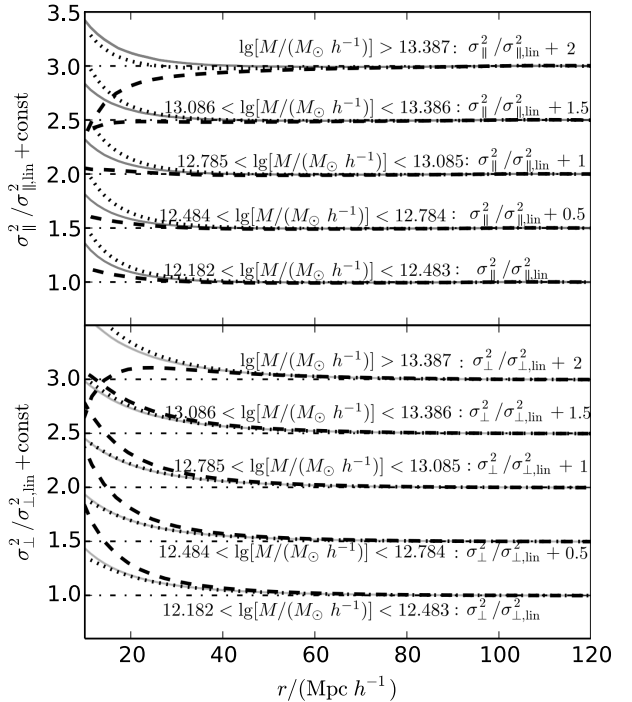
#### 6.1.1 Real-space auto-correlation statistics

We present the CLPT predictions of real-space statistics in this section, which will be used as the “input” of the Gaussian streaming redshift-space distortion model. All calculations are compared with pertinent results in Section 3.

We treat  $\langle F' \rangle$  and  $\langle F'' \rangle$  as free parameters in our model and fit them to the real-space correlation function,  $\xi(r)$ , measured in the  $N$ -body simulations for each halo mass bin. We treat all of the  $\xi(r)$  bins as independent and use the inverse variance obtained from the simulation. While it is incorrect to neglect the correlations, one can see by eye that the resulting best-fit (and the value of  $\chi^2$ ) are entirely reasonable. The resulting values are listed in Table 1. We



**Figure 2.** Pairwise infall velocity of halos in five different mass bins. Our CLPT results are shown by heavy dashed curves and the SPT results (Reid & White 2011) are shown by heavy dotted curves. Both are to be compared with the simulation results shown by shaded bands (indicating the error on the mean of the simulations). All results are divided by the absolute value of linear theory results ( $|v_{12,\text{lin}}(r)|$ ) for better comparison, and are elevated by different constants.



**Figure 3.** Pairwise velocity dispersion for halos in five different mass bins. For clearer presentation we show the values of  $\sigma_{\parallel}^2/\sigma_{\parallel,\text{lin}}^2$  (upper panel) and  $\sigma_{\perp}^2/\sigma_{\perp,\text{lin}}^2$  (lower panel). Shaded bands (very narrow) show the simulation results, the SPT results (Reid & White 2011) are presented by heavy dotted curves, and our theory prediction is presented by dashed curves. Different curves are elevated by different constants.

<sup>1</sup> The code is available at [https://github.com/wl1745881210/CLPT\\_GSRSD.git](https://github.com/wl1745881210/CLPT_GSRSD.git)



**Table 1.** Bias parameters, as well as the uncertainty (denoted by  $\sigma$ ) obtained by fitting the real-space correlation functions for different halo mass bins. Please note that  $\langle F' \rangle$  and  $\langle F'' \rangle$  are both free parameters determined by fitting and  $\sigma$  is the formal error on the fit assuming Gaussian, uncorrelated errors on  $\xi$  as determined from the variance in the simulations.

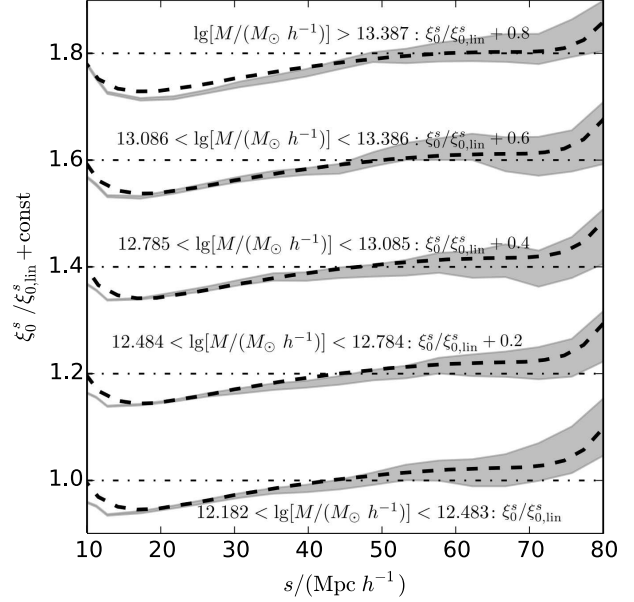
$\lg[M/(M_\odot h^{-1})]$	$\langle F' \rangle$	$\sigma(\langle F' \rangle)$	$\langle F'' \rangle$	$\sigma(\langle F'' \rangle)$
12.182 – 12.483	0.341	0.004	0.06	0.06
12.484 – 12.784	0.435	0.004	0.16	0.06
12.785 – 13.085	0.652	0.005	0.16	0.01
13.086 – 13.386	0.965	0.006	0.14	0.09
> 13.387	1.738	0.007	-0.10	0.11

note here that it is also possible to obtain  $\langle F'' \rangle$  as a function of  $\langle F' \rangle$  using the peak-background split relation (as in Matsubara 2008a). While the relation between our best-fit  $\langle F' \rangle$  and halo mass is close to that obtained from the peak-background split the values of  $\langle F'' \rangle$  can differ significantly. Imposing the peak-background split value of  $\langle F'' \rangle$  has only a modest effect on the shape of the correlation function on the scales of interest however, and does not change our conclusions in any qualitative way.  $\langle F'' \rangle$  is also not well constrained in our fitting (see Table 1 for the uncertainty of  $\langle F'' \rangle$ ), which confirms that  $\langle F'' \rangle$  does not have a considerable impact on the correlation function on the scales of interest.

Fig. 1 compares the real-space correlation function predicted by CLPT with that measured in the simulations. Note that the consistency between CLPT results and simulations is almost perfect from  $\lesssim 10h^{-1}$  Mpc through the BAO scale ( $\sim 110h^{-1}$  Mpc), as also seen in Carlson et al. (2012). The redshift-space correlation function predicted directly from CLPT was presented in Carlson et al. (2012). Here we want to examine the velocity statistics themselves.

We assume that the values of  $\langle F' \rangle$  and  $\langle F'' \rangle$  obtained by fitting the real-space correlation functions are the right ones for evaluating the velocity statistics. Using these values in Eq. (32) and (38), we obtain the scale dependence of the pairwise infall velocity and velocity dispersion. The CLPT results (divided by the linear theory as fiducial values) are compared with simulations in Figs. 2 and 3. The CLPT predictions for  $v_{12}$  are better than the SPT predictions with first order bias presented in Reid & White (2011) for all but the highest mass bin. The CLPT predictions of the pairwise infall velocity statistics can be slightly improved by varying  $\langle F'' \rangle$ , but the prediction of velocity dispersion is quite insensitive to  $\langle F'' \rangle$ .

When comparing the CLPT result to  $\sigma_{\parallel}^2(r)$  and  $\sigma_{\perp}^2(r)$  we add a constant to the predictions so that they take the same value as the simulation at  $r = 130h^{-1}$  Mpc. The constant offsets for  $\sigma_{\parallel}^2(r)$  and  $\sigma_{\perp}^2(r)$  are almost the same, with only  $\sim 1$  per cent relative difference. These two constants are similar to what the authors referred to in Reid & White (2011): the CLPT prediction of absolute value of  $\sigma_{\parallel}^2(r)$  and  $\sigma_{\perp}^2(r)$  is not correct, but a constant shift over the whole range of scales reveals that the CLPT results have correct trend. The possible reason for this is that the velocity dispersion component yielded by gravitational evolution on smallest scales, which should be separated from the overall scale



**Figure 4.** Redshift-space statistics obtained by Gaussian streaming model specified in Reid & White (2011), showing  $\xi_0^s/\xi_{0,\text{lin}}^s$ . The shaded bands present simulation values (showing the error range) and our CLPT values are presented by heavy dashed curve. Each mass bin is elevated by a different constant.

dependence of pairwise velocity dispersion, is not able to be predicted by perturbation theory: this suggests that we should evaluate the constant shift as a fitting parameter. Our CLPT predictions have a similar accuracy to the SPT predictions in Reid & White (2011) for the second highest mass bin. However, it is not clear that CLPT accurately captures the bias-dependence of the deviations from linear theory for  $\sigma_{\parallel}^2(r)$  and  $\sigma_{\perp}^2(r)$ .

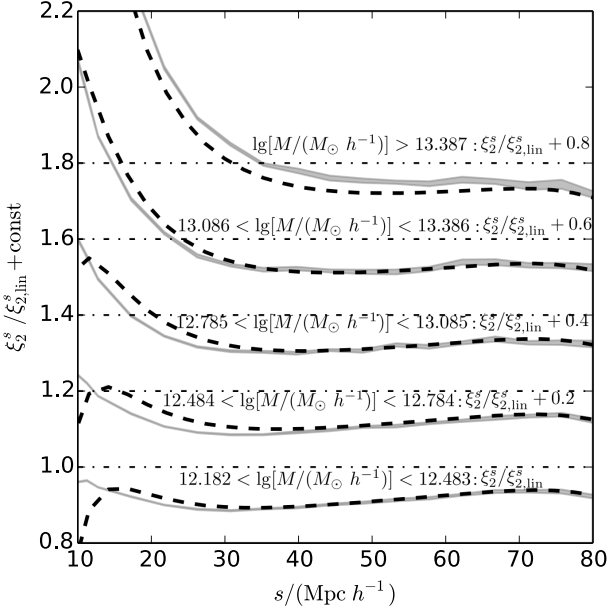
### 6.1.2 Redshift-space distortion for auto-correlation

The redshift-space correlation functions depend on the angle between separation vectors and LOS. This “direction dependency” can be expanded into series with respect to Legendre polynomials (Eq. 4), or, equivalently:

$$\xi_l^s(s) = \frac{2l+1}{2} \int_{-1}^1 d\mu_s L_l(\mu_s) \xi_l^s(s, \mu), \quad (47)$$

where  $s = (r_p^2 + r_{\parallel}^2)^{1/2}$  is the redshift-space distance, and  $\mu_s = r_{\parallel}/s$  is the cosine of the angle between separation vector and LOS. Generally we are most interested in the lowest non-zero moments, i.e.  $l = 0, 2$ . In Figs. 4 and 5, we present the lowest two non-zero multipole moments divided by linear theory results (see Fisher 1995; Reid & White 2011, for the linear theory expressions).

We show the results of the multipole expansion in Figs. 4 and 5. The accuracy of the prediction of CLPT with the Gaussian streaming model is at the several per cent level on scales larger than  $\gtrsim 20h^{-1}$  Mpc, and no worse than  $\sim 10$  per cent even down to  $\sim 10h^{-1}$  Mpc. The agreement remains equally good at BAO scales, but we only show  $10h^{-1}$  Mpc  $< s < 70h^{-1}$  Mpc for clearer presentation at smaller scales. The theory breaks down at  $r \lesssim 10h^{-1}$  Mpc where the correlation function amplitude is approaching  $O(1)$ . For the



**Figure 5.** Redshift-space results of  $\xi_2^s / \xi_{2,\text{lin}}^s$ . Labels and curve indications are identical to Fig. 4.

quadrupole moment ( $\ell = 2$ ), we observe that the model has reasonable overlap with the simulations throughout the scales of general interest ( $10h^{-1} \text{ Mpc} \lesssim s \lesssim 120h^{-1} \text{ Mpc}$ ).

To further isolate the regions where the theory and N-body simulations are in good agreement we additionally examine the “wedge” statistics (e.g. Kazin et al. 2012), defined by

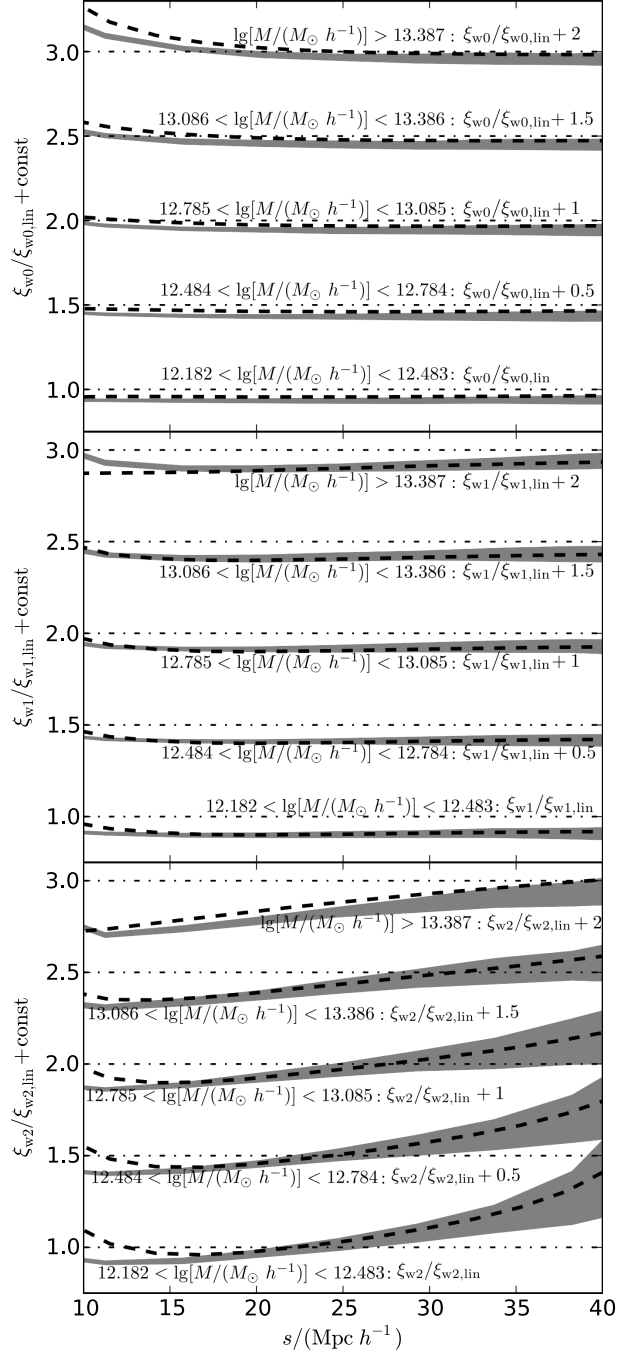
$$\xi_{\text{wedge}}^s(s, \mu_{\text{min}}, \mu_{\text{max}}) = \frac{1}{\Delta\mu} \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} \xi^s(s, \mu) d\mu, \quad (48)$$

where  $\Delta\mu = \mu_{\text{max}} - \mu_{\text{min}}$ . In this paper we use three such “wedges”, which are denoted by  $\xi_{w0} = \xi_{\text{wedge}}^s(s, 0, 1/3)$ ,  $\xi_{w1} = \xi_{\text{wedge}}^s(s, 1/3, 2/3)$  and  $\xi_{w2} = \xi_{\text{wedge}}^s(s, 2/3, 1)$ . The predictions for the  $\xi_{wi}$  are compared to N-body simulations in Fig. 6. Note that the fractional deviations from linear theory are largest on small scales and when  $\mu \simeq 1$ . In addition the inaccuracy of our theoretical prediction for the quadrupolar moment on about  $\sim 10h^{-1} \text{ Mpc}$  can be attributed to the disagreement near  $\mu \simeq 1$  (please note that  $\xi_2^s$  is negative around  $r \sim 10h^{-1} \text{ Mpc}$  but  $\xi_{w2}$  is positive there). On scales above  $20h^{-1} \text{ Mpc}$  our model works well, the difference between the model and N-body results is less than 5 per cent for all three wedges.

In order to provide another view of the disagreement between the model and simulations, we show in Fig. 7 contours of  $\xi$  predicted by the analytic model (dashed contours) and N-body simulations (solid contours) for two bins in halo mass. We can clearly observe that, for the halos in the lower mass bin,  $\xi^s$  is less precisely predicted around  $\mu \simeq 1$ . On larger scales ( $s \gtrsim 20h^{-1} \text{ Mpc}$ ) simulation results are accurately predicted for both mass bins.

## 6.2 Cross correlation between halos and dark matter particles

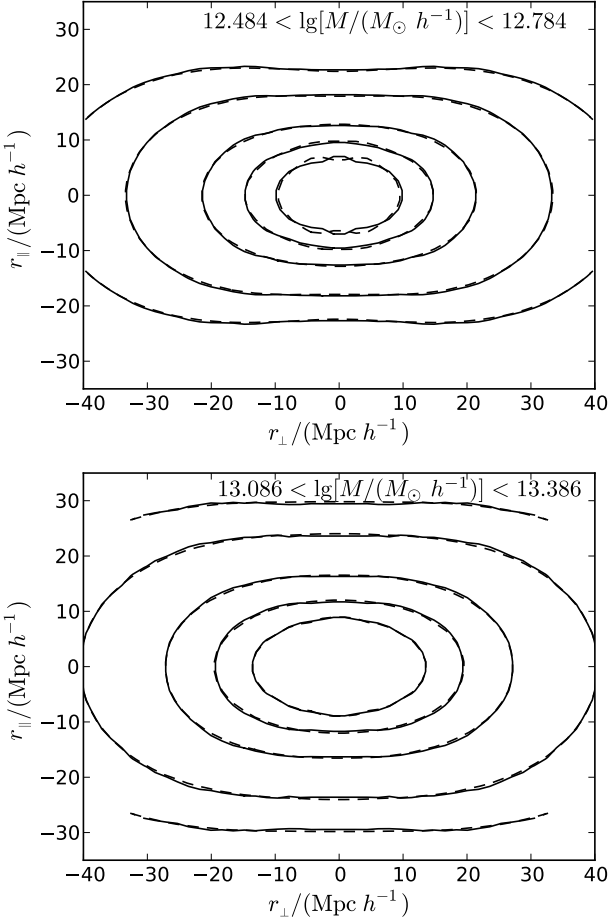
As shown in subsection 4.4, CLPT theory is also capable of making predictions for cross-correlations. Here we compare



**Figure 6.** Wedge statistics showing  $\xi_{w0}$ ,  $\xi_{w1}$  and  $\xi_{w2}$  (see Eq. 48) in redshift-space. The scheme of presentation is similar to Fig. 4.

statistics predicted by CLPT and the Gaussian streaming model with those given by simulations, for cross-correlations of halos with dark matter particles in the simulations. In this subsection we adopt the same bias parameters as in Table 1. We can obtain a better match to the cross-correlation infall velocities by adjusting  $\langle F' \rangle$  and  $\langle F'' \rangle$ , however those values do not provide a good match to the real-space cross-correlation function suggesting either that our bias model is too simple or the improved agreement reflects a breakdown of perturbation theory.

Real-space statistics are presented by Figs. 8 through

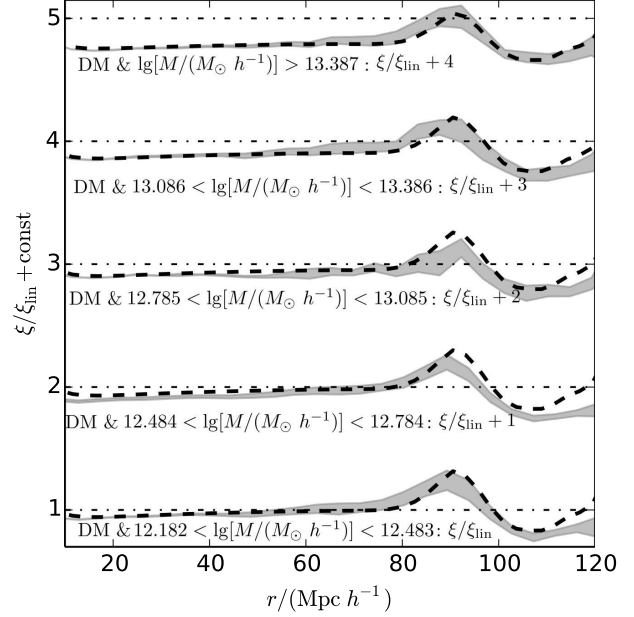


**Figure 7.** Contour plots that compare theoretical predictions (dashed) and simulation results (solid). The contour lines for  $\xi^s = 0.8, 0.4, 0.2, 0.08$ , and  $0.04$  are presented in the figure. We note that the dashed and solid contours are nicely overlapped in most areas, except in the regions that  $s \lesssim 10 h^{-1} \text{Mpc}$  and  $\mu \lesssim 1$ .

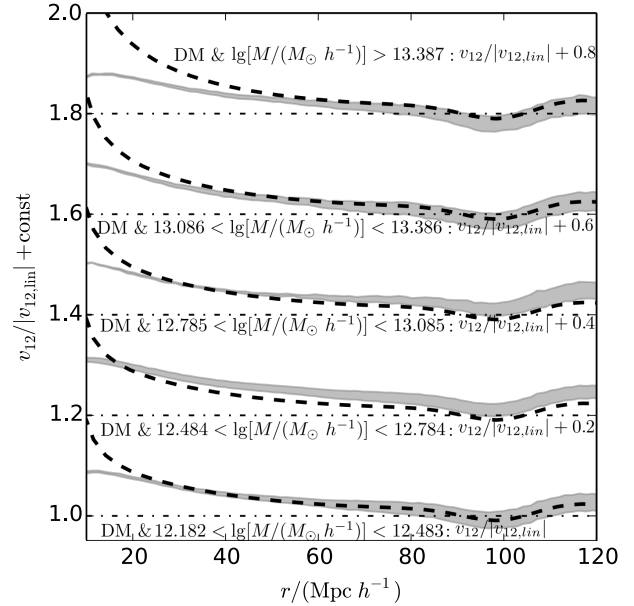
10. Predictions of CLPT for pairwise infall velocity (Fig. 9) in such a cross-correlation case is not as good as the prediction for auto-correlation (Fig. 2), but still satisfactory; the discrepancy is  $\lesssim 5$  per cent throughout the scales of interest. The real-space correlation function,  $\xi(r)$ , and the velocity dispersion,  $\sigma_{12}^2(r)$ , on the other hand, are still accurately predicted by the theory.

Inserting the real-space statistics shown by Figs. 8 through 10 into Eq. (44), we get the redshift-space correlation function, which is also expanded with respect to Legendre polynomials as in Eq. (47). Similar to Section (6.1.2), this section also presents monopole and quadrupole moments in Figs. 11 and 12. While we focus on  $10h^{-1} \text{Mpc} < s < 70h^{-1} \text{Mpc}$ , the agreement remains good on BAO scales. Although the predictions for the velocity statistics are not as good as in the auto-correlation case, the behavior of the multipole moments is still well sketched by CLPT and the Gaussian streaming model (to  $\sim 10$  per cent, even on scales of  $\sim 10h^{-1} \text{Mpc}$ ). The manner in which  $\xi_0^s/\xi_{0,\text{lin}}^s$  and  $\xi_2^s/\xi_{2,\text{lin}}^s$  vary with  $s$  is still correct to quite small  $s$ .

Similar procedures also produce predictions for halo-halo cross correlations. In Fig. 13 we compare the statistics as cross-correlations between halos in two differ-

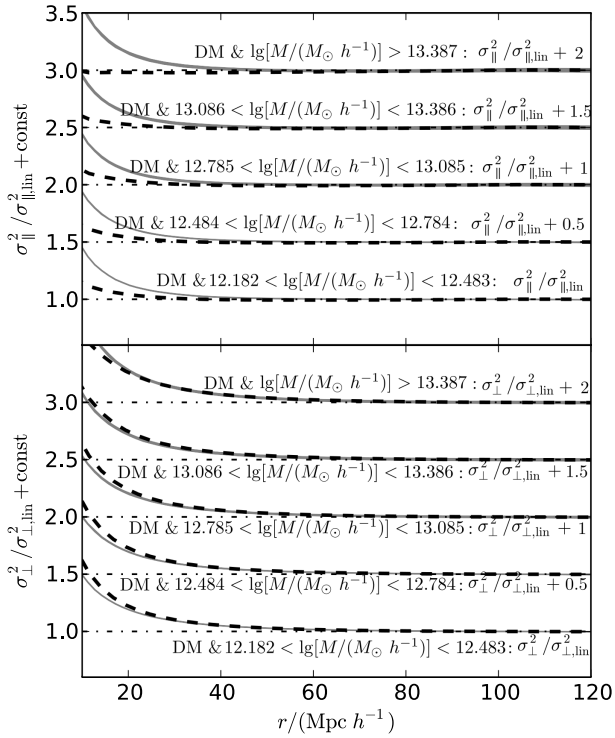


**Figure 8.** Real-space cross-correlation function between dark matter and halos (in five different mass bins; each mass bin is elevated by a specific constant). The scheme of presentation is similar to Fig. 1.

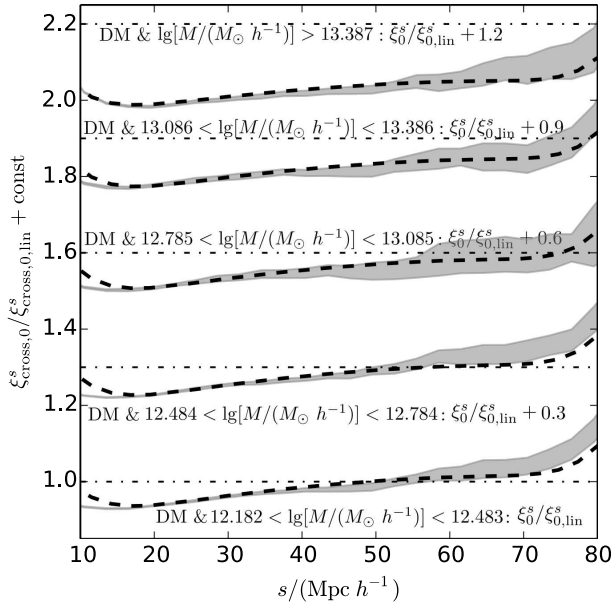


**Figure 9.** Pairwise infall velocity as cross-correlations. Our CLPT results are shown by a heavy solid curve, compared with simulations presented by shaded bands showing the error range. The scheme of presentation is similar to Fig. 2.

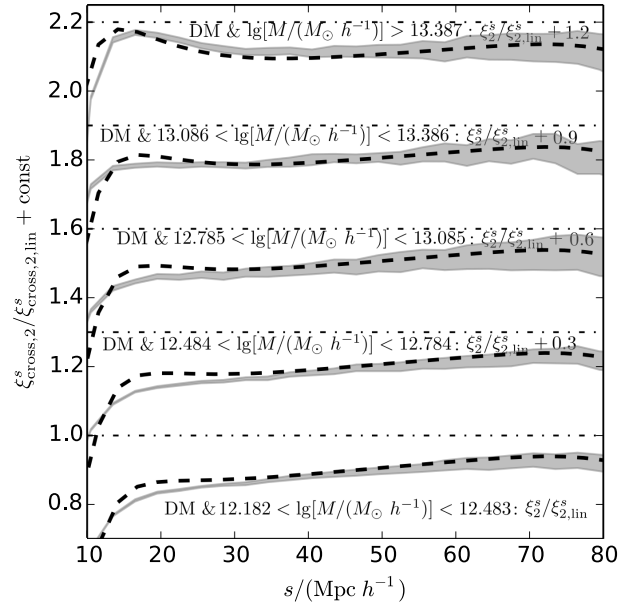
ent mass bins:  $12.182 < \lg[M/(M_\odot h^{-1})] < 12.483$  and  $\lg[M/(M_\odot h^{-1})] > 13.387$ . There are no adjustable parameters in this comparison, because the values of  $\langle F' \rangle$  and  $\langle F'' \rangle$  are fixed by the auto-correlations. Good agreement between our theoretical model and the simulations is still observed, even down to the scale of  $s \gtrsim 10h^{-1} \text{Mpc}$ .



**Figure 10.** Pairwise velocity dispersion as cross-correlations. Similar to Fig. 3, we also show the values of  $\sigma_{\parallel}^2/\sigma_{\parallel,lin}^2$  (upper panel) and  $\sigma_{\perp}^2/\sigma_{\perp,lin}^2$  (lower panel), with shaded bands (simulation results) and dashed curves (theoretical predictions). Different mass bins are elevated by different constants.



**Figure 11.** Monopole moment of redshift-space cross-correlation function between halos and dark matter. Results are divided by fiducial linear theory results (i.e.  $\xi_{cross,0}^s/\xi_{cross,0,lin}^s$  is presented). The scheme of presentation is similar to Fig. 4.



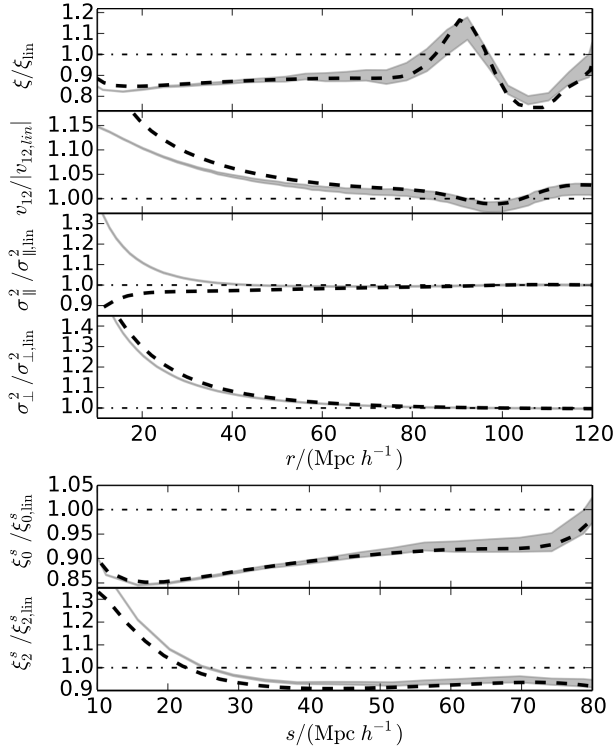
**Figure 12.** Quadrupole moment of redshift-space cross-correlation function between halos and dark matter (linear theory results as fiducial value)  $\xi_{cross,2}^s/\xi_{cross,2,lin}^s$ . The scheme of presentation is similar to Fig. 11.

## 7 DISCUSSION AND SUMMARY

By introducing an auxiliary term  $\mathbf{J}$  in the generating function, we generalize the CLPT scheme elaborated in Carlson et al. (2012) to estimate the pairwise infall velocity and velocity dispersion as functions of pair separation. This allows a self-consistent calculation of these statistics for biased tracers, including scale-dependent or higher-order bias terms. Indeed we find that CLPT gives better estimates for the magnitude of the pairwise infall velocity,  $v_{12}(r)$ , than to quasi-linear theory with a “linear” bias (Reid & White 2011) for a wide range of halo masses.

The  $\xi(r)$ ,  $v_{12}(r)$  and  $\sigma_{12}^2(r)$  predicted by CLPT can be used as inputs to Gaussian streaming model (Eq. 44) to obtain predictions for the redshift-space correlation function of halos. For the monopole and quadrupole moments of the correlation function the agreement between theory and N-body simulation is at the few per cent level down to  $\sim 15h^{-1}$  Mpc, and  $\sim 10$  per cent at  $\sim 10h^{-1}$  Mpc. We infer that the Gaussian streaming model of redshift-space distortion is not sensitively affected by  $\sigma_{12}^2(r)$ , but the small scale statistics are enhanced by better estimations of  $v_{12}(r)$ , compared with semi-linear results in Reid & White (2011). We attribute the enhanced results to our inclusion of higher order (one-loop) terms and the resummation scheme employed in CLPT.

It is worth noting that the argument between the CLPT-Gaussian streaming quadrupolar moment and N-body simulation (e.g. Fig. 5) is considerably better than the “original” SPT scheme in Reid & White (2011). From Figs. 6 and 7 we observe that the theoretical predictions are not sufficiently accurate only around the region where  $s \lesssim 20h^{-1}$  Mpc and  $\mu \simeq 1$  (or  $r_p \simeq 0$ ) for the lower mass bins. It was shown in Reid & White (2011) (e.g. their figure 6) that the Gaussian streaming model predicted the redshift-space correlation function well when accurate “inputs” (i.e.



**Figure 13.** Statistic functions as cross-correlations between halos in two different mass bins:  $12.182 < \lg[M/(M_{\odot} h^{-1})] < 12.483$  and  $\lg[M/(M_{\odot} h^{-1})] > 13.387$ . The curves, lines and shaded bands in the panels have similar indications to Figs. 1 through 5, which compare our theoretical predictions with simulations. From top to bottom: real-space correlation function (first panel); pairwise infall velocity (second panel); velocity dispersion parallel (third panel) and perpendicular (fourth panel) to separation vector; monopole (fifth panel) and quadrupole (sixth panel) moment of redshift-space correlation function.

$v_{12}(r)$  and  $\sigma_{12}^2(r)$  were used. Prediction of these inputs using CLPT seems a reliable way of computing redshift-space statistics for tracers with a local Lagrangian bias.

We also extended the CLPT-Gaussian streaming model to cross-correlations between differently biased tracers. As an example, we modelled the monopole and quadrupole moments of the redshift-space cross-correlation function between halos and dark matter, and between halos in different mass bins. The agreement with N-body simulations for  $v_{12}(r)$  was not as good as in the auto-correlation case, as expected, but the distortions were still accurately revealed in monopole and quadrupole moments. This is not unexpected: it was already noted by Reid & White (2011) that the Gaussian approximation for the velocity PDF worked much better for halos in simulations than for the dark matter particles themselves.

## REFERENCES

- Alcock C., Paczynski B., 1979, *Nature*, 281, 358
- Baldauf T., Seljak U., Desjacques V., McDonald P., 2012, ArXiv e-prints
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Phys.Rep.*, 367, 1
- Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, *A&A*, 296, 575
- Buchert T., 1992, *MNRAS*, 254, 729
- Buchert T., 1994, *MNRAS*, 267, 811
- Carlson J., Reid B., White M., 2012, ArXiv e-prints
- Carlson J., White M., Padmanabhan N., 2009, *Phys Rev D*, 80, 043531
- Chan K. C., Scoccimarro R., Sheth R. K., 2012, *Phys Rev D*, 85, 083509
- Fisher K. B., 1995, *ApJ*, 448, 494
- Goroff M. H., Grinstein B., Rey S.-J., Wise M. B., 1986, *ApJ*, 311, 6
- Gorski K., 1988, *ApJL*, 332, L7
- Gorski K. M., Davis M., Strauss M. A., White S. D. M., Yahil A., 1989, *ApJ*, 344, 1
- Guzzo L., Pierleoni M., Meneux B., et al. 2008, *Nature*, 451, 541
- Hamilton A. J. S., 1998, in Hamilton D., ed., *The Evolving Universe Vol. 231 of Astrophysics and Space Science Library, Linear Redshift Distortions: a Review.* p. 185
- Jain B., Bertschinger E., 1994, *ApJ*, 431, 495
- Juszkiewicz R., 1981, *MNRAS*, 197, 931
- Kazin E. A., Sánchez A. G., Blanton M. R., 2012, *MNRAS*, 419, 3223
- Makino N., Sasaki M., Suto Y., 1992, *Phys Rev D*, 46, 585
- Matsubara T., 2008a, *Phys Rev D*, 78, 083519
- Matsubara T., 2008b, *Phys Rev D*, 77, 063530
- Pápai P., Szapudi I., 2008, *MNRAS*, 389, 292
- Peebles P. J. E., 1980, *The large-scale structure of the universe.* Princeton University Press, Princeton, NJ
- Percival W. J., White M., 2009, *MNRAS*, 393, 297
- Reid B. A., White M., 2011, *MNRAS*, 417, 1913
- Reid B. A. et al., 2012, *MNRAS*, 426, 2719
- Roth N., Porciani C., 2011, *MNRAS*, 415, 829
- Samushia L., Percival W. J., Guzzo L., Wang Y., Cimatti A., Baugh C., Geach J. E., Lacey C., Majerotto E., Mukherjee P., Orsi A., 2011, *MNRAS*, 410, 1993
- Samushia L., Percival W. J., Raccanelli A., 2012, *MNRAS*, 420, 2102
- Scoccimarro R., 2004, *Phys. Rev. D*, 70, 083007
- Sheth R. K., Chuen Chan K., Scoccimarro R., 2012, ArXiv e-prints
- Vishniac E. T., 1983, *MNRAS*, 203, 345
- Wang X., Szalay A., 2012, *Phys Rev D*, 86, 043508
- White M., et al. 2011, *ApJ*, 728, 126
- White M., Song Y.-S., Percival W. J., 2009, *MNRAS*, 397, 1348