

# Galaxy Clustering in Redshift Space

Shun Saito<sup>1, \*</sup>

<sup>1</sup>*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany*  
(Dated: June 8, 2016)

This short note is based on my day-long lecture for ‘Lecture Series on Cosmology’ at MPA on 9th June in 2016. The aim of this note is to review the galaxy clustering in redshift space, focusing mainly on the Redshift-Space Distortion (RSD) on cosmological scales from both model and measurement points of view in a self-consistent manner. The basic goal is to provide a brief overview of recent developments on RSD and to present the most updated BOSS DR12 result. It is true that there exist too many equations, but don’t worry! I will try to keep my explanations as simple as possible.

## Contents

<b>I. Preface</b>	2
<b>II. Introduction and illustrative pictures</b>	2
<b>III. Theory: modeling RSD in linear regime</b>	6
<b>IV. Theory: modeling RSD in nonlinear regime</b>	8
<b>V. Theory: impact of RSD on the projected clustering</b>	15
<b>VI. Analysis (Theory): quantifying the RSD information in an ideal survey</b>	17
<b>VII. Analysis: measuring RSD from the multipole in BOSS</b>	20
<b>VIII. Concluding remark</b>	25
<b>Acknowledgments</b>	25
<b>A. Convention and useful formula</b>	25
<b>B. Perturbation Theory basics</b>	26
<b>C. The redshift-space power spectrum model adopted in the actual analysis for BOSS</b>	28
<b>D. Derivation of Eq. (17)</b>	29
<b>References</b>	29

---

\*E-mail: [ssaito@mpa-garching.mpg.de](mailto:ssaito@mpa-garching.mpg.de)

## I. PREFACE

The goal of this internal lecture is to provide a simplified and pedagogical review of the redshift-space distortion (**RSD**) which is one of the main scientific targets in ongoing and forthcoming galaxy redshift surveys such as BOSS in SDSS-III, eBOSS in SDSS-IV, HETDEX, PFS, DESI, WFIRST and EUCLID.

I surely begin with the very basics assuming a textbook-level knowledge of cosmology. On the other hand, I try to include recent developments of both modeling and measurement efforts (although I apologize that these might be technical and advanced topics for general audience). As far as I am aware, there is no recent comprehensive and self-consistent review which focuses only on RSD in the literature. There was one by Hamilton [1] for the linear RSD almost 20 years ago and the review of the large-scale structure by Bernardeau et al. about 15 years ago [2] still remains standard as an introductory reading. This already reflects the fact that this field is still in a developing phase and not well matured yet. Nevertheless I personally think that it would be good to summarize the current status and even hope that the lecture is somehow extended to further collaboration.

This lecture is heavily based on our experiences and contributions to the field through the modeling works [3–5] and the observational analysis in BOSS [6, 7] (and hence could be somewhat biased), although I try to include as many relevant references as possible for further reading (of course, the list is likely to be very incomplete).

## II. INTRODUCTION AND ILLUSTRATIVE PICTURES

The goal of this lecture note is to answer or to help one better understand the approaches to answer the following questions:

- *What is RSD?*

When we map out objects like galaxies in 3-dimensional space, the radial (comoving) distance to the object is determined by its measured redshift,  $z_{\text{obs}}$ . However, we should remember that there are always two contributions to  $z_{\text{obs}}$ : Hubble flow,  $r(z_{\text{cos}}) = \int_0^{z_{\text{cos}}} cdz/H(z)$  and the peculiar velocity of the object as

$$1 + z_{\text{obs}} = (1 + z_{\text{cos}}) \left( 1 - \frac{v_{\parallel}(\mathbf{r})}{c} \right)^{-1}, \quad (1)$$

$$\mathbf{s} = \mathbf{r} + \frac{(1 + z_{\text{cos}})v_{\parallel}(\mathbf{r})}{H(z_{\text{cos}})} \hat{r}, \quad (2)$$

where  $v_{\parallel}$  denotes the line-of-sight (LOS) component of the peculiar velocity. The 2nd term is usually ignored in astronomy. For instance, in a flat  $\Lambda$ CDM universe with  $\Omega_{\text{m}0}=0.3$ ,  $r(z_{\text{cos}} = 0.5) \simeq 1.32 \text{ Gpc}/h$  while the second term is evaluated as

$$\left. \frac{(1 + z_{\text{cos}})v_{\parallel}(\mathbf{r})}{H(z_{\text{cos}})} \right|_{z_{\text{cos}}=0.5} \simeq 1.18 \frac{v_{\parallel}}{100 \text{ km/s}} [\text{Mpc}/h], \quad (3)$$

which typically amounts to  $\mathcal{O}(1 \text{ Mpc}/h)$ . Nevertheless, the existence of the 2nd term has non-negligible impact on the clustering statistics of the matter density field,

$$\delta_{\text{m}}(\mathbf{x}) = \frac{\rho_{\text{m}}(\mathbf{x})}{\bar{\rho}_{\text{m}}} - 1. \quad (4)$$

Its 2-pt correlation in Fourier space, the so-called power spectrum defined by

$$\langle \delta_{\text{m}}(\mathbf{k}) \delta_{\text{m}}(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') P_{\text{m}}(\mathbf{k}), \quad (5)$$

can be reduced to  $P_{\text{m}}(\mathbf{k}) = P_{\text{m}}(k)$  due to the cosmological principle. However, this does not hold in the case of the observed power spectrum in redshift space, simply because the peculiar velocity term

obviously breaks down rotational invariance. Therefore, *the peculiar velocity makes the redshift-space clustering anisotropic*. The same story shall hold for number counts for dark matter halos and galaxies. This is the so-called RSD.

The reason why this anisotropy is quantitatively non-negligible will be shortly explained in Sec. III. Here in turn let me discuss a schematic picture shown in Fig. 1 which illustrates the anisotropic clustering caused by RSD. At large scales, objects tend to coherently infall into high density region and hence the density field becomes *squashed* hence the clustering amplitude becomes *stronger* along LOS, so-called the Kaiser effect [8]. On the other hand, at small scales objects are virialized and hence have random motions. In this case, the density field becomes *stretched* hence the clustering amplitude becomes *smaller* along LOS, so-called the Finger-of-God (FoG) effect [9].

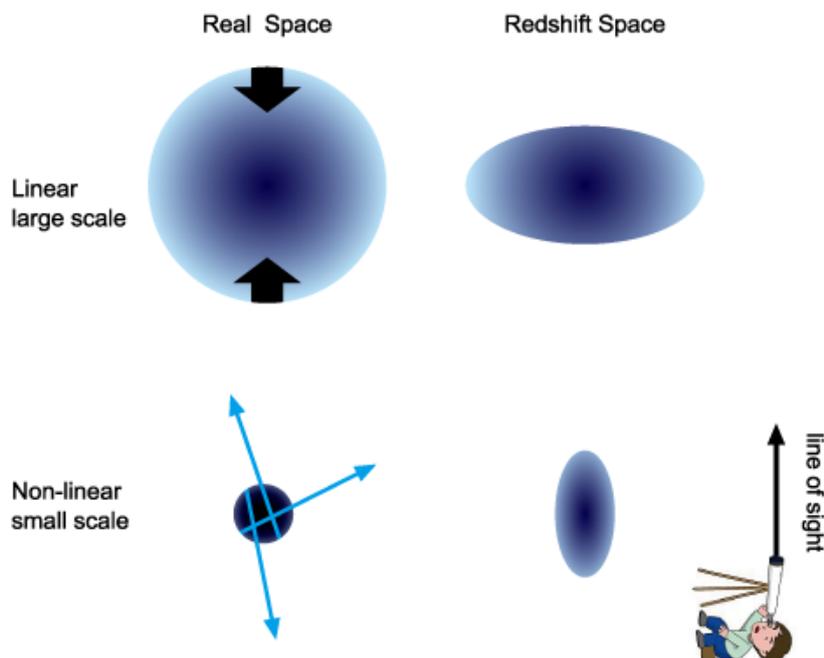


FIG. 1: The schematic picture of RSD. Picture courtesy of my wife, Kimika Saito.

- Why is RSD important and useful?

RSD is important because we can measure any quantities only as a function of *redshift-space distance*. In other words, for any cosmological observables involving radial distance, we should take RSD into account. Such observables include galaxy number count, Ly $\alpha$  forest, 21cm, and intensity mapping etc. I will also argue that RSD cannot be negligible in the case of projection onto the sky under some conditions (see Sec. V).

RSD is useful, because RSD is again the measurement of *cosmological velocity field* which is determined only through gravitational potential. In linear theory, the Euler equation is given by

$$v' + aHv = -\nabla\Psi, \quad (6)$$

where  $'$  denotes derivative w.r.t conformal time. Historically, RSD is proposed as a probe of density parameter, since the linear velocity field is directly proportional to growth function (e.g., [10]):

$$f \equiv \frac{d \ln D_1}{d \ln a} \approx \Omega_m(z)^{0.545}, \quad (7)$$

where  $D_1(a)$  is linear growth rate and  $\Omega_m(z) = H_0^2 \Omega_{m0} (1+z)^3 / H(z)^2$ . As far as I know, this idea is first introduced by Sargent and Turner (1977) [11] (see Fig. 2) rather than Kaiser (1987) [8].

A STATISTICAL METHOD FOR DETERMINING THE COSMOLOGICAL DENSITY PARAMETER  
 FROM THE REDSHIFTS OF A COMPLETE SAMPLE OF GALAXIES

WALLACE L. W. SARGENT

Hale Observatories, California Institute of Technology and Carnegie Institution of Washington

AND

EDWIN L. TURNER

Institute for Advanced Study, Princeton, New Jersey

Received 1976 July 12; revised 1976 November 2

ABSTRACT

The distribution of galaxies in space is approximated by their distribution in a “redshift space” in which their radial coordinate is  $cz/H_0$ . Deviations from a smooth and uniform Hubble expansion, due either to perturbations arising from density fluctuations in the distribution of galaxies or to virial motions in bound groups and clusters, cause characteristic distortions in “redshift space.” A method of detecting and measuring these distortions (anisotropies) from the relative redshifts and positions on the sky of pairs of galaxies is proposed. An approximate and a more powerful general method of relating these characteristic distortions to their associated density enhancements (and hence  $\Omega$ ) are presented. The limited data presently available are used to illustrate the approximate method, and a very tentative result of  $\Omega \approx 0.07$  is obtained. The data requirements for a strong test of  $\Omega$  using the general method are discussed.

*Subject headings:* cosmology — galaxies: redshifts

FIG. 2: Title and abstract of the 1st RSD paper by Sargent and Turner (1977) [11].

Nowadays RSD is paid more attentions to as a probe of gravity theory at cosmological scales because of the same reason above. Note that RSD is sensitive to *the dynamical mass* determined by  $\Psi$ , while weak lensing is sensitive to *the lensing mass* proportional to  $(\Phi + \Psi)$ .

- Why is it difficult to model RSD even at large scales  $\gtrsim \mathcal{O}(10 \text{ Mpc})$ ?

This is the main topic which I address in the first half of the lecture (Sec. III, IV, and V). A short answer is because the RSD involves the nonlinear mapping in terms of peculiar velocity as follows. Since the density field should be preserved,  $\rho_m^s(\mathbf{s})d\mathbf{s} = \rho_m(\mathbf{r})d\mathbf{r}$ . So the Jacobian is given by

$$J = \left| \frac{d\mathbf{r}}{d\mathbf{s}} \right| = \frac{r^2 dr}{s^2 ds} = \left\{ 1 - \frac{(1 + z_{\text{cos}}) v_{\parallel}}{H(z_{\text{cos}}) r} \right\}^{-2} \left\{ 1 - \frac{(1 + z_{\text{cos}}) \partial v_{\parallel}}{H(z_{\text{cos}}) \partial r} \right\}^{-1}. \quad (8)$$

In Sec. IV, I discuss the nonlinear RSD model on the basis of perturbation theory (PT) proposed by our paper, Taruya, Nishimichi, Saito (TNS, 2010) [3]. The TNS model has been applied to several galaxy surveys to extract the RSD information. Although its derivation is a bit technical, I think it is helpful to understand why modeling nonlinear RSD is such a difficult task when going through the derivation of the TNS model.

- How and to what extent can we extract cosmological information from the RSD measurement?

In this lecture, I mainly focus on modeling the redshift-space power spectrum,  $P^s(\mathbf{k})$  which is the Fourier transform of the two-point correlation function in configuration space. Although this is just my personal preference, I try to address the advantage of the redshift-space power spectrum in Sec. VI.

- What is the current result of the RSD measurements?

The current status of the RSD measurements is well summarized by Fig. 3 [12]. The meaning of  $f\sigma_8$  will be introduced shortly. It is worth noting that this is not really a fair comparison in the sense that a way to analyze and extract the RSD information is not exactly same.

As I mentioned earlier, I will briefly show the updated results from BOSS DR12 in Sec. VII. Since the DR12 results are not allowed to be public at this point, I do not present them in this note but show them in the lecture.

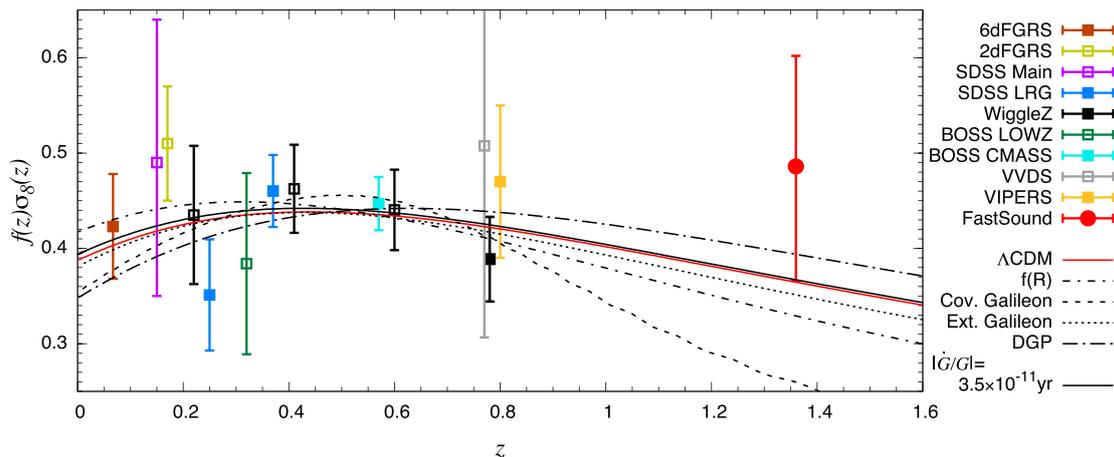


FIG. 3: A compilation of the recent RSD measurements [12].

### A. Assumptions

Here I summarize the assumptions which further simplify the analysis in both modeling and measurement.

- *Distant observer and global plain-parallel approximations in modeling RSD.*

Eq. (8) tells us that the velocity term in the first bracket can be ignored when

$$\frac{v_{\parallel}}{L} \ll kv_{\parallel} \Leftrightarrow kL \gg 1, \quad (9)$$

where  $k$  is wavenumber for the Fourier mode of interest and  $L$  is the characteristic size of the survey. Namely, the distant observer approximation is valid when the scale of interest is well within the survey size. Then the Jacobian is simply approximated by

$$J \simeq \left\{ 1 - \frac{(1 + z_{\text{cos}})}{H(z_{\text{cos}})} \frac{\partial v_{\parallel}}{\partial r} \right\}^{-1}. \quad (10)$$

This Jacobian formula guarantees that the RSD depends only on one direction, and hence we can fix LOS as one global direction such as  $\hat{z}$ ,

$$\hat{k} \cdot \hat{x} \approx \hat{k} \cdot \hat{z}. \quad (11)$$

I call this as the global observer approximation. I am going to assume the distant observer and the global plain-parallel approximations in modeling RSD in the following section.

In other words, the distant observer and the global plain-parallel approximations break down for the extremely large scale mode. Once the global plain-parallel approximation is dropped out, it does make more sense to interpret the clustering by making a radial-angular decomposition [13, 14]. Furthermore, keeping the  $v_{\parallel}/r$  terms leads to an additional correction term, the so-called *wide angle effect* (see e.g., [15–17]).

- *Local plain-parallel approximation in measuring the redshift-space power spectrum.*

Historically, the power-spectrum estimator introduced by Feldman, Kaiser and Peacock (the so-called FKP estimator) assumed the global plane-parallel approximation [18]. It turns out that this is no longer valid once one measures the anisotropic part of the power spectrum especially in the large-angle survey like BOSS [19, 20]. In Sec. VII, I will introduce more refined estimators which assume the local plain-parallel approximation,

$$\hat{k} \cdot \hat{x}_1 \approx \hat{k} \cdot \hat{x}_2 \approx \hat{k} \cdot \hat{x}_h, \quad (12)$$

where  $\mathbf{x}_1$  and  $\mathbf{x}_2$  describe two galaxy positions and  $\mathbf{x}_h = (\mathbf{x}_1 + \mathbf{x}_2)/2$ .

- *No velocity bias, i.e.,  $\mathbf{v}_g = \mathbf{v}_m$ .*

This means that it will be straightforward to go from matter to halo/galaxy in the RSD modeling as long as one has a good prescription of biased tracer,  $\delta_{h/g}(\mathbf{x}) = \mathcal{F}[\delta_m(\mathbf{x}), \mathcal{O}(\mathbf{x})]$ , in real space (see e.g., [21–26] and those references therein for recent efforts). Although the velocity field is determined by gravitational potential governed by matter density, it is not necessarily true that velocity bias is negligible (see [27, 28] for recent discussions).

## B. Uncovered topics

Of course it is impossible to cover all the topics on the galaxy clustering in redshift space in this lecture. Here is the list of topics which is relevant particularly to RSD but uncovered by this lecture:

- Other nonlinear RSD models of the power spectrum or 2pt correlation function which include integrated Lagrangian Perturbation Theory (iLPT) [29], the Convolution Lagrangian Perturbation Theory (CLPT) [30], other LPT [31], Effective Field Theory (EFT) [32], the Distribution Function approach [33–38] etc.
- Higher-order statistics such as the bispectrum (the 3pt correlation function). See e.g., [39].
- Horizon effect such as GR, the wide-angle effect etc. See e.g., [40].
- Small-scale physics such as the galaxy-halo connection, impact of baryon etc. See e.g., [41–43].

## III. THEORY: MODELING RSD IN LINEAR REGIME

### A. The real-space power spectrum in a nutshell

I assume everyone is quite familiar with the linear theory in real space.

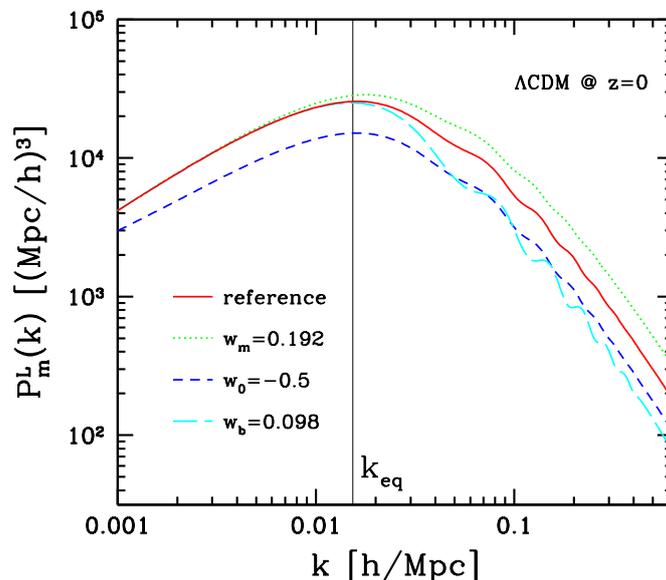


FIG. 4: The linear matter power spectrum in a  $\Lambda$ CDM universe with  $\Omega_{m0}h^2 = 0.147$ ,  $\Omega_{m0}h^2 = 0.0245$ ,  $\Omega_\Lambda = 0.7$ ,  $h = 0.7$ ,  $\Delta_{\mathcal{R}}^2(k_*) = 2.35 \times 10^{-9}$  and  $n_s = 0.95$ .

## B. Linear RSD: Kaiser formula

For convenience, I define the velocity divergence field as

$$\theta(\mathbf{x}) \equiv -\frac{\nabla \cdot \mathbf{v}(\mathbf{x})}{aHf}, \quad (13)$$

and its Fourier transform is give by

$$\mathbf{v}(\mathbf{k}) = -iaHf \frac{\mathbf{k}}{k^2} \theta(\mathbf{k}). \quad (14)$$

Then the linear continuity equation,  $\delta_m^{L'} + \nabla \cdot \mathbf{v}^L = 0$  becomes

$$\delta_m^L(\mathbf{k}) = \theta^L(\mathbf{k}). \quad (15)$$

Now the density conservation from real to redshift space implies that

$$\delta_m^s(\mathbf{s}) = \left| \frac{d\mathbf{s}}{d\mathbf{r}} \right|^{-1} \{1 + \delta_m(\mathbf{r})\} - 1, \quad (16)$$

and its Fourier component is obtained as

$$\delta_m^s(\mathbf{k}) = \int d^3x \left\{ \delta_m(\mathbf{x}) - \frac{1}{aH} \frac{\partial v_z(\mathbf{x})}{\partial z} \right\} e^{i\mathbf{k} \cdot \mathbf{x} + ik\mu v_z/(aH)}, \quad (17)$$

where I fix the LOS direction as  $\hat{\mathbf{z}}$  and define the directional cosine as  $\mu \equiv \hat{\mathbf{k}} \cdot \hat{\mathbf{z}}$ . Note that this expression is *exact* under the distant observer and global plain-parallel approximations (see Appendix. D for detailed derivation).

Now let me derive the RSD correction at linear order, known as the famous Kaiser formula [8]. At linear order in terms of  $\delta$  and  $v$ , the 2nd term in the power of the exponential factor is dropped out, and hence one obtains

$$\begin{aligned} \delta_m^{s,L}(\mathbf{k}) &= \delta_m(\mathbf{k}) - \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} \frac{1}{aH} \frac{\partial}{\partial z} \int \frac{d^3k'}{(2\pi)^3} e^{-i\mathbf{k}' \cdot \mathbf{x}} v_z(\mathbf{k}') \\ &= \delta_m(\mathbf{k}) + f \int \frac{d^3k'}{(2\pi)^3} \int d^3x e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}} \frac{k_z k'_z}{k^2} \theta(\mathbf{k}) \\ &= \delta_m(\mathbf{k}) + f\mu^2 \theta(\mathbf{k}) \\ &= (1 + f\mu^2) \delta_m^L(\mathbf{k}), \end{aligned} \quad (18)$$

and the redshift-space power spectrum at linear order is given by

$$P_m^{s,L}(\mathbf{k}) = P_m^{s,L}(k, \mu) = (1 + f\mu^2)^2 P_m^L(k). \quad (19)$$

In the case of galaxy number density with  $\delta_g = b\delta_m$ , similarly one obtains

$$P_g^{s,L}(\mathbf{k}) = P_g^{s,L}(k, \mu) = b^2(1 + \beta\mu^2)^2 P_m^L(k), \quad (20)$$

where  $\beta \equiv f/b$ . It is important to realize that the anisotropic term originates from the velocity and hence it does not depend on bias. This is the reason why the RSD measurement is often parametrized by the amplitude of the peculiar velocity field  $f\sigma_8(z_{\text{cos}})$ .

How significant is the RSD correction? In order to see this, let me expand the anisotropic power spectrum with the Legendre polynomials,

$$P^s(k, \mu) = \sum_{\ell} P_{\ell}(k) \mathcal{L}_{\ell}(\mu), \quad (21)$$

$$P_{\ell}(k) = \frac{2\ell + 1}{2} \int_{-1}^1 d\mu P^s(k, \mu) \mathcal{L}_{\ell}(\mu). \quad (22)$$

Since the Kaiser formula contains terms only up to  $\mu^4$ , only  $\ell = 0$  (monopole), 2 (quadrupole) and 4 (hexadecapole) are non-vanishing (e.g., [44]):

$$P_{g,\ell=0}(k) = \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right) b^2 P_m^L(k), \quad (23)$$

$$P_{g,\ell=2}(k) = \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right) b^2 P_m^L(k), \quad (24)$$

$$P_{g,\ell=4}(k) = \frac{8}{35}\beta^2 b^2 P_m^L(k). \quad (25)$$

Suppose that  $f = 0.774$  and  $b = 2$  at  $z = 0.57$  (roughly corresponding to the BOSS CMASS sample), the Kaiser factor is evaluated as  $P_{g,\ell=0}(k)/P_g \simeq 1.288$ . This means that RSD introduces overall correction by a factor of 1.3 even for the monopole, i.e., isotropic part. This is a significant effect! Then a next and natural question is how well we can measure the anisotropic part such as quadrupole and hexadecapole, which will be answered in Sec. VI.

Let me make a comment on other (but highly related) two-point statistics. The multipole of the correlation function in configuration space is simply related to the power spectrum multipole,

$$\xi_\ell^s(s) = i^\ell \int \frac{k^2 dk}{2\pi^2} P_\ell^s(k) j_\ell(ks), \quad (26)$$

where  $j_\ell(x)$  is the spherical Bessel function at  $\ell$ -th order. I should also mention that the clustering wedges (see e.g., [45, 46]) defined by simple average within certain range of  $\mu$ ,

$$P_{\mu_1}^{s,\mu_2}(k) \equiv \frac{1}{\mu_2 - \mu_1} \int_{\mu_1}^{\mu_2} d\mu P^s(k, \mu). \quad (27)$$

#### IV. THEORY: MODELING RSD IN NONLINEAR REGIME

This section more or less highlights our paper, Taruya, Nishimichi, and Saito (TNS, 2010) [3]. I also try to include a recent work by Zheng and Song (2016) [47].

##### A. How hard is RSD to predict?

- First of all, how good is the Kaiser formula? See Fig. 5.
- What about next-to-leading order calculation in standard PT (SPT) [48]? See Fig. 5.

$$P_{\text{SPT}}^s(k, \mu) = (1 + f\mu^2)^2 P^L(k) + P_{1\text{loop}}^s(k, \mu). \quad (28)$$

- What about phenomenological models (e.g., [44, 49])? See Fig. 6.

$$P_{\text{pheno}}^s(k, \mu) = D_{\text{FoG}}(k\mu f\sigma_v) P_{\text{Kaiser}}^s(k, \mu), \quad (29)$$

where

$$P_{\text{Kaiser}}^s(k, \mu) = \begin{cases} (1 + f\mu^2)^2 P_{\delta\delta}(k) & ; \text{ linear} \\ P_{\delta\delta}(k) + 2f\mu^2 P_{\delta\theta}(k) + f^2\mu^4 P_{\theta\theta}(k) & ; \text{ non-linear} \end{cases} \quad (30)$$

$$D_{\text{FoG}}(x) = \begin{cases} \exp(-x^2) & ; \text{ Gaussian} \\ 1/(1+x^2) & ; \text{ Lorentzian} \end{cases} \quad (31)$$

Note that here the velocity dispersion,  $\sigma_v^2$ ,

$$\sigma_v^2 = \frac{1}{3} \int \frac{d^3q}{(2\pi)^3} \frac{P_{\theta\theta}(q)}{q^2}, \quad (32)$$

is treated as a free parameter but sometimes fixed by the linear one,

$$\sigma_v^{L^2} = \frac{1}{6\pi^2} \int dq P_m^L(q). \quad (33)$$

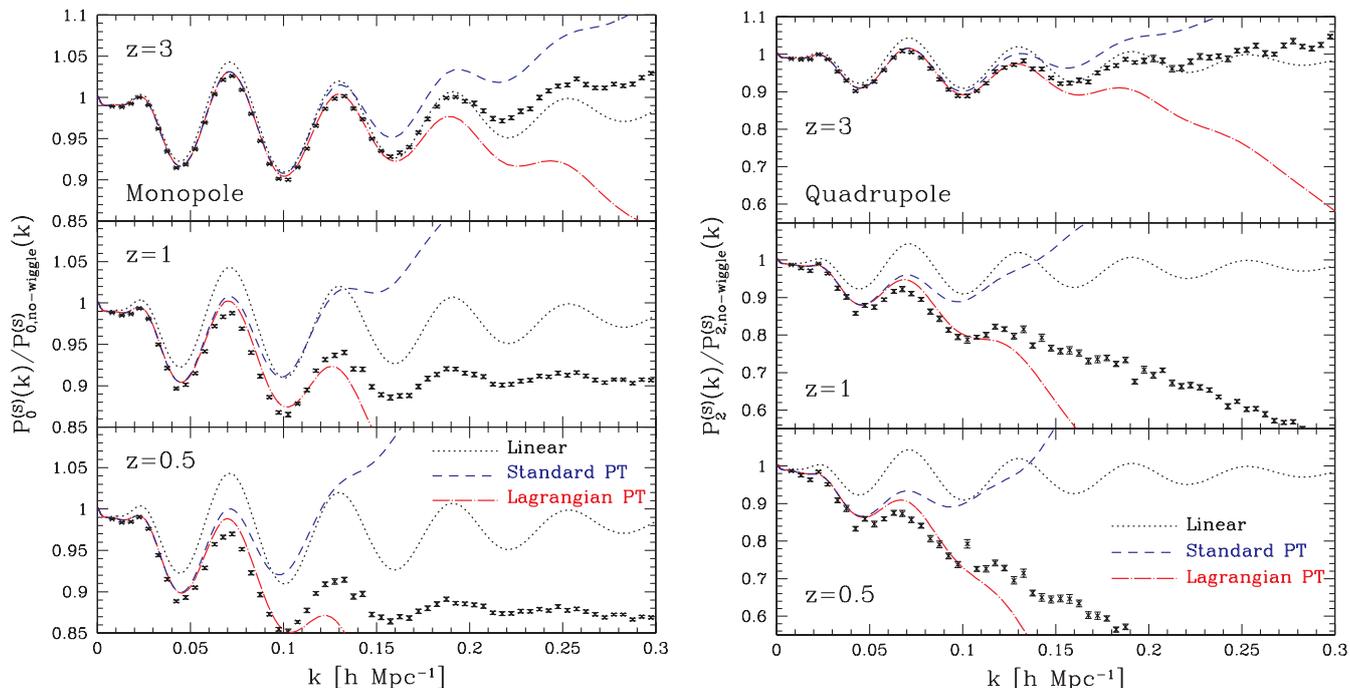


FIG. 5: Comparison between PT predictions and N-body simulation [3]. Ratio of power spectra to smoothed reference spectra in redshift space are plotted.  $P_\ell^{(S)}(k)/P_{\ell,\text{no-wiggle}}^{(S)}(k)$ . N-body results are taken from the `wmap5` simulations of Ref. [50]. The reference spectrum  $P_{\ell,\text{no-wiggle}}^{(S)}$  is calculated from the no-wiggle approximation of the linear transfer function, and the linear theory of the Kaiser effect is taken into account. Short dashed and dot-dashed lines respectively indicate the results of one-loop PT and Lagrangian PT calculations for redshift-space power spectrum.

## B. Derivation of the TNS model (and beyond)

It is always good to begin with the exact expression. For simplicity, I here focus on the matter density field and omit the subscript, ‘m’. Using Eq. (17), the redshift-space power spectrum is exactly given by

$$P^s(\mathbf{k}) = \int d^3r e^{i\mathbf{k}\cdot\mathbf{r}} \langle e^{-ik\mu f \Delta u_z} \{ \delta(\mathbf{x}) + f \nabla_z u_z(\mathbf{x}) \} \{ \delta(\mathbf{x}') + f \nabla_z u_z(\mathbf{x}') \} \rangle \quad (34)$$

$$= \int d^3r e^{i\mathbf{k}\cdot\mathbf{r}} \langle e^{j_1 A_1} A_2 A_3 \rangle, \quad (35)$$

where I define the following variables to simplify the equation;  $\mathbf{u} \equiv -\mathbf{v}/(aH)$ ,  $\mathbf{r} \equiv \mathbf{x} - \mathbf{x}'$ ,  $\Delta u_z \equiv u_z(\mathbf{x}) - u_z(\mathbf{x}')$ , and

$$j_1 = -ik\mu f, \quad (36)$$

$$A_1 = \Delta u_z = u_z(\mathbf{x}) - u_z(\mathbf{x}'), \quad (37)$$

$$A_2 = \delta(\mathbf{x}) + f \nabla_z u_z(\mathbf{x}), \quad (38)$$

$$A_3 = \delta(\mathbf{x}') + f \nabla_z u_z(\mathbf{x}'). \quad (39)$$

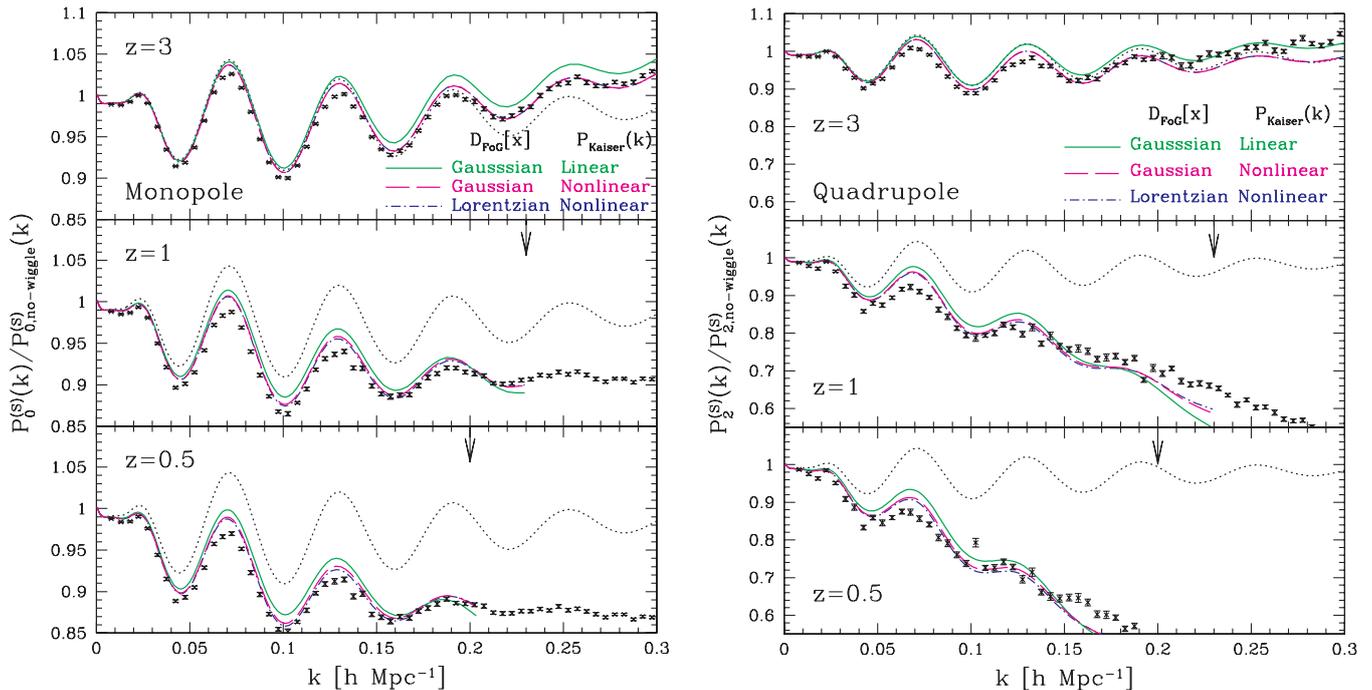


FIG. 6: Comparison between phenomenological models and N-body simulation [3]. Same as in Fig. 5, but we here plot the results of phenomenological model predictions. The three different predictions depicted as solid, dashed, dot-dashed lines are based on the phenomenological model of redshift distortion with various choices of Kaiser and Finger-of-God terms (Eq. (31)). Left panel shows the monopole power spectra ( $\ell = 0$ ), and the right panel shows the quadrupole spectra ( $\ell = 2$ ). In all cases, one-dimensional velocity dispersion  $\sigma_v$  was determined by fitting the predictions to the N-body simulations. In each panel, vertical arrow indicates the maximum wavenumber  $k_{1\%}$  for valid range of the improved PT.

Why is it very difficult to exactly evaluate Eq. (35)?

- Even if one assumes the density and velocity fields are Gaussian, Eq. (35) still contains the exponential prefactor. This physically means that *nonlinear mapping in terms of velocity cannot be avoided*.
- Irrespective to the pair separation scale,  $\mathbf{r}$ , Eq. (35) involves the terms like  $\int d^3r \sum_n \langle \Delta u_z^n \rangle$  which is in principle sensitive to small-scale physics and hence hopeless to evaluate on the PT basis. This physically corresponds to the fact *the FoG effect due to the random motion of virialized object is important even at large scales and hard to be modeled*.

The cumulant expansion theorem tells us that (e.g., [49])

$$\langle e^{\mathbf{j} \cdot \mathbf{A}} \rangle = \exp\{\langle e^{\mathbf{j} \cdot \mathbf{A}} \rangle_c\}. \quad (40)$$

By taking the derivative twice w.r.t  $j_2$  and  $j_3$  and then setting  $j_2 = j_3 = 0$ , one obtains

$$\langle e^{j_1 A_1} A_2 A_3 \rangle = \exp\{\langle e^{j_1 A_1} \rangle_c\} [\langle e^{j_1 A_1} A_2 A_3 \rangle_c + \langle e^{j_1 A_1} A_2 \rangle_c \langle e^{j_1 A_1} A_3 \rangle_c]. \quad (41)$$

Note that Eq. (41) is an exact expression, and the question is how to evaluate Eq. (41). TNS's approach is as follows:

- What we want is an expression at large scale limit,  $k\mu \rightarrow 0$ , i.e.,  $j_1 \rightarrow 0$  where PT should work well. However, from the considerations in the previous subsection, we see that *the naive PT expansion (SPT) does not work*.
- Therefore, the exponential prefactor is decided to be left. We assume that spatial correlations between  $u_z(\mathbf{x})$  and  $u_z(\mathbf{x}')$  are ignored, and also

$$\langle A_1^n \rangle_c \simeq 2 \langle u_z(\mathbf{x})^n \rangle_c = 2c_n \sigma_v^n, \quad (42)$$

for even  $n$  with  $c_n$  being constants. We further simplify the exponential prefactor by assuming that  $c_2 = 1$ ,  $c_{2n} = (2n - 1)!$  for  $n > 2$ ,  $c_{2n-1} = 0$ . These approximations result in

$$\exp\{\langle e^{j_1 A_1} \rangle_c\} \approx \exp[-k^2 \mu^2 f^2 \sigma_{v,\text{eff}}^2]. \quad (43)$$

I write the velocity dispersion parameter as  $\sigma_{v,\text{eff}}^2$  because Eq. (43) is not obviously an exact expression and  $\sigma_{v,\text{eff}}^2$  is treated as a free parameter. To be fair, this is the biggest disadvantage of the TNS model. I will get back to the validity of these approximation by referring to the speculations in [47].

- On the other hand, we expand the 2nd bracket in Eq. (41) in terms of  $j_1$  as

$$\langle e^{j_1 A_1} A_2 A_3 \rangle_c + \langle e^{j_1 A_1} A_2 \rangle_c \langle e^{j_1 A_1} A_3 \rangle_c \simeq \langle A_2 A_3 \rangle_c + j_1 \langle A_1 A_2 A_3 \rangle_c + j_1^2 \left\{ \frac{1}{2} \langle A_1^2 A_2 A_3 \rangle_c + \langle A_1 A_2 \rangle_c \langle A_1 A_3 \rangle_c \right\} + \mathcal{O}(j_1^3), \quad (44)$$

where we are going to ignore the term,  $\langle A_1^2 A_2 A_3 \rangle_c$ , since the leading contribution in PT is the tree-level trispectrum roughly proportional to  $\mathcal{O}(P^L(k)^3)$ . This term is actually evaluated in other works and turned out to be negligible at scales of interest (e.g., the ‘D’ term in [51], also see [47] and Eq. (54)).

As a result, we finally derive the TNS formula:

$$P_{\text{m,TNS}}^s(k, \mu) = \exp[-k^2 \mu^2 f^2 \sigma_{v,\text{eff}}^2] \{ P_{\delta\delta}(k) + 2f\mu^2 P_{\delta\theta}(k) + f^2 \mu^4 P_{\theta\theta}(k) + A(k, \mu; f) + B(k, \mu; f) \}, \quad (45)$$

where the new correction terms,  $A(k, \mu)$  and  $B(k, \mu)$ , are given by

$$A(k, \mu; f) = j_1 \int d^3 r e^{i\mathbf{k}\cdot\mathbf{r}} \langle A_1 A_2 A_3 \rangle_c = k\mu f \int \frac{d^3 p}{(2\pi)^3} \frac{p_z}{p^2} \{ B_\sigma(\mathbf{p}, \mathbf{k} - \mathbf{p}, -\mathbf{k}) - B_\sigma(\mathbf{p}, \mathbf{k}, -\mathbf{k} - \mathbf{p}) \}, \quad (46)$$

$$B(k, \mu; f) = j_1^2 \int d^3 r e^{i\mathbf{k}\cdot\mathbf{r}} \langle A_1 A_2 \rangle_c \langle A_1 A_3 \rangle_c = (k\mu f)^2 \int \frac{d^3 p}{(2\pi)^3} F_\sigma(\mathbf{p}) F_\sigma(\mathbf{k} - \mathbf{p}), \quad (47)$$

where the function  $B_\sigma$  and  $F_\sigma$  are defined by

$$(2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_\sigma(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \left\langle \theta(\mathbf{k}_1) \left\{ \delta(\mathbf{k}_2) + f \frac{k_{2z}^2}{k_2^2} \theta(\mathbf{k}_2) \right\} \left\{ \delta(\mathbf{k}_3) + f \frac{k_{3z}^2}{k_3^2} \theta(\mathbf{k}_3) \right\} \right\rangle, \quad (48)$$

$$F_\sigma(\mathbf{p}) = \frac{p_z}{p^2} \left\{ P_{\delta\theta}(p) + f \frac{p_z^2}{p^2} P_{\theta\theta}(p) \right\}. \quad (49)$$

In the TNS paper, we follow the standard PT technique to compute  $A$  and  $B$  correction terms up to next-to-leading order (i.e.,  $\mathcal{O}(P^L(k)^2)$ ). Here I omit the full expressions (which are quite long!), and refer them to [3]. If you hesitate to follow the formulas, just use the public code (there is a public version on Atsushi’s personal website, and ask me if you want a CAMB-integrated version). Just to provide a sense,  $A(k, \mu)$  and  $B(k, \mu)$  terms contain up to  $f^3 \mu^6$  and  $f^4 \mu^8$ , respectively, since one velocity divergence term has  $f\mu^2$  dependence. Also it is worth mentioning that Ref. [51] made an attempt to improve the evaluation of  $A$  and  $B$  correction terms with the multi-point propagator, but showed that the difference is basically absorbed into the FoG factor. Switching to a biased tracer such as galaxy in the case of linear bias, one find

$$P_{\text{g,TNS}}^s(k, \mu) = \exp[-k^2 \mu^2 f^2 \sigma_{v,\text{eff}}^2] \{ b^2 P_{\delta\delta}(k) + 2bf\mu^2 P_{\delta\theta}(k) + f^2 \mu^4 P_{\theta\theta}(k) + b^3 A(k, \mu; \beta) + b^4 B(k, \mu; \beta) \}. \quad (50)$$

Note that the bias dependence of  $b^3$  and  $b^4$  in  $A$  and  $B$  correction terms is just an artifact of  $\beta$  parametrization and they indeed contains terms only up to  $b^2$ .

- How well does the TNS model perform?

See Fig. 7. It looks certainly better than the previous figure. We indeed show that the TNS model better recovers the input  $f\sigma_8$  value in the  $N$ -body simulations. Also it is worth pointing out that the TNS formula performs worse at larger  $k\mu$  and hence higher-order multipole.

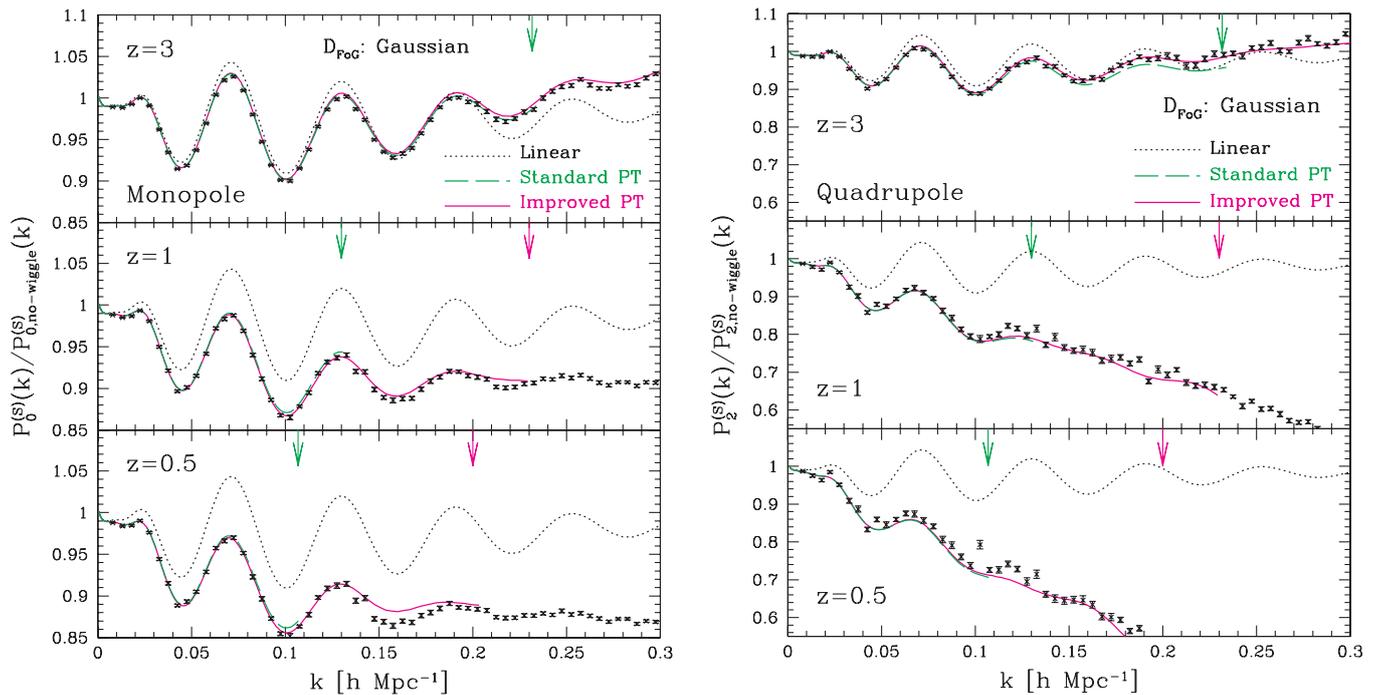


FIG. 7: Same as in Fig. 6, but we here adopt the TNS model.

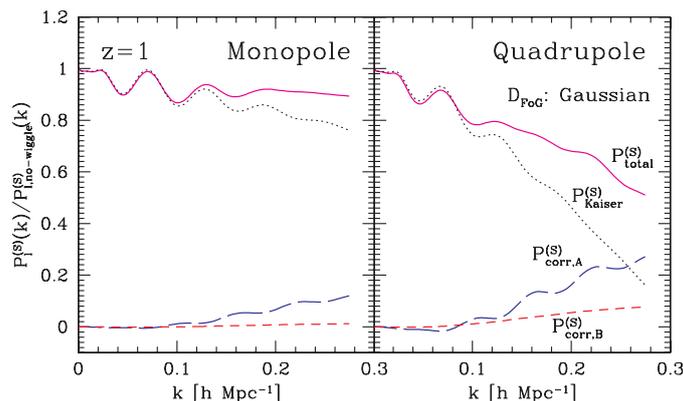


FIG. 8: Contribution of each correction term in redshift-space power spectrum. For monopole ( $\ell = 0$ , left) and quadrupole ( $\ell = 2$ , right) spectra of the improved model prediction at  $z = 1$  shown in solid lines, we divide the total power spectrum  $P_{\text{total}}^s$  (solid) into the three pieces as  $P_{\text{total}}^s = P_{\text{Kaiser}}^s + P_{\text{corr,A}}^s + P_{\text{corr,B}}^s$ , and each contribution is separately plotted dividing by smoothed reference spectra,  $P_{\ell,\text{no-wiggle}}^s$ . Here, the spectrum  $P_{\text{Kaiser}}^s$  (dotted) is the contribution of non-linear Kaiser term convolved with the Finger-of-God damping and the corrections  $P_{\text{corr,A}}^s$  and  $P_{\text{corr,B}}^s$ .

- How do  $A(k, \mu)$  and  $B(k, \mu)$  look like?

See Fig. 8. As expected, the correction terms become more significant at higher-order multipole. The A term is more important to better recover the BAO features.

#### Further investigation by Zheng and Song (2016)

There is one recent work by Zheng and Song (ZS, 2016) [47] which further tries to improve the TNS model. Here let me briefly discuss their ideas and approaches. They realize that the exponential prefactor can be

further decomposed into two parts: Eq. (40) can be rewritten as

$$\begin{aligned} \exp\{\langle e^{\mathbf{j}\cdot\mathbf{A}} \rangle_c\} &= \exp\left[\sum_{n=1}^{\infty} j_1^n \frac{\langle A_1^n \rangle_c}{(n)!}\right] \\ &= \underbrace{\exp\left[\sum_{n=1}^{\infty} j_1^{2n} \frac{2\langle u_z(\mathbf{x})^{2n} \rangle_c}{(2n)!}\right]}_{D_{\text{nonlocal}}^{\text{FoG}}(k, \mu)} \underbrace{\exp\left[\sum_{n=1}^{\infty} j_1^{2n} \frac{\langle \{u_z(\mathbf{x}) - u_z(\mathbf{x}')\}^{2n} \rangle_c - \langle u_z(\mathbf{x})^{2n} \rangle_c - \langle u_z(\mathbf{x}')^{2n} \rangle_c}{(2n)!}\right]}_{D_{\text{local}}^{\text{FoG}}(k, \mu; \mathbf{r})} \end{aligned} \quad (51)$$

Since the first term is the cumulant of the one-point distribution function and does not depend on the separation scale, it can be integrated out in Eq. (35). Then they consistently keep the leading-order term proportional to  $j_1^2$  in each FoG function as

$$D_{\text{nonlocal}}^{\text{FoG}}(k, \mu) \simeq \exp[-k^2 \mu^2 f^2 \sigma_{v, \text{eff}}^2] \quad (52)$$

$$D_{\text{local}}^{\text{FoG}}(k, \mu; \mathbf{r}) \simeq \exp[-k^2 \mu^2 f^2 \langle u_z(\mathbf{x}) u_z(\mathbf{x}') \rangle_c], \quad (53)$$

and then they provide the modified formula,

$$\begin{aligned} P_{\text{m,ZS}}^s(k, \mu) &= \exp[-k^2 \mu^2 f^2 \sigma_{v, \text{eff}}^2] \{P_{\delta\delta}(k) + 2f\mu^2 P_{\delta\theta}(k) + f^2 \mu^4 P_{\theta\theta}(k) + A(k, \mu; f) + B(k, \mu; f) \\ &\quad + C_{\text{TNS}}(k, \mu; f) + D_{\text{TNS}}(k, \mu; f)\}, \end{aligned} \quad (54)$$

where  $C_{\text{TNS}}$  and  $D_{\text{TNS}}$  correction terms are given by

$$D_{\text{TNS}}(k, \mu; f) = T_{\text{ZS}}(k, \mu; f) = \frac{1}{2} j_1^2 \int d^3 r e^{i\mathbf{k}\cdot\mathbf{r}} \langle A_1^2 A_2 A_3 \rangle_c, \quad (55)$$

$$C_{\text{TNS}}(k, \mu; f) = F_{\text{ZS}}(k, \mu; f) = -j_1^2 \int d^3 r e^{i\mathbf{k}\cdot\mathbf{r}} \langle u_z(\mathbf{x}) u_z(\mathbf{x}') \rangle_c \langle A_2 A_3 \rangle_c. \quad (56)$$

The tree-level PT expressions for  $C$  and  $D$  terms can be found in [3, 47]. Note that ZS call  $D_{\text{TNS}}$  as  $T_{\text{ZS}}$ , and  $C_{\text{TNS}}$  as  $F_{\text{ZS}}$ . I choose this convention because we already discuss that  $C$  and  $D$  correction terms are subdominant at least for monopole and quadrupole in the TNS paper. Indeed it is not new to introduce the  $C$  term because the SPT expression already includes this term as

$$P_{\text{m,SPT}}^s(k, \mu) = [1 - k^2 \mu^2 f^2 \sigma_v^{L2}] \{P_{\delta\delta}(k) + 2f\mu^2 P_{\delta\theta}(k) + f^2 \mu^4 P_{\theta\theta}(k) + A(k, \mu; f) + B(k, \mu; f) + C_{\text{TNS}}(k, \mu; f)\}. \quad (57)$$

Nevertheless I appreciate the fact that they realize that every correction term can be directly measured from simulations (see e.g., Fig. 9). They indeed show that each correction term starts to deviate from the simulation results at larger  $k\mu$ , and the Gaussian FoG prefactor is a good approximation up at a certain  $k\mu$ . For more detail I refer to the ZS paper.

### C. Nonlinear RSD from another different point of view

One may have heard of the so-called *streaming model* as one of the nonlinear RSD model [49, 52]:

$$1 + \xi^s(s_{\parallel}, s_{\perp}) = \int dr_{\parallel} [1 + \xi(r)] \mathcal{P}(r_{\parallel} - s_{\parallel}, \mathbf{r}), \quad (58)$$

where  $r^2 = r_{\perp}^2 + r_{\parallel}^2$ ,  $s_{\perp} = r_{\perp}$ , and  $\mathcal{P}(v_{\parallel}, \mathbf{r})$  denotes the pairwise velocity probability distribution function defined by

$$\mathcal{P}(v, \mathbf{r}) \equiv \int \frac{d\gamma}{2\pi} e^{i\gamma v} \mathcal{M}(-if\gamma, \mathbf{r}), \quad (59)$$

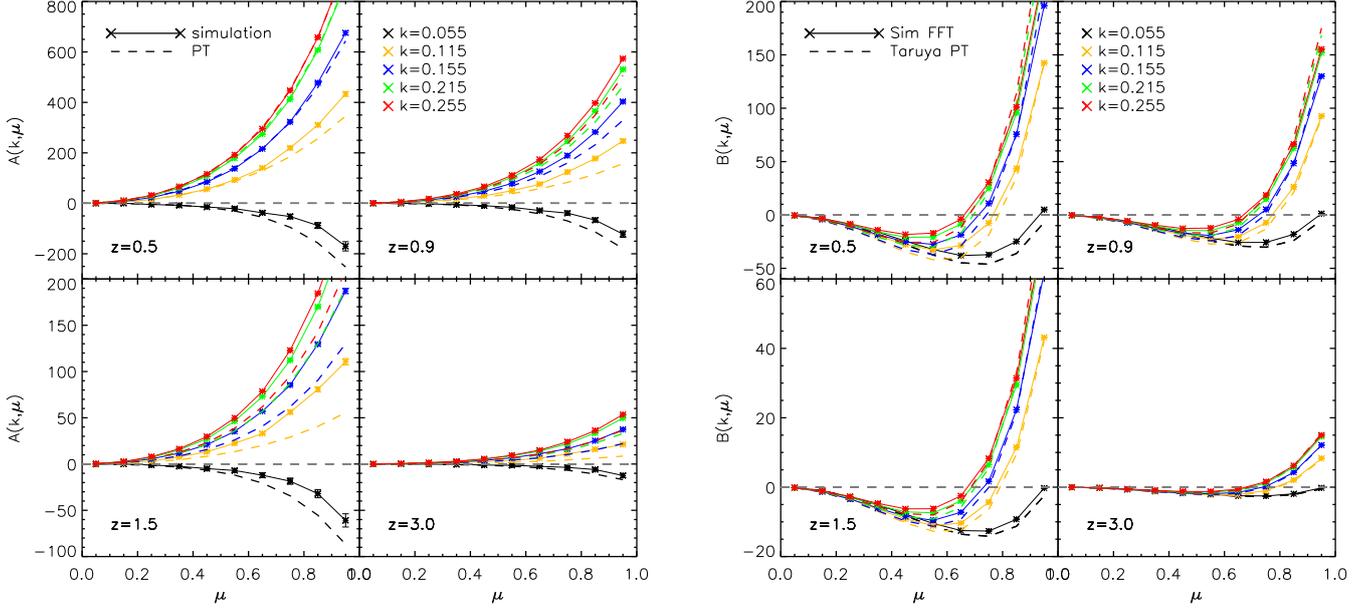


FIG. 9: Comparison of  $A(k, \mu)$  (left) and  $B(k, \mu)$  (right) between PT and direct measurements in simulations [47].

where  $\mathcal{M}$  is the pairwise velocity generating function,

$$\mathcal{Z}(\lambda, \mathbf{r}) \equiv [1 + \xi(\mathbf{r})]\mathcal{M}(\lambda, \mathbf{r}) \equiv \left\langle e^{\lambda \Delta u_z} [1 + \delta(\mathbf{x})][1 + \delta(\mathbf{x}')] \right\rangle. \quad (60)$$

This is useful because the line-of-sight pairwise velocity moments are obtained by its derivative as

$$\mathbf{v}_{12}(\mathbf{r}) \equiv \left( \frac{\partial \mathcal{M}}{\partial \lambda} \right)_{\lambda=0}, \quad (61)$$

$$\sigma_{12}^2(\mathbf{r}) \equiv \left( \frac{\partial^2 \mathcal{M}}{\partial \lambda^2} \right)_{\lambda=0}. \quad (62)$$

Also, the function  $\mathcal{Z}$  is useful when comparing different models as follows. After the cumulant expansion, the exact expression of  $\mathcal{Z}$  is given by

$$\begin{aligned} \mathcal{Z}^{\text{exact}}(\lambda, \mathbf{r}) = & \exp \left[ \left\langle e^{\lambda \Delta u_z} \right\rangle_c \right] \left[ 1 + \left\langle e^{\lambda \Delta u_z} \delta \right\rangle_c + \left\langle e^{\lambda \Delta u_z} \delta' \right\rangle_c \right. \\ & \left. + \left\langle e^{\lambda \Delta u_z} \delta \right\rangle_c \left\langle e^{\lambda \Delta u_z} \delta' \right\rangle_c + \left\langle e^{\lambda \Delta u_z} \delta \delta' \right\rangle_c \right] \end{aligned} \quad (63)$$

and therefore the approximation in the TNS model corresponds to

$$\begin{aligned} \mathcal{Z}^{\text{TNS}}(\lambda, \mathbf{r}) = & \exp \left[ \left\langle e^{\lambda \Delta u_z} \right\rangle_c \right] \left[ \langle (1 + \delta)(1 + \delta') \rangle_c \right. \\ & + \lambda \langle (1 + \delta)(1 + \delta') \Delta u_z \rangle_c \\ & \left. + \frac{\lambda^2}{2} \langle (1 + \delta)(1 + \delta') (\Delta u_z)^2 - \delta \delta' (\Delta u_z)^2 \rangle_c \right]. \end{aligned} \quad (64)$$

This can be compared with the configuration space model proposed by Reid and White (RW, 2012) [53]:

$$\mathcal{P}^{\text{RW}}(v, \mathbf{r}) = \frac{1}{\sqrt{2\pi} f \sigma_{12}(\mathbf{r})} \exp \left[ -\frac{\{v - f v_{12}(\mathbf{r})\}^2}{2 f^2 \sigma_{12}(\mathbf{r})^2} \right]. \quad (65)$$

Then we have

$$\mathcal{M}^{\text{RW}}(\lambda, \mathbf{r}) = \exp \left[ v_{12}(\mathbf{r}) \lambda + \frac{1}{2} \{ \sigma_{12}(\mathbf{r})^2 - v_{12}(\mathbf{r})^2 \} \lambda^2 \right], \quad (66)$$

and

$$\mathcal{Z}^{\text{RW}}(\lambda, \mathbf{r}) = [1 + \xi(r)] \exp \left[ v_{12}(\mathbf{r})\lambda + \frac{1}{2} \{ \sigma_{12}(\mathbf{r})^2 - v_{12}(\mathbf{r})^2 \} \lambda^2 \right]. \quad (67)$$

Suppose the bracket in the exponential factor is expanded in terms of  $\lambda$  up to the second order, we have

$$\mathcal{Z}^{\text{RW}}(\lambda, \mathbf{r}) \simeq [1 + \xi(r)] \times [v_{12}(\mathbf{r})\lambda + \frac{1}{2} \sigma_{12}(\mathbf{r})^2 \lambda^2] \quad (68)$$

$$\begin{aligned} &= \langle (1 + \delta)(1 + \delta') \rangle \\ &\quad + \lambda \langle (1 + \delta)(1 + \delta') \Delta u_z \rangle \\ &\quad + \frac{\lambda^2}{2} \langle (1 + \delta)(1 + \delta') (\Delta u_z)^2 \rangle, \end{aligned} \quad (69)$$

where we have used definition of  $v_{12}$  and  $\sigma_{12}$ ,

$$v_{12}(\mathbf{r}) = \frac{\langle \Delta u_z (1 + \delta)(1 + \delta') \rangle}{1 + \xi(\mathbf{r})}, \quad (70)$$

$$\sigma_{12}(\mathbf{r})^2 = \frac{\langle (\Delta u_z)^2 (1 + \delta)(1 + \delta') \rangle}{1 + \xi(\mathbf{r})}. \quad (71)$$

A comparison between Eqs. (64) and (69) shows that two expressions are quite similar at very small  $\lambda$  except for the exponential prefactor in TNS and unconnected pieces in RW. At least, as for the bias dependence, the term,  $\lambda \langle \delta \delta' \Delta u_z \rangle_c$  which is exactly corresponding to the A-term in TNS, is proportional to  $b^2$ , and is the second term in RW as well. As shown in RW, if the large-scale limit is applied to RW in Eq. (67), this term can be interpreted as  $\xi(r)v_{12}(r)$  which is proportional to  $b^3$ . This is an artificial consequence of the fact that  $v_{12}$  is exponentiated.

## V. THEORY: IMPACT OF RSD ON THE PROJECTED CLUSTERING

So far I have considered the RSD in the 3D clustering and shown that RSD imprints a characteristic signal along LOS. However, it is common to project the density field onto the 2D sky and to measure the angular power spectrum especially in the case of imaging surveys due to the less accuracy in photometric redshifts. The simple and non-trivial (at least to me) question is whether RSD has an impact on the projected angular power spectrum. Here I briefly address this question, following [5, 54].

The 2D projected density field is written as

$$1 + \delta_{2\text{D}}^s(\hat{\mathbf{n}}) = \int ds \Pi(s) \{1 + \delta^s(s, \hat{\mathbf{n}})\} = \int dr \Pi(r) \{1 + \delta(r, \hat{\mathbf{n}})\}, \quad (72)$$

where  $\Pi(r)$  is the normalized radial selection function (often written as  $dN(z)/dz$ ) such that  $\int dr \Pi(r) = 1$ . As is clearly seen, there is already a notable difference compared to the 3D case. Namely, the Jacobian of real-to-redshift-space mapping is cancelled out in the integral and the 2D density field is affected by RSD only through the radial selection function. Therefore it is much easier to handle the RSD correction in the 2D case.

Taylor-expanding the radial selection function yields to

$$\Pi(s) \simeq \Pi(r) - f \frac{\partial \Pi}{\partial r} \{ \mathbf{u}(\mathbf{r}, \hat{\mathbf{n}}) \cdot \hat{\mathbf{n}} \} + \frac{f^2}{2} \frac{\partial^2 \Pi}{\partial r^2} \{ \mathbf{u}(\mathbf{r}, \hat{\mathbf{n}}) \cdot \hat{\mathbf{n}} \}^2 + \dots \quad (73)$$

Now it is obvious the the expansion parameter is the balance between the characteristic displacement by the velocity field and the slice width. The 1st term is the real-space part which is usually considered:

$$\begin{aligned} \delta_\ell &= \frac{1}{2} \int_{-1}^1 d\mu \delta_{2\text{D}}(\hat{\mathbf{n}}) \mathcal{L}_\ell(\mu) \\ &= \frac{1}{2} \int_{-1}^1 d\mu \left[ \int dr \Pi(r) \int \frac{d^3 k}{(2\pi)^3} \delta(\mathbf{k}; r) \sum_L (-i)^L (2L + 1) j_L(kr) \mathcal{L}_L(\mu) \right] \mathcal{L}_\ell(\mu) \\ &= (-i)^\ell \int \frac{d^3 k}{(2\pi)^3} \delta(\mathbf{k}) W_\ell(k), \end{aligned} \quad (74)$$

where  $W_\ell(k)$  is the radial window function which describes the projected weight along the LOS direction,

$$W_\ell(k) \equiv \int dr \Pi(r) j_\ell(kr). \quad (75)$$

Similar calculations show that the 2nd term that is the leading-order RSD correction is described as

$$\theta_{2D}(\hat{\mathbf{n}}) \equiv -f \int dr \frac{\partial \Pi}{\partial r} \{\mathbf{u} \cdot \hat{\mathbf{n}}\}, \quad (76)$$

and then

$$\theta_\ell = (-i)^\ell f \int \frac{d^3 k}{(2\pi)^3} \theta(\mathbf{k}) W_\ell^I(k), \quad (77)$$

$$W_\ell^I(k) = \int dr \Pi(r) \left[ \frac{(2\ell^2 + 2\ell - 1)}{(2\ell + 3)(2\ell - 1)} j_\ell(kr) - \frac{\ell(\ell - 1)}{(2\ell - 1)(2\ell + 1)} j_{\ell-2}(kr) - \frac{(\ell + 1)(\ell + 2)}{(2\ell + 3)(2\ell + 1)} j_{\ell+2}(kr) \right]. \quad (78)$$

Thus the angular power spectrum with the linear RSD correction included is derived as [54]

$$C_\ell^{\text{1st RSD}} = \frac{2}{\pi} \int k^2 dk \{ P_{\delta\delta}(k) W_\ell(k)^2 + 2f P_{\delta\theta}(k) W_\ell(k) W_\ell^I(k) + f^2 P_{\theta\theta}(k) W_\ell^I(k)^2 \}. \quad (79)$$

It is useful to compare this expression with the Kaiser formula, Eq. (30). The different  $\mu^2$  dependence of the anisotropic terms is described by the different projection effect with the different radial window function. In fact it is straightforward to derive the 2nd-order RSD correction terms [5],

$$\Delta C_\ell^{\text{2nd RSD}} = \frac{2}{\pi} \int k^2 dk [2fQ(k)W_\ell(k)W_\ell^I(k) + f^2R(k)W_\ell(k)W_\ell^I(k) + f^2S(k)W_\ell(k)W_\ell^{II}(k) + f^2T(k)W_\ell^I(k)^2] \quad (80)$$

where the functions,  $Q(k)$ ,  $R(k)$ ,  $S(k)$ , and  $T(k)$  are correction terms originating the bispectrum (therefore, in fact, corresponding to  $A(k, \mu)$  term in the TNS model), and the 2nd-order radial window function is given by

$$\begin{aligned} W_\ell^{II}(k) = & \int dr \Pi(r) \left[ \frac{\ell(\ell - 3)(\ell - 2)(\ell - 1)}{(2\ell - 5)(2\ell - 3)(2\ell - 1)(2\ell + 1)} j_{\ell-4}(kr) - \frac{2\ell(\ell - 1)(2\ell^2 - 2\ell - 7)}{(2\ell - 5)(2\ell - 1)(2\ell + 1)(2\ell + 3)} j_{\ell-2}(kr) \right. \\ & + \frac{3(2\ell^4 + 4\ell^3 - 6\ell^2 - 8\ell + 3)}{(2\ell - 3)(2\ell - 1)(2\ell + 3)(2\ell + 5)} j_\ell(kr) - \frac{2(\ell + 1)(\ell + 2)(2\ell^2 + 6\ell - 3)}{(2\ell - 1)(2\ell + 1)(2\ell + 3)(2\ell + 7)} j_{\ell+2}(kr) \\ & \left. + \frac{(\ell + 1)(\ell + 2)(\ell + 3)(\ell + 4)}{(2\ell + 1)(2\ell + 3)(2\ell + 5)(2\ell + 7)} j_{\ell+4}(kr) \right]. \quad (81) \end{aligned}$$

In the following, I examine the impact of RSD on the angular power spectrum assuming

$$\Pi(z) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp \left\{ -\frac{(z - z_*)^2}{2\sigma_z^2} \right\}, \quad (82)$$

where  $\sigma_z = \sigma_{z0}(1 + z_*)$  and we set  $\sigma_{z0} = 0.04$  unless specifically quoted.

- How strong is the projection effect in each radial window function? See Fig. 10.
- How significant are the RSD correction terms? See Fig. 11.
- Take-home message:

*The angular power spectrum is NOT a projected version of the 3D redshift-space power spectrum!* The Kaiser-like enhancement cannot be ignored at large scales, and nonlinear RSD corrections could add small contributions at mildly nonlinear regime. Much easier than the 3D case to handle the RSD correction as long as the width of slice is sufficiently large.

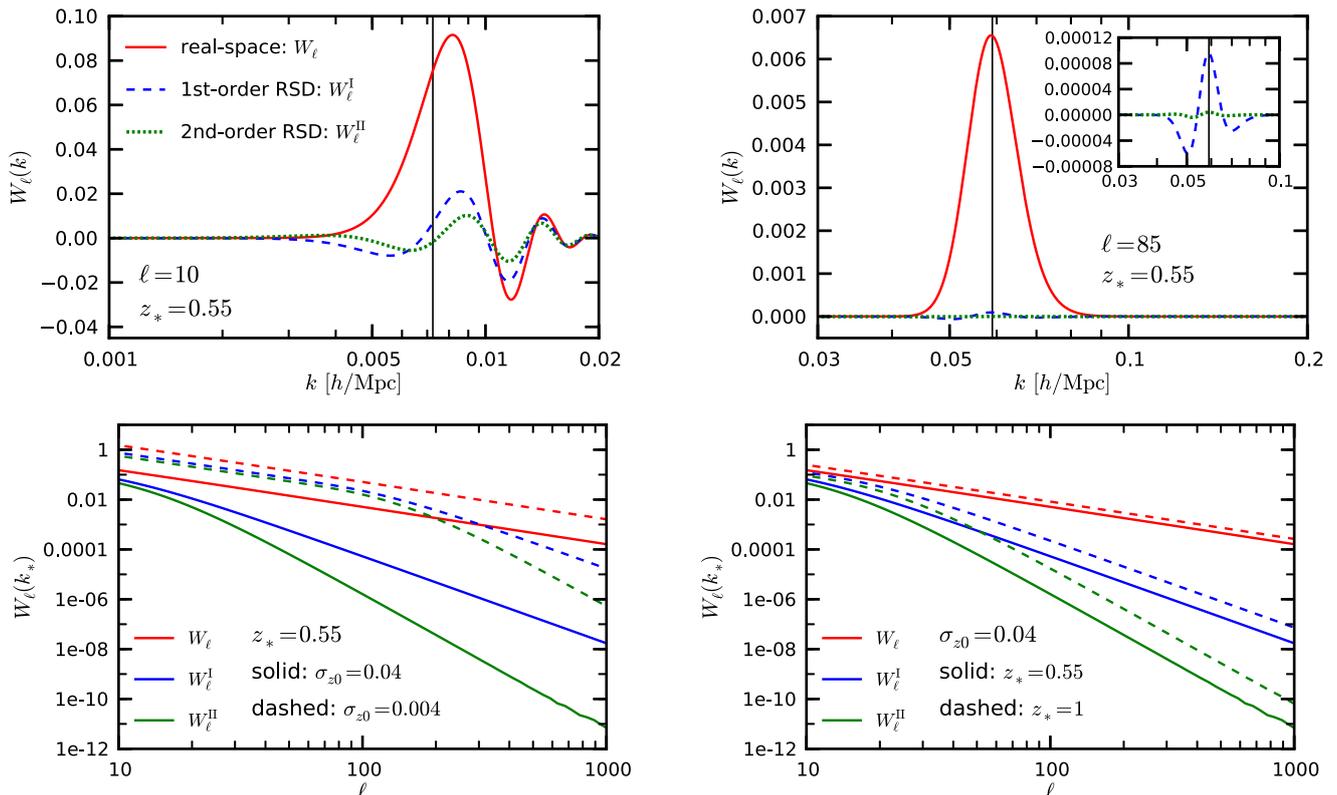


FIG. 10: (Upper panels) The radial window functions at  $\ell = 10$  (left) and  $\ell = 85$  (right) at  $z_* = 0.55$ . We show 3 radial window functions as the real-space part (solid red), the 1st-order RSD (dashed blue), and the 2nd-order RSD (green dotted). At  $\ell = 85$ , small RSD ones are zoomed up for clarification purpose. Vertical thin lines indicate  $k_* = (\ell + 1/2)/y_*$  corresponding to the scale where the Limber approximation is estimated.

(Lower panels) Comparison of radial window functions amplitude at the scale  $k_* = (\ell + 1/2)/y_*$  corresponding to peak of radial window functions. 3 radial window functions are shown as the real-space part (red), the 1st-order RSD (blue), and the 2nd-order RSD (green). The amplitude shown here partly explains to what extent projection to the 2D-sky suppress the density or the velocity fields. Hence we expect the amplitude in the case of thinner slices becomes bigger, which can be indeed confirmed from the 10 times thinner case drawn with dashed lines. The ratios between real-space and RSD correction part, i.e.,  $W_\ell^I/W_\ell$  or  $W_\ell^{II}/W_\ell$ , are apparently smaller for thinner slice.

## VI. ANALYSIS (THEORY): QUANTIFYING THE RSD INFORMATION IN AN IDEAL SURVEY

Let me go back to RSD in the 3D case. Once one is convinced that the TNS model is an okay description of the nonlinear RSD, it is interesting to ask the following questions:

- How well can we measure the anisotropic power spectrum, and hence constrain  $f\sigma_8$  etc given a galaxy redshift survey?
- What is the efficient way to compress the data to fully extract cosmological information on RSD? For example, how many multipoles are necessary in nonlinear regime?

These questions can be answered (not perfectly, though) within a theoretical framework by combining a simple calculation of the power spectrum covariance with the Fisher matrix formalism. Here I summarize our findings in our paper, Taruya, Saito, Nishimichi (2011) [4]. One may find similar efforts to address these questions in the literature (see e.g., [55–58]).

### The Alcock-Paczynski effect

Before proceeding the Fisher forecast, I discuss another source of the anisotropy. In making a 3D map of galaxies, we should assume cosmology to convert redshift to radial comoving distance. Then the measured

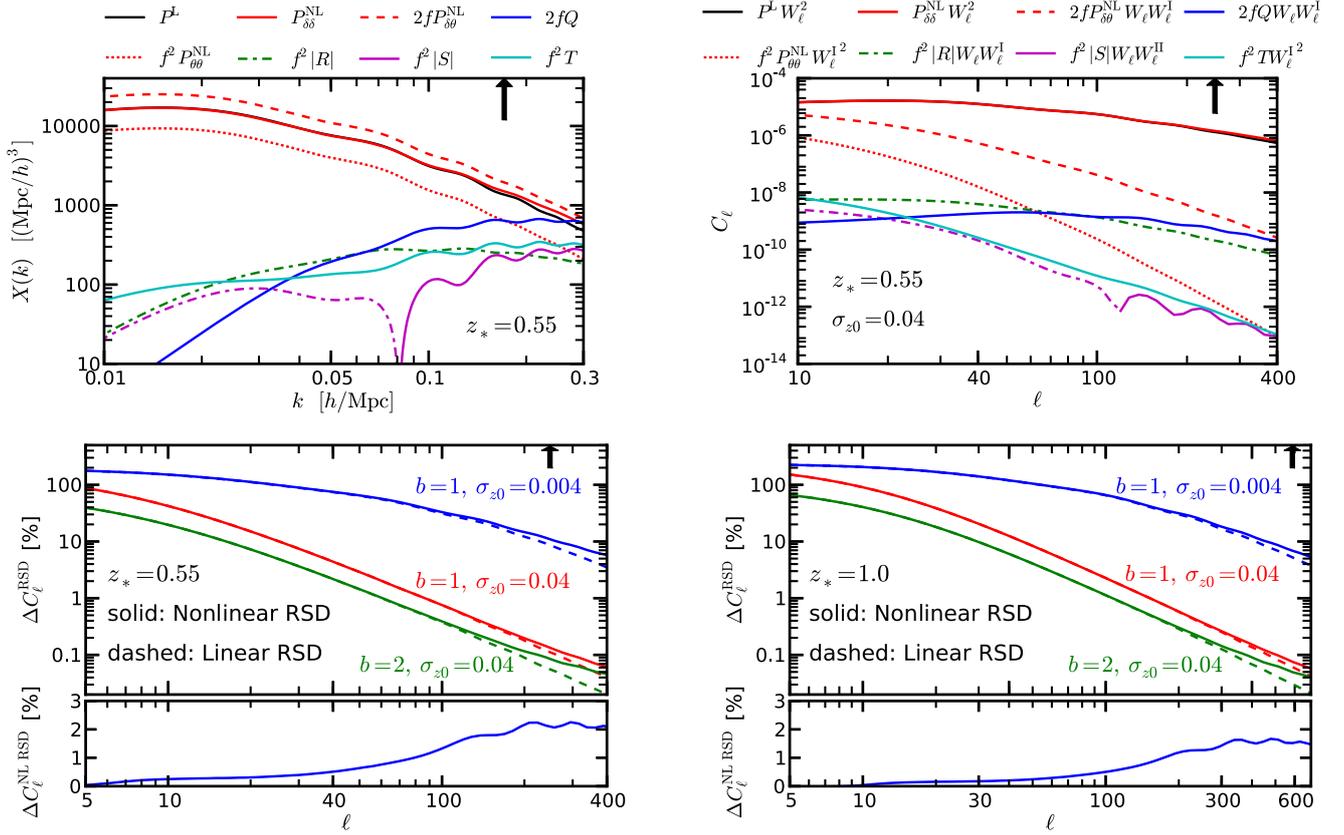


FIG. 11: (Upper left panel) Comparisons of the 3D power spectra which appear in Eqs. (79) and (80). (Upper right panel) Contributions from each term in the angular power spectrum in Eqs. (79) and (80). (Lower panels) The fractional contributions from the RSD correction terms.

distance assuming wrong cosmology can differ from true distance scale. This is the so-called *Alcock-Paczynski (AP) effect* which makes the density distribution anisotropic, since the distortion of the scale perpendicular to LOS is proportional to the angular diameter distance,  $D_A(z)$ , while the distortion of the scale along LOS is proportional to the Hubble distance,  $c/H(z)$  [59, 60]. More explicitly, the measured scale,  $(k_{\parallel}, k_{\perp})$ , is related to the true distance scale  $(q_{\parallel}, q_{\perp})$  through  $k_{\perp} D_A^{\text{fid}} = q_{\perp} D_A$  and  $k_{\parallel}/H^{\text{fid}} = q_{\parallel}/H$ . Therefore, the observed power spectrum is rewritten as

$$P^{\text{obs}}(k, \mu) = \frac{H(z)}{H^{\text{fid}}(z)} \left[ \frac{D_A^{\text{fid}}(z)}{D_A(z)} \right]^2 P^{\text{true}}(q, \nu), \quad (83)$$

where

$$q \equiv \sqrt{q_{\parallel}^2 + q_{\perp}^2} = k \left[ \left( \frac{D_A^{\text{fid}}}{D_A} \right)^2 + \left\{ \left( \frac{H}{H^{\text{fid}}} \right) - \left( \frac{D_A^{\text{fid}}}{D_A} \right)^2 \right\} \mu^2 \right]^{1/2}, \quad (84)$$

$$\nu \equiv \frac{q_{\parallel}}{q} = \left( \frac{H}{H^{\text{fid}}} \right) \mu \left[ \left( \frac{D_A^{\text{fid}}}{D_A} \right)^2 + \left\{ \left( \frac{H}{H^{\text{fid}}} \right) - \left( \frac{D_A^{\text{fid}}}{D_A} \right)^2 \right\} \mu^2 \right]^{-1/2}. \quad (85)$$

Note that the AP effect makes the clustering anisotropic even without RSD. Ref. [61] shows that the isotropic part (i.e., monopole) constrains the dilation parameter,  $\alpha \propto (D_A^2/H)^{1/3}$ , and the anisotropic components constrain the deformation parameter,  $\epsilon \propto D_A H$ . The main cosmological interests in modern galaxy surveys are BAO and RSD, and the parameter constraints are presented in combination of  $(D_A, H, f\sigma_8)$  or  $(\alpha, \epsilon, f\sigma_8)$ .

Gaussian covariance of the redshift-space power spectrum

The statistical error of the redshift-space power spectrum, i.e., the covariance, is given by [18]

$$\text{Cov}[P_g^s(k, \mu), P_g^s(k', \mu')] = [\Delta P_g^s(k, \mu)]^2 \delta_D(\mathbf{k} - \mathbf{k}') = \frac{2}{N_k} \left[ P_g^s(k, \mu) + \frac{1}{\bar{n}_g} \right]^2 \delta_D(\mathbf{k} - \mathbf{k}'), \quad (86)$$

where  $\bar{n}_g$  is the mean number density of galaxies (assumed to be constant here) and the term  $1/\bar{n}_g$  is the Poisson shot noise. The factor  $N_k$  is the number of Fourier modes in a given survey volume  $V_s$ , given by

$$N_k = 2\pi k^2 \Delta k \Delta \mu \left( \frac{2\pi}{V_s^{1/3}} \right)^{-3} = \frac{k^2 \Delta k \Delta \mu}{4\pi^2} V_s. \quad (87)$$

Strictly speaking, this expression holds only for the Gaussian density field which is not true for the galaxy density field in redshift space. This is one of the reasons why the covariance matrix should be estimated with many realizations of realistic mock galaxy catalog. Nevertheless, Eq. (86) is still useful to study feasibility of a given galaxy survey. The cumulative signal-to-noise ratio of the redshift-space power spectrum is then evaluated as

$$\left( \frac{S}{N} \right)_{k < k_{\max}}^2 = \frac{V_s}{8\pi^2} \int_{k_{\min}}^{k_{\max}} k^2 dk \int_{-1}^1 d\mu \left[ \frac{\bar{n}_g P_g^s(k, \mu)}{\bar{n}_g P_g^s(k, \mu) + 1} \right]^2. \quad (88)$$

Therefore the Fisher matrix in Gaussian likelihood is simply given by (e.g., [62])

$$F_{\alpha\beta} = - \left\langle \frac{\partial^2 \ln L}{\partial p_\alpha \partial p_\beta} \right\rangle = \frac{V_s}{8\pi^2} \int_{k_{\min}}^{k_{\max}} k^2 dk \int_{-1}^1 d\mu \frac{\partial \ln P_g^s(k, \mu)}{\partial p_\alpha} \frac{\partial \ln P_g^s(k, \mu)}{\partial p_\beta} \left[ \frac{\bar{n}_g P_g^s(k, \mu)}{\bar{n}_g P_g^s(k, \mu) + 1} \right]^2. \quad (89)$$

The Cramer-Rao bound tells that the  $1\sigma$  error on  $p_\alpha$  marginalized over the other parameters is computed by  $\sigma(p_\alpha)^2 = F_{\alpha\alpha}^{-1}$ . Similarly, the 2D error contours can be estimated by the inverse submatrix of  $F^{-1}$  (but, when plotting error contours, be aware that the value of  $\Delta\chi^2$  in  $1\sigma$  confidence region is not 1.0 but actually 2.3 in the 2D case. See e.g., [63].) Also, I define the Figure-of-Merit (FoM) parameter as an area of the error ellipse in  $n$ -dimensional space,

$$\text{FoM}(p_1, p_2, \dots, p_n) \equiv \frac{1}{\sqrt{|\tilde{F}^{-1}|}}, \quad (90)$$

where  $\tilde{F}^{-1}$  is the  $n \times n$  submatrix of the inverse Fisher matrix [4].

So far I have considered the full anisotropic power spectrum,  $P_g^s(k, \mu)$ . This is motivated by the fact that Eq. (86) suggests that the two-point statistics becomes the most diagonal in the case of the full anisotropic power spectrum. However, this is an ideal case, and many realistic conditions (e.g., the survey window function as discussed in next section) make such analysis much more complicated. From a data-compression point view, it would be a better idea to choose a different base statistics. A natural candidate is the multipole power spectrum, since all the cosmological information is encoded in the multipole moment up to  $\ell = 4$  in linear theory. Notice that the multipole power spectrum is no longer diagonal even for the Gaussian approximation [64, 65],

$$\text{Cov}[P_{g,\ell}^s(k), P_{g,\ell'}^s(k')] = \frac{2}{N_k} \text{Cov}'_{\ell\ell'} \delta_D(k-k') = \frac{2}{N_k} \frac{(2\ell+1)(2\ell'+1)}{2} \int_{-1}^1 d\mu \mathcal{L}_\ell(\mu) \mathcal{L}_{\ell'}(\mu) \left[ P_g^s(k, \mu) + \frac{1}{\bar{n}_g} \right]^2 \delta_D(k-k'), \quad (91)$$

where now the number of mode becomes  $N_k = V_s k^2 \Delta k / (2\pi^2)$ . Eq. (91) shows that the error on any order of the multipole has a contribution from the constant shot noise. This means that a higher-order multipole has a lower signal-to-noise ratio, since a higher-order multipole has lower amplitude. The Fisher matrix for the multipole is also written as

$$F_{\alpha\beta}^{\text{multipole}} = \frac{V_s}{4\pi^2} \int_{k_{\min}}^{k_{\max}} k^2 dk \sum_{\ell, \ell'} \frac{\partial P_{g,\ell}^s(k)}{\partial p_\alpha} [\text{Cov}'_{\ell\ell'}]^{-1} \frac{\partial P_{g,\ell'}^s(k)}{\partial p_\beta}. \quad (92)$$

Now it is ready to perform the Fisher forecast. Namely, given a hypothetical galaxy survey with  $(V_s, \bar{n}_g)$  (and redshift range), one can estimate how well the free parameter set,  $p_\alpha = (D_A, H, f, b, \sigma_v)$  is simultaneously constrained. Our finding is summarized as follows:

- What is the contribution from each multipole to constrain  $(D_A, H, b)$ ?

See Fig. 12. The different multipole power spectra contribute to the parameter constraints with different degeneracy directions. Hence combining them is powerful to break such degeneracies.

- How many multipole are necessary to constrain  $(D_A, H, b)$  as well as the full 2D case?

See Fig. 13. A short answer is that *it is still sufficient to measure the multipole up to  $\ell = 4!$*  However, keep in mind that the parameter constraints could be more biased when higher-order multipole is included if an imperfect RSD model is applied.

- What is the best tracer for the purpose of RSD in terms of bias with survey parameter being fixed?

See the right panel of Fig. 13. A higher bias parameter makes the real-space amplitude larger, while the  $\beta = f/b$  parameter smaller and hence the anisotropic part becomes smaller. This detailed balance results in a peak at  $b \sim 1.2$ .

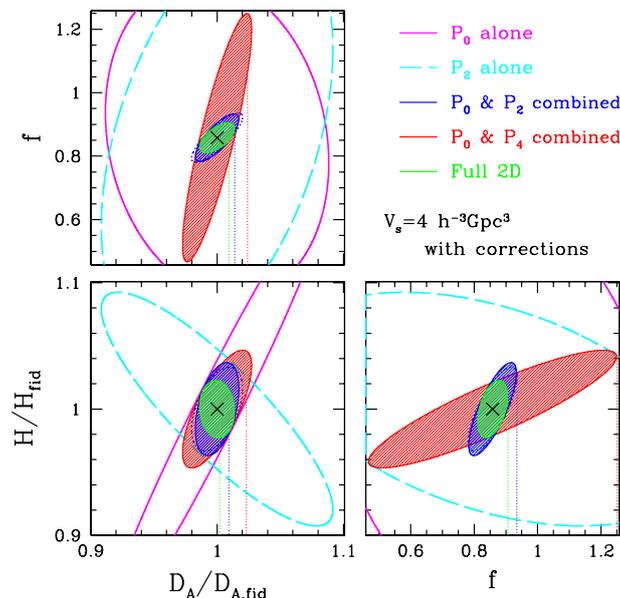


FIG. 12: The marginalized  $1\sigma$  contour in 2D parameter space as a function of  $(D_A, H, f)$ . Here we consider a hypothetical survey with  $z = 1$ ,  $V_s = 4 [(Gpc/h)^3]$ , and  $\bar{n}_g = 5 \times 10^{-4} [(h/Mpc)^3]$ , and consider galaxy samples with  $b = 2$  and  $\sigma_v = 395$  [km/s]. We also set  $k_{\max} = 0.2$  [h/Mpc].

## VII. ANALYSIS: MEASURING RSD FROM THE MULTIPOLE IN BOSS

In previous sections, we developed the refined RSD model (TNS model, but include a bunch of nonlinear galaxy bias term as well in the real data analysis. See Appendix. C for the full expression), and learned how useful the multipole power spectra are. Now it is time to face the real data! Here I present the updated RSD measurement from BOSS DR12 (final dataset!) in SDSS-III [7]. However, even if the galaxy survey is done and its catalog is already available (which is basically a list of  $(ra, dec, z)$ . Of course getting this involves tremendous efforts by many people! I refer to [66] for the galaxy target selection algorithm in BOSS, and to [67, 68] for its stellar-mass completeness using the S82MGC catalog.), there are still several steps to reach the  $f\sigma_8$  constraint as follows:

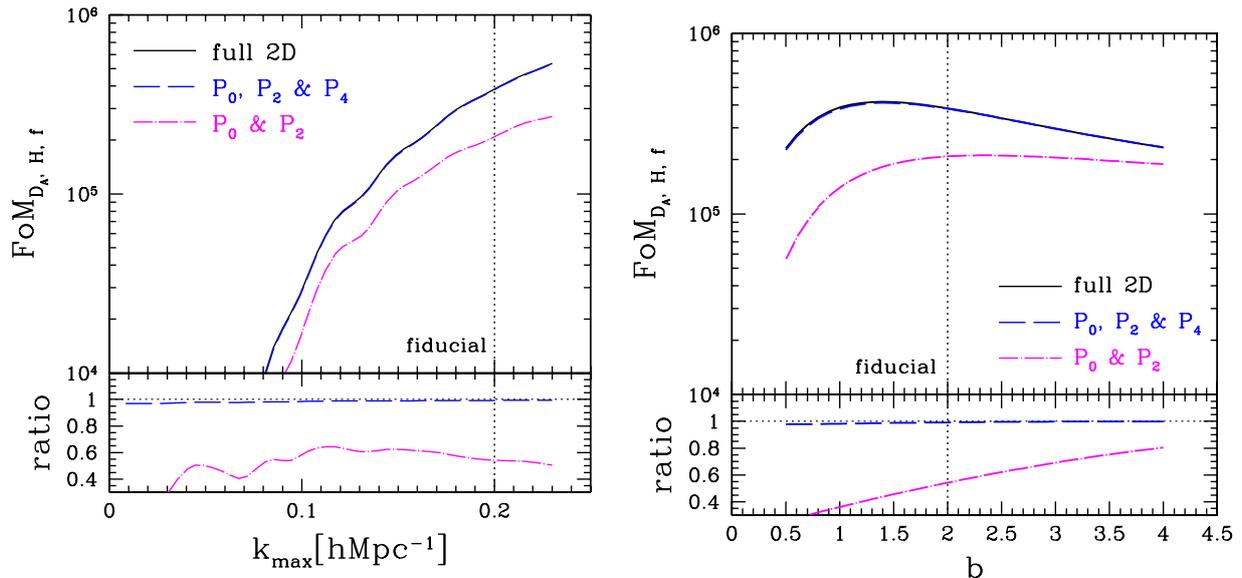


FIG. 13: FoM( $D_A, H, f$ ) with the same hypothetical survey in Fig. 12 but varying parameters such as  $k_{\max}$  (left) and  $b$  (right).

1. First of all, we should measure  $\hat{P}_\ell(k)$  in an unbiased way.
2. In Fourier space, the measured power spectrum is convolved with the survey window function. Therefore we should evaluate the survey window function given a survey geometry.
3. The error covariance matrix,  $\text{Cov}[\hat{P}_\ell(k), \hat{P}_{\ell'}(k')]$  should be estimated by realistic mock catalogs.
4. The theoretical model and every analysis pipeline should be checked against such mock catalogs *before* they are applied to the real data.
5. Finally perform the MCMC parameter estimation to get the  $f\sigma_8$  constraint.

Also I should mention that there are other approaches to constrain RSD with the same BOSS dataset, which include the multipole in configuration space, the wedges in configuration space, the wedges in Fourier space, and combining with the bispectrum etc. I refer to the main alphabetical DR12 paper for the complete reference list. Note that DR12 CMASS and LOWZ papers are already out [69, 70]. In the following, I am going to explain the basic methodology especially in the first three steps in more detail.

### A. The multipole power spectrum estimator

The galaxy power spectrum estimator is first developed by the famous Feldman-Kaiser-Peacock (FKP) paper [18]. Their estimator is designed to implement Fast Fourier Transform (FFT) by assuming the global plain-parallel approximation. As mentioned earlier, however, this is no longer a good approximation when one is interested in measuring the anisotropic component of the galaxy clustering [19, 20]. In order to overcome this, Yamamoto [64] extended the FKP estimator to the multipole by assuming the local plain-parallel approximation in turn:

$$\hat{P}_\ell(\mathbf{k}) = \frac{2\ell + 1}{2A} \left[ \int d^3x \int d^3x' e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} F(\mathbf{x})F(\mathbf{x}') \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}}_h) - S_\ell \right]. \quad (93)$$

Here the density field,  $F(\mathbf{x})$ , and the normalization,  $A$ , are given by

$$F(\mathbf{x}) = w_{\text{FKP}}(\mathbf{x}) [n'_g(\mathbf{x}) - \alpha n_{\text{rand}}(\mathbf{x})] \quad (94)$$

$$A = \int d^3x [n_g(\mathbf{x})^2 w_{\text{FKP}}(\mathbf{x})]^2, \quad (95)$$

respectively.  $n'_g(\mathbf{x})$  denotes the observed galaxy number density but is corrected by weight function as

$$n'_g(\mathbf{x}) = w_c(\mathbf{x})n_g(\mathbf{x}) \quad (96)$$

$$w_c(\mathbf{x}) = (w_{\text{rf}}(\mathbf{x}) + w_{\text{fc}}(\mathbf{x}) - 1)w_{\text{sys}}(\mathbf{x}), \quad (97)$$

where  $w_{\text{rf}}$ ,  $w_{\text{fc}}$ , and  $w_{\text{sys}}$  are the weights which correct the redshift failure, the fiber collision, and systematics that combines a stellar density and seeing condition, respectively, in the particular case of BOSS.  $w_{\text{FKP}}(\mathbf{x})$  is the weight which makes the variance of measured power spectrum minimum, derived as [18, 64]

$$w_{\text{FKP}}(\mathbf{x}) = \frac{1}{1 + n'_g(\mathbf{x})P_0}. \quad (98)$$

$n_{\text{rand}}(\mathbf{x})$  denotes the random field such that  $\langle n'_g \rangle = \alpha \langle n_{\text{rand}} \rangle$  at a given redshift and  $\alpha$  is the ratio of number of galaxies to number of random particles (usually  $\alpha \sim 1/50$ ). Finally  $S_\ell$  denotes the Poisson shot noise, defined by

$$S_\ell = \int d^3x (1 + \alpha)n'_g(\mathbf{x})w_{\text{FKP}}^2(\mathbf{x})\mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}}). \quad (99)$$

Note that there is a subtlety on whether or not the weighted galaxies missed due to the fiber collision is counted in Poisson statistics, which affects the shot noise definition and the FKP weight (see more discussion in [6]). In the Yamamoto estimator it is further simplified by switching the integral to the sum, i.e.,  $\int d^3x n'_g(\mathbf{x}) \dots \rightarrow \sum_i^{N_g} w_c(\mathbf{x}_i) \dots$  or  $\rightarrow \alpha \sum_i^{N_{\text{rand}}} \dots$ , yielding to

$$\hat{P}_\ell(\mathbf{k}) = \frac{2\ell + 1}{2A} [F_\ell(\mathbf{k})F_0(\mathbf{k})^* - S_\ell], \quad (100)$$

$$\begin{aligned} F_\ell(\mathbf{k}) &= \int d^3x F(\mathbf{x})e^{i\mathbf{k} \cdot \mathbf{x}} \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}}) \\ &= \sum_i^{N_g} w_c(\mathbf{x}_i)w_{\text{FKP}}(\mathbf{x}_i)e^{i\mathbf{k} \cdot \mathbf{x}_i} \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}}_i) - \alpha \sum_j^{N_{\text{rand}}} w_{\text{FKP}}(\mathbf{x}_j)e^{i\mathbf{k} \cdot \mathbf{x}_j} \mathcal{L}_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}}_j), \end{aligned} \quad (101)$$

where the local plain-parallel approximation is adopted. Notice that this estimator has  $\mathbf{k}$  dependence in the integrand and hence FFT cannot be applied. Therefore the Yamamoto estimator has a computing cost of  $\mathcal{O}(N^2)$  but was anyhow used in the DR11 analysis [6].

However, Bianchi et al. (2015) [71] and Scoccimarro (2015) [72] have recently realized that there is actually a way to implement FFT on Eq. (101). The idea is very simple: once the Legendre polynomial is explicitly written down, the  $\mathbf{k}$  dependence in the integrand is factored out and hence FFT can be safely applied. The resultant expressions are

$$\hat{P}_0(\mathbf{k}) = \frac{1}{2A} [F_0(\mathbf{k})F_0(\mathbf{k})^* - S_0], \quad (102)$$

$$\hat{P}_2(\mathbf{k}) = \frac{5}{4A} F_0(\mathbf{k}) [3F_2(\mathbf{k})^* - F_0(\mathbf{k})^*], \quad (103)$$

$$\hat{P}_4(\mathbf{k}) = \frac{9}{16A} F_0(\mathbf{k}) [35F_4(\mathbf{k})^* - 30F_2(\mathbf{k})^* + 3F_0(\mathbf{k})^*], \quad (104)$$

where

$$F_0(\mathbf{k}) = A_0(\mathbf{k}), \quad (105)$$

$$F_2(\mathbf{k}) = \frac{1}{k^2} \sum_{p,q=x,y,z} k_p k_q B_{pq}, \quad (106)$$

$$F_4(\mathbf{k}) = \frac{1}{k^4} \sum_{p,q,r=x,y,z} k_p^2 k_q k_r C_{pqr}, \quad (107)$$

and each coefficient can be estimated with FFT,

$$A_0(\mathbf{k}) = \int d^3r F(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (108)$$

$$B_{ij}(\mathbf{k}) = \int d^3r \frac{r_i r_j}{r^2} F(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (109)$$

$$C_{ijk}(\mathbf{k}) = \int d^3r \frac{r_i^2 r_j r_k}{r^4} F(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (110)$$

Now the estimator has a computational complexity of  $\mathcal{O}(N_c \log N_c)$  with  $N_c$  being the number of grid cells, and is adopted in the DR12 analysis [7]. Finally, the multipole power spectrum is evaluated by averaging over spherical shell in Fourier space,

$$\hat{P}_\ell(k) = \langle \hat{P}_\ell(\mathbf{k}) \rangle = \frac{1}{N_{\text{mode}}} \sum_{\mathbf{k} \text{ bin}} \hat{P}_\ell(\mathbf{k}). \quad (111)$$

## B. The survey window function

Because Fourier transform involves the integral over the infinite space while we can observe only finite volume, the measured power spectrum is always convolved with the survey window function,

$$P^{\text{conv}}(\mathbf{k}) = \int d^3k' P^{\text{true}}(\mathbf{k}') |W(\mathbf{k} - \mathbf{k}')|^2 - \frac{|W(\mathbf{k})|^2}{|W(0)|^2} \int d^3k' P^{\text{true}}(\mathbf{k}') |W(\mathbf{k}')|^2, \quad (112)$$

where the 2nd term is the so-called integral constraint which ensures that  $P^{\text{conv}} = 0$  at  $\mathbf{k} \rightarrow 0$ . This means that one needs to estimate the survey window function a priori and to convolve with the theoretical model spectrum. One immediately sees from Eq. (112) that it is complicated to estimate the window function which is as a function of  $\mathbf{k}$  and  $\mathbf{k}'$ . However, we found in the DR11 paper the way to handle this issue on the basis of multipole [6]. With a bit lengthy calculation one obtain a simplified formula (see [6] for detailed derivation),

$$P_\ell^{\text{conv}}(k) = \int \frac{k'^2 dk'}{2\pi^2} \sum_L P_L^{\text{true}}(k') |W(k, k')|_{\ell L}^2, \quad (113)$$

$$|W(k, k')|_{\ell L}^2 = i^\ell (-i)^L (2\ell + 1) \sum_{\substack{i, j \\ i \neq j}}^{N_{\text{rand}}} w_{\text{FKP}}(\mathbf{x}_i) w_{\text{FKP}}(\mathbf{x}_j) j_\ell(k|\Delta\mathbf{x}|) j_L(k'|\Delta\mathbf{x}|) \mathcal{L}_\ell(\hat{\mathbf{x}}_h \cdot \Delta\hat{\mathbf{x}}) \mathcal{L}_L(\hat{\mathbf{x}}_h \cdot \Delta\hat{\mathbf{x}}) \quad (114)$$

Here  $\Delta\mathbf{x} \equiv \mathbf{x}_i - \mathbf{x}_j$ . A similar formula can be also found for the integral constraint term. This simplification is one of the reasons why I prefer to work on the multipole in Fourier space. Interestingly, Eq. (114) tells that *monopole cannot be decoupled from other multipole due to the survey geometry and vice versa*. A physical understanding of this is rather clear: anisotropic survey geometry can make the anisotropic contribution as a leakage from monopole, even if there is no RSD.

One downside of this approach turned out to be the fact that it is hard to obtain the well-converging window function for the hexadecapole. To overcome this, we decide to follow the approach recently proposed by [73]. This indeed leads to more stable results of the hexadecapole window function. The basic idea is following:

- Firstly, Fourier transform the model power spectrum to obtain the correlation function,  $\xi_\ell(s)$ .
- Secondly, multiply the window function,  $W_\ell(s)^2$  in configuration space to obtain the ‘convolved’ correlation function,  $\xi_\ell^{\text{conv}}(s)$ . For the explicit expression for the convolution, see [73].
- Finally, Fourier transform the ‘convolved’ correlation function to obtain the convolved power spectrum,  $P_\ell^{\text{conv}}(k)$ .

Here the window function in configuration space is given by

$$W_\ell(s)^2 \propto \sum_{\mu} \sum_{\mathbf{x}_1} \sum_{\mathbf{x}_2} RR(s, \mu) \mathcal{L}(\mu). \quad (115)$$

Although [73] assumes the global plain parallel approximation, we prove in [7] that the same formula can be used under the local plain parallel approximation as well.

### C. Estimating the covariance matrix and fitting procedure

Compared to an ideal setting in the previous section, the observed galaxy density field is not perfectly Gaussian. Also, the modes of interest is  $\mathcal{O}(10\text{-}100\text{ Mpc})$  and comparable to the survey size, and hence internal methods such as jackknife or bootstrap do not work well [74]. Therefore the covariance matrix is directly estimated from a large number of realizations of realistic mock galaxy catalogs. In the case of the BOSS DR12 analysis, we make use of 2048 realizations of the *Multidark Patchy* mock catalogs which is based on combination of approximated but quite fast  $N$ -body simulations and the subhalo abundance matching [75, 76] (however, see [77] for a similar work but with different results). The realistic survey geometry is applied to light-cone output of each realization. Namely, the covariance matrix is simply estimated by

$$\text{Cov}'_{XY} = \frac{1}{N_s - 1} \sum_{n=1}^{N_s} [\hat{\mathbf{P}}_{\ell,n}(X) - \bar{\mathbf{P}}_\ell(X)][\hat{\mathbf{P}}_{\ell,n}(Y) - \bar{\mathbf{P}}_\ell(Y)], \quad (116)$$

where the vector  $\mathbf{P}_\ell$  contains monopole, quadrupole, and hexadecapole. The fiducial fitting range is  $k = 0.01\text{-}k = 0.15\ h/\text{Mpc}$  with  $\Delta k = 0.01$  for monopole and quadrupole (i.e.,  $n_{\text{bin},\ell=0} = n_{\text{bin},\ell=2} = 14$ ) and  $k = 0.01\text{-}k = 0.10\ h/\text{Mpc}$  with  $\Delta k = 0.01$  for hexadecapole (i.e.,  $n_{\text{bin},\ell=4} = 9$ ). Hence the index in the matrix is defined by  $(X, Y) = (n_{\text{bin},\ell}/2 + i, n_{\text{bin},\ell'}/2 + j)$  which describes the covariance between  $P_\ell(k_i)$  and  $P_{\ell'}(k_j)$ . Note that estimating the covariance matrix with the finite-volume simulation could be underestimated due to the so-called super sample mode [78, 79]. The parameter fitting is performed by minimizing

$$\chi^2 = \Delta \mathbf{P}_\ell(X) \text{Cov}_{XY}^{-1} \Delta \mathbf{P}_\ell(Y), \quad (117)$$

where  $\Delta \mathbf{P}_\ell(X)$  denotes the difference between measurement and model, and also the Hartlap factor should be multiplied to correct the skewness in the inverse covariance matrix [80],

$$\text{Cov}_{XY}^{-1} = \frac{N_s - n_{\text{bin}}^{\text{tot}} - 2}{N_s - 1} \text{Cov}'_{XY}{}^{-1}. \quad (118)$$

Finally, the error in the covariance matrix should be further multiplied to the variance of the derived parameters by [81, 82]

$$M_1 = \sqrt{\frac{1 + a(n_{\text{bin}}^{\text{tot}} - n_p)}{1 + a + b(n_p + 1)}}, \quad (119)$$

where

$$a = \frac{2}{(N_s - n_{\text{bin}}^{\text{tot}} - 1)(N_s - n_{\text{bin}}^{\text{tot}} - 4)}, \quad (120)$$

$$b = \frac{N_s - n_{\text{bin}}^{\text{tot}} - 2}{(N_s - n_{\text{bin}}^{\text{tot}} - 1)(N_s - n_{\text{bin}}^{\text{tot}} - 4)}. \quad (121)$$

At each redshift bin, we have  $n_{\text{bin}}^{\text{tot}} = 37 \times 2 = 74$  (for NGC and SGC),  $n_p = 11$  which results in a very minor correction,  $M_1 \approx 1.01$ . At each redshift bin, the 11 free-parameter set includes  $(b_1\sigma_8, b_2\sigma_8, N, \sigma_v)$  for NGC and SGC separately, and  $(f\sigma_8, D_A/D_A^{\text{fid}}, H/H^{\text{fid}})$  for common cosmological parameters of interest. See also Appendix. C.

## D. BOSS DR12 Result

See another handout in the lecture. Enjoy the most recent RSD measurement! All the results will appear on arXiv at the end of June in 2016.

## E. The optimal estimator

My final quote in the analysis section is on the optimal estimator. It is shown that the FKP estimator (and hence Yamamoto and its variants as well) is optimal only for the modes much smaller than the survey size ( $k \gg L$ ) and the so-called *quadratic estimator* is optimal otherwise [62]. As far as I know, the exact quadratic estimator has never been successfully applied to the actual galaxy survey, although there are a couple of attempts in an approximated way [83–85]. This means that there is still a way to extract more information from the redshift-space power spectrum even from the same BOSS data, and I hope this is achieved in a near future.

## VIII. CONCLUDING REMARK

RSD is one of the main scientific targets for ongoing and forthcoming cosmological surveys. Here I discuss recent efforts for both modeling and measurement which highlights what I have been involved in this several years. I expect there will be substantial progresses in many aspects in this field in coming years. Even though there are tons of topics which I cannot cover in this lecture, I hope this is helpful for you to learn something about RSD. At least this note should be quite helpful to remind me a lot of things!

### Acknowledgments

I acknowledge Eiichiro Komatsu for giving me such an excellent opportunity. I also thank my wonderful collaborators, especially, Atsushi Taruya, Takahiro Nishimichi, and Florian Beutler, who have led most of the works presented in this note.

### Appendix A: Convention and useful formula

#### Cosmology

Hubble equation in a  $\Lambda$ CDM universe

$$H(z)^2 = H_0^2 \{ \Omega_{m0}(1+z)^3 + \Omega_\Lambda \}. \quad (\text{A1})$$

Friedman-Robertson-Walker (FRW) metric

$$ds^2 = a(\tau)^2 \{ -(1+2\Psi)d\tau^2 + (1-2\Phi)d\mathbf{x}^2 \}. \quad (\text{A2})$$

The amplitude of matter fluctuation is often characterized by

$$\sigma_8^2(z) = \int \frac{k^2 dk}{2\pi^2} P_m^L(k; z) W_8(k)^2, \quad (\text{A3})$$

where  $W_8(k)$  is the Fourier transform of the top-hat window function of width 8 Mpc/ $h$ .

#### Mathematics

Fourier transformation

$$A(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} A(\mathbf{k}). \quad (\text{A4})$$

$$A(\mathbf{k}) = \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} A(\mathbf{x}). \quad (\text{A5})$$

the Legendre polynomial:

$$\mathcal{L}_0(\mu) = 1, \quad (\text{A6})$$

$$\mathcal{L}_2(\mu) = \frac{3\mu^2 - 1}{2}, \quad (\text{A7})$$

$$\mathcal{L}_4(\mu) = \frac{35\mu^4 - 30\mu^2 + 3}{8}. \quad (\text{A8})$$

And its orthogonality

$$\frac{2\ell + 1}{2} \int_{-1}^1 d\mu \mathcal{L}_\ell(\mu) \mathcal{L}_{\ell'}(\mu) = \delta_{\ell\ell'}. \quad (\text{A9})$$

The partial-wave expansion

$$e^{i\mathbf{k}\cdot\hat{\mathbf{r}}r} = \sum_L i^L (2L + 1) j_L(kr) \mathcal{L}_L(\mu). \quad (\text{A10})$$

Recursion relation of the spherical Bessel function

$$\ell j_{\ell-1}(x) - (\ell + 1) j_{\ell+1}(x) = (2\ell + 1) j_\ell(x)'. \quad (\text{A11})$$

### Statistics

Moments and cumulants of the one-point distribution function [2]:

$$\langle \delta \rangle_c = \langle \delta \rangle \rightarrow 0, \quad (\text{A12})$$

$$\langle \delta^2 \rangle_c = \sigma^2 = \langle \delta^2 \rangle - \langle \delta \rangle_c^2 \rightarrow \langle \delta^2 \rangle, \quad (\text{A13})$$

$$\langle \delta^3 \rangle_c = \langle \delta^3 \rangle - 3\langle \delta \rangle_c^2 \langle \delta \rangle_c - \langle \delta \rangle_c^3 \rightarrow \langle \delta^3 \rangle, \quad (\text{A14})$$

$$\langle \delta^4 \rangle_c = \langle \delta^4 \rangle - 4\langle \delta \rangle_c^3 \langle \delta \rangle_c - 3\langle \delta^2 \rangle_c^2 - 6\langle \delta^2 \rangle_c \langle \delta \rangle_c^2 - \langle \delta \rangle_c^4 \rightarrow \langle \delta^4 \rangle - 3\sigma^4, \quad (\text{A15})$$

where the case with  $\langle \delta \rangle = 0$  is shown after the arrow.

## Appendix B: Perturbation Theory basics

In this appendix we summarize basic equations in perturbation theory, taken over from [24].

### 1. Matter density

A matter density in Fourier space is perturbatively expanded into

$$\begin{aligned} \delta_m(\mathbf{k}) &= \delta_0(\mathbf{k}) \\ &+ \int \frac{d^3q}{(2\pi)^3} F_S^{(2)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \delta_0(\mathbf{q}) \delta_0(\mathbf{k} - \mathbf{q}) \\ &+ \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} F_S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\ &+ \mathcal{O}(\delta_0^4), \end{aligned} \quad (\text{B1})$$

where  $\delta_0$  is the linear density perturbation and the symmetrized PT kernels are given by

$$\begin{aligned} F_S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) &= \frac{1}{2} \left\{ F^{(2)}(\mathbf{q}_1, \mathbf{q}_2) + F^{(2)}(\mathbf{q}_2, \mathbf{q}_1) \right\} \\ &= \frac{5}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \left( \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \right)^2, \end{aligned} \quad (\text{B2})$$

$$G_S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{3}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{4}{7} \left( \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \right)^2, \quad (\text{B3})$$

$$\begin{aligned} F_S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) &= \frac{1}{3!} \left\{ F^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + \text{cyclic} \right\} \\ &= \frac{1}{6} \left[ \frac{7}{9} \frac{\mathbf{q}_{123} \cdot \mathbf{q}_3}{q_3^2} F_S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) + \left\{ \frac{7}{9} \frac{\mathbf{q}_{123} \cdot (\mathbf{q}_1 + \mathbf{q}_2)}{|\mathbf{q}_1 + \mathbf{q}_2|^2} + \frac{2}{9} \frac{q_{123}^2 \mathbf{q}_3 \cdot (\mathbf{q}_1 + \mathbf{q}_2)}{|\mathbf{q}_1 + \mathbf{q}_2|^2 \cdot q_3^2} \right\} G_S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) \right] \\ &\quad + \text{cyclic}, \end{aligned} \quad (\text{B4})$$

$$\begin{aligned} G_S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) &= \frac{1}{6} \left[ \frac{1}{3} \frac{\mathbf{q}_{123} \cdot \mathbf{q}_3}{q_3^2} F_S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) + \left\{ \frac{1}{3} \frac{\mathbf{q}_{123} \cdot (\mathbf{q}_1 + \mathbf{q}_2)}{|\mathbf{q}_1 + \mathbf{q}_2|^2} + \frac{2}{3} \frac{q_{123}^2 \mathbf{q}_3 \cdot (\mathbf{q}_1 + \mathbf{q}_2)}{|\mathbf{q}_1 + \mathbf{q}_2|^2 \cdot q_3^2} \right\} G_S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) \right] \\ &\quad + \text{cyclic}, \end{aligned} \quad (\text{B5})$$

where  $\mathbf{q}_{123} = \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$ . The unsymmetrized kernels are given by

$$F^{(2)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{5}{7} \alpha(\mathbf{q}_1, \mathbf{q}_2) + \frac{2}{7} \beta(\mathbf{q}_1, \mathbf{q}_2), \quad (\text{B6})$$

$$G^{(2)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{3}{7} \alpha(\mathbf{q}_1, \mathbf{q}_2) + \frac{4}{7} \beta(\mathbf{q}_1, \mathbf{q}_2), \quad (\text{B7})$$

$$\alpha(\mathbf{q}_1, \mathbf{q}_2) = \frac{(\mathbf{q}_1 + \mathbf{q}_2) \cdot \mathbf{q}_1}{q_1^2}, \quad (\text{B8})$$

$$\beta(\mathbf{q}_1, \mathbf{q}_2) = \frac{1}{2} (\mathbf{q}_1 + \mathbf{q}_2)^2 \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1^2 q_2^2}. \quad (\text{B9})$$

## 2. Biased tracer's density

Following an ansatz in McDonald & Roy (2010) [21], a halo density field (or generally biased tracer) is written as

$$\begin{aligned} \delta_h(\mathbf{x}) &= c_\delta \delta_m(\mathbf{x}) \\ &\quad + \frac{1}{2} c_{\delta^2} \delta_m(\mathbf{x})^2 + \frac{1}{2} c_{s^2} s(\mathbf{x})^2 \\ &\quad + \frac{1}{3!} c_{\delta^3} \delta_m(\mathbf{x})^3 + \frac{1}{2} c_{\delta s^2} \delta_m(\mathbf{x}) s(\mathbf{x})^2 + c_\psi \psi(\mathbf{x}) + c_{st} s(\mathbf{x}) t(\mathbf{x}) + \frac{1}{3!} c_{s^3} s(\mathbf{x})^3 \\ &\quad + c_\epsilon \epsilon + \dots, \end{aligned} \quad (\text{B10})$$

where each independent variable is defined as

$$s_{ij}(\mathbf{x}) \equiv \partial_i \partial_j \phi(\mathbf{x}) - \frac{1}{3} \delta_{ij}^K \delta_m(\mathbf{x}) = \left[ \partial_i \partial_j \partial^{-2} - \frac{1}{3} \delta_{ij}^K \right] \delta_m(\mathbf{x}), \quad (\text{B11})$$

$$t_{ij}(\mathbf{x}) \equiv \partial_i v_j - \frac{1}{3} \delta_{ij}^K \theta_m(\mathbf{x}) - s_{ij}(\mathbf{x}) = \left[ \partial_i \partial_j \partial^{-2} - \frac{1}{3} \delta_{ij}^K \right] [\theta(\mathbf{x}) - \delta_m(\mathbf{x})], \quad (\text{B12})$$

$$\psi(\mathbf{x}) \equiv [\theta(\mathbf{x}) - \delta_m(\mathbf{x})] - \frac{2}{7} s(\mathbf{x})^2 + \frac{4}{21} \delta_m(\mathbf{x})^2. \quad (\text{B13})$$

Note that  $t_{ij}$  is zero at first order, and  $\psi$  is zero up to second order. In Fourier space, the halo density contrast is given by

$$\begin{aligned}
\delta_h(\mathbf{k}) = & c_\delta \delta_0(\mathbf{k}) \\
& + c_\delta \int \frac{d^3q}{(2\pi)^3} F_S^{(2)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \delta_0(\mathbf{q}) \delta_0(\mathbf{k} - \mathbf{q}) \\
& + \frac{1}{2} c_{\delta^2} \int \frac{d^3q}{(2\pi)^3} \delta_0(\mathbf{q}) \delta_0(\mathbf{k} - \mathbf{q}) \\
& + \frac{1}{2} c_{s^2} \int \frac{d^3q}{(2\pi)^3} S^{(2)}(\mathbf{q}, \mathbf{k} - \mathbf{q}) \delta_0(\mathbf{q}) \delta_0(\mathbf{k} - \mathbf{q}) \\
& + c_\delta \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} F_S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + c_{\delta^2} \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} F_S^{(2)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + \frac{1}{3!} c_{\delta^3} \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + c_{s^2} \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} S^{(2)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1) F_S^{(2)}(\mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + \frac{1}{3!} c_{s^3} \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + \frac{1}{2} c_{\delta s^2} \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} S^{(2)}(\mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + c_\psi \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} \left\{ D_S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) - 2F_S^{(2)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) D_S^{(2)}(\mathbf{q}_2, \mathbf{k} - \mathbf{q}_2) \right\} \\
& \quad \times \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2) \\
& + 2c_{st} \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} S^{(2)}(\mathbf{q}_1, \mathbf{k} - \mathbf{q}_1) D_S^{(2)}(\mathbf{q}_2, \mathbf{q}_1 - \mathbf{q}_2) \delta_0(\mathbf{q}_1) \delta_0(\mathbf{q}_2) \delta_0(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2), \tag{B14}
\end{aligned}$$

where

$$S^{(2)}(\mathbf{q}_1, \mathbf{q}_2) = \left( \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \right)^2 - \frac{1}{3}, \tag{B15}$$

$$S^{(3)}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)(\mathbf{q}_2 \cdot \mathbf{q}_3)(\mathbf{q}_3 \cdot \mathbf{q}_1)}{q_1^2 q_2^2 q_3^2} - \frac{1}{3} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} - \frac{1}{3} \frac{(\mathbf{q}_2 \cdot \mathbf{q}_3)^2}{q_2^2 q_3^2} - \frac{1}{3} \frac{(\mathbf{q}_3 \cdot \mathbf{q}_1)^2}{q_3^2 q_1^2} + \frac{2}{9}, \tag{B16}$$

$$D^{(N)} \equiv G^{(N)} - F^{(N)}. \tag{B17}$$

### Appendix C: The redshift-space power spectrum model adopted in the actual analysis for BOSS

The model adopted in the BOSS analysis is an extended version of the TNS model, Eq. (50):

$$P_{g,\text{TNS}}^s(k, \mu) = \exp[-k^2 \mu^2 f^2 \sigma_{v,\text{eff}}^2] \left\{ P_{g,\delta\delta}(k) + 2f\mu^2 P_{g,\delta\theta}(k) + f^2 \mu^4 P_{\theta\theta}(k) + b_1^3 A(k, \mu; \beta) + b_1^4 B(k, \mu; \beta) \right\}. \tag{C1}$$

Our galaxy bias model is based on the bias renormalization including local and nonlocal bias term proposed in [21] is validated by [24]:

$$\begin{aligned}
P_{g,\delta\delta}(k) = & b_1^2 P_{\delta\delta}^{NL}(k) + 2b_1 b_2 P_{b_2,\delta}(k) + 2b_1 b_{s^2} P_{b_{s^2},\delta}(k) + 2b_1 b_{3nl} \sigma_3^2(k) P(k) \\
& + b_2^2 P_{b_{22}}(k) + 2b_2 b_{s^2} P_{b_{2s^2}}(k) + b_{s^2}^2 P_{s_{22}}(k) + N, \tag{C2}
\end{aligned}$$

$$\begin{aligned}
P_{g,\delta\theta}(k) = & b_1 P_{\delta\theta}^{NL}(k) + b_2 P_{b_2,\theta}(k) \\
& + b_{s^2} P_{b_{s^2},\theta}(k) + b_{3nl} \sigma_3^2(k) P^L(k). \tag{C3}
\end{aligned}$$

The exact expression of the bias correction terms are redundant here, and I refer to [6]. One approximation which is not exactly true but can reduce the free parameter is the local Lagrangian bias for the nonlocal bias [24]:

$$b_{s^2} \approx -\frac{4}{7}(b_1 - 1), \quad (\text{C4})$$

$$b_{3nl} \approx \frac{32}{315}(b_1 - 1). \quad (\text{C5})$$

Strictly speaking, we should consistently include the bias term up to 2nd order in  $A$  and  $B$  correction terms but simply ignore them here. As a summary, the model include 4 free parameter,  $b_1, b_2, N$ , and  $\sigma_{v,\text{eff}}$ .

#### Appendix D: Derivation of Eq. (17)

Starting from Eq. (16), one finds

$$\begin{aligned} \delta^s(\mathbf{k}) &= \int d^3s \delta^s(\mathbf{s}) e^{i\mathbf{k}\cdot\mathbf{s}} \\ &= \int d^3x e^{i\mathbf{k}\cdot\mathbf{s}} \{1 + \delta(\mathbf{x})\} - \int d^3s e^{i\mathbf{k}\cdot\mathbf{s}} \\ &= \int d^3x e^{i\mathbf{k}\cdot\mathbf{s}} \{1 + \delta(\mathbf{x})\} - \int d^3x \frac{d^3s}{d^3x} e^{i\mathbf{k}\cdot\mathbf{s}} \\ &= \int d^3x e^{i\mathbf{k}\cdot\mathbf{s}} \{1 + \delta(\mathbf{x})\} - \int d^3x \left\{ 1 + \frac{1}{aH} \frac{\partial v_z(\mathbf{x})}{\partial z} \right\} e^{i\mathbf{k}\cdot\mathbf{s}} \\ &= \int d^3x \left\{ \delta(\mathbf{x}) - \frac{1}{aH} \frac{\partial v_z(\mathbf{x})}{\partial z} \right\} e^{i\mathbf{k}\cdot\mathbf{x} + ik\mu v_z/(aH)}. \end{aligned} \quad (\text{D1})$$

- 
- [1] A. J. S. Hamilton, in *The Evolving Universe*, edited by D. Hamilton (1998), vol. 231 of *Astrophysics and Space Science Library*, p. 185, astro-ph/9708102. [I](#)
- [2] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, *Phys. Rep.* **367**, 1 (2002), arXiv:astro-ph/0112551. [I](#), [A](#)
- [3] A. Taruya, T. Nishimichi, and S. Saito, *Phys. Rev. D* **82**, 063522 (2010), 1006.0699. [I](#), [II](#), [IV](#), [5](#), [6](#), [IVB](#), [IVB](#)
- [4] A. Taruya, S. Saito, and T. Nishimichi, *Phys. Rev. D* **83**, 103527 (2011), 1101.4723. [VI](#), [VI](#)
- [5] S. Saito, should have been published before and kept being in preparation (201x). [I](#), [V](#), [V](#)
- [6] F. Beutler, S. Saito, H.-J. Seo, J. Brinkmann, K. S. Dawson, D. J. Eisenstein, A. Font-Ribera, S. Ho, C. K. McBride, F. Montesano, et al., *MNRAS* **443**, 1065 (2014), 1312.4611. [I](#), [VIIA](#), [VIIA](#), [VIIB](#), [VIIB](#), [C](#)
- [7] F. Beutler, H.-J. Seo, and S. Saito et al., in preparation (2016), submitted to the BOSS collaboration. [I](#), [VII](#), [VIIA](#), [VIIB](#)
- [8] N. Kaiser, *MNRAS* **227**, 1 (1987). [II](#), [II](#), [IIIB](#)
- [9] J. C. Jackson, *MNRAS* **156**, 1P (1972). [II](#)
- [10] O. Lahav, P. B. Lilje, J. R. Primack, and M. J. Rees, *MNRAS* **251**, 128 (1991). [II](#)
- [11] W. L. W. Sargent and E. L. Turner, *ApJ* **212**, L3 (1977). [II](#), [2](#)
- [12] T. Okumura, C. Hikage, T. Totani, M. Tonegawa, H. Okada, K. Glazebrook, C. Blake, P. G. Ferreira, S. More, A. Taruya, et al., *PASJ* **68**, 38 (2016), 1511.08083. [II](#), [3](#)
- [13] A. F. Heavens and A. N. Taylor, *MNRAS* **275**, 483 (1995), astro-ph/9409027. [IIA](#)
- [14] A. J. S. Hamilton and M. Culhane, *MNRAS* **278**, 73 (1996), arXiv:astro-ph/9507021. [IIA](#)
- [15] I. Szapudi, *ApJ* **614**, 51 (2004), arXiv:astro-ph/0404477. [IIA](#)
- [16] P. Pápai and I. Szapudi, *MNRAS* **389**, 292 (2008), 0802.2940.
- [17] A. Raccanelli, L. Samushia, and W. J. Percival, *MNRAS* **409**, 1525 (2010), 1006.1652. [IIA](#)
- [18] H. A. Feldman, N. Kaiser, and J. A. Peacock, *ApJ* **426**, 23 (1994), arXiv:astro-ph/9304022. [IIA](#), [VI](#), [VIIA](#), [VIIA](#)
- [19] L. Samushia, W. J. Percival, and A. Raccanelli, *MNRAS* **420**, 2102 (2012), 1102.1014. [IIA](#), [VIIA](#)
- [20] J. Yoo and U. Seljak, *MNRAS* **447**, 1789 (2015), 1308.1093. [IIA](#), [VIIA](#)

- [21] P. McDonald and A. Roy, *J. Cosmology Astropart. Phys.* **8**, 20 (2009), 0902.0991. [II A](#), [B 2](#), [C](#)
- [22] K. C. Chan, R. Scoccimarro, and R. K. Sheth, *Phys. Rev. D* **85**, 083509 (2012), 1201.3614.
- [23] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, *Phys. Rev. D* **86**, 083540 (2012), 1201.4827.
- [24] S. Saito, T. Baldauf, Z. Vlah, U. Seljak, T. Okumura, and P. McDonald, *ArXiv e-prints* (2014), 1405.1447. [B](#), [C](#), [C](#)
- [25] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga (2014), 1412.5169v1.
- [26] F. Schmidt, *ArXiv e-prints* (2015), 1511.02231. [II A](#)
- [27] M. Biagetti, V. Desjacques, A. Kehagias, and A. Riotto, *Phys. Rev. D* **90**, 103529 (2014), 1408.0293. [II A](#)
- [28] T. Baldauf, V. Desjacques, and U. Seljak, *Phys. Rev. D* **92**, 123507 (2015), 1405.5885. [II A](#)
- [29] T. Matsubara, *ArXiv e-prints* (2013), 1304.4226. [II B](#)
- [30] L. Wang, B. Reid, and M. White, *ArXiv e-prints* (2013), 1306.1804. [II B](#)
- [31] N. S. Sugiyama, *ApJ* **788**, 63 (2014), 1311.0725. [II B](#)
- [32] M. Lewandowski, L. Senatore, F. Prada, C. Zhao, and C.-H. Chuang, *ArXiv e-prints* (2015), 1512.06831. [II B](#)
- [33] U. Seljak and P. McDonald, *J. Cosmology Astropart. Phys.* **11**, 039 (2011), 1109.1888. [II B](#)
- [34] T. Okumura, U. Seljak, and V. Desjacques, *J. Cosmology Astropart. Phys.* **11**, 014 (2012), 1206.4070.
- [35] T. Okumura, U. Seljak, P. McDonald, and V. Desjacques, *J. Cosmology Astropart. Phys.* **2**, 010 (2012), 1109.1609.
- [36] T. Okumura, N. Hand, U. Seljak, Z. Vlah, and V. Desjacques, *Phys. Rev. D* **92**, 103516 (2015), 1506.05814.
- [37] Z. Vlah, U. Seljak, P. McDonald, T. Okumura, and T. Baldauf, *J. Cosmology Astropart. Phys.* **11**, 009 (2012), 1207.0839.
- [38] Z. Vlah, U. Seljak, T. Okumura, and V. Desjacques, *J. Cosmology Astropart. Phys.* **10**, 053 (2013), 1308.6294. [II B](#)
- [39] H. Gil-Marín, C. Wagner, J. Noreña, L. Verde, and W. Percival (2014), 1407.1836v1. [II B](#)
- [40] J. Yoo and V. Desjacques, *ArXiv e-prints* (2013), 1301.4501. [II B](#)
- [41] H. Guo, Z. Zheng, I. Zehavi, K. Dawson, R. A. Skibba, J. L. Tinker, D. H. Weinberg, M. White, and D. P. Schneider, *MNRAS* **446**, 578 (2015), 1407.4811. [II B](#)
- [42] W. A. Hellwing, M. Schaller, C. S. Frenk, T. Theuns, J. Schaye, R. G. Bower, and R. A. Crain, *ArXiv e-prints* (2016), 1603.03328.
- [43] B. A. Reid, H.-J. Seo, A. Leauthaud, J. L. Tinker, and M. White, *MNRAS* **444**, 476 (2014), 1404.3742. [II B](#)
- [44] W. J. Percival and M. White, *MNRAS* **393**, 297 (2009), 0808.0003. [III B](#), [IV A](#)
- [45] E. A. Kazin, A. G. Sánchez, and M. R. Blanton, *MNRAS* **419**, 3223 (2012), 1105.2037. [III B](#)
- [46] G. J. Niklas and A. G. Sánchez et al., in preparation (2016), submitted to the BOSS collaboration. [III B](#)
- [47] Y. Zheng and Y.-S. Song, *ArXiv e-prints* (2016), 1603.00101. [IV](#), [IV B](#), [IV B](#), [IV B](#), [IV B](#), [9](#)
- [48] A. F. Heavens, S. Matarrese, and L. Verde, *MNRAS* **301**, 797 (1998), arXiv:astro-ph/9808016. [IV A](#)
- [49] R. Scoccimarro, *Phys. Rev. D* **70**, 083007 (2004), arXiv:astro-ph/0407214. [IV A](#), [IV B](#), [IV C](#)
- [50] A. Taruya, T. Nishimichi, S. Saito, and T. Hiramatsu, *Phys. Rev. D* **80**, 123503 (2009), 0906.0507. [5](#)
- [51] A. Taruya, T. Nishimichi, and F. Bernardeau, *Phys. Rev. D* **87**, 083509 (2013), 1301.3624. [IV B](#), [IV B](#)
- [52] M. Davis and P. J. E. Peebles, *ApJL* **267**, 465 (1983), URL <http://dx.doi.org/10.1086/160884>. [IV C](#)
- [53] B. A. Reid and M. White, *ArXiv e-prints* (2011), 1105.4165. [IV C](#)
- [54] N. Padmanabhan, D. J. Schlegel, and U. Seljak et al., *MNRAS* **378**, 852 (2007), arXiv:astro-ph/0605302. [V](#), [V](#)
- [55] H. Seo and D. J. Eisenstein, *ApJ* **598**, 720 (2003), arXiv:astro-ph/0307460. [VI](#)
- [56] H. Seo and D. J. Eisenstein, *ApJ* **665**, 14 (2007), arXiv:astro-ph/0701079.
- [57] M. Shoji, D. Jeong, and E. Komatsu, *ApJ* **693**, 1404 (2009), 0805.4238.
- [58] M. White, Y. Song, and W. J. Percival, *MNRAS* **397**, 1348 (2009), 0810.1518. [VI](#)
- [59] C. Alcock and B. Paczynski, *Nature* **281**, 358 (1979). [VI](#)
- [60] T. Matsubara and Y. Suto, *ApJ* **470**, L1 (1996), arXiv:astro-ph/9604142. [VI](#)
- [61] N. Padmanabhan and M. White, *Phys. Rev. D* **77**, 123540 (2008), 0804.0799. [VI](#)
- [62] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, *ApJ* **499**, 555 (1998), arXiv:astro-ph/9708020. [VI](#), [VII E](#)
- [63] L. Verde, *ArXiv e-prints* (2007), 0712.3028. [VI](#)
- [64] K. Yamamoto, *ApJ* **595**, 577 (2003), arXiv:astro-ph/0208139. [VI](#), [VII A](#), [VII A](#)
- [65] K. Yamamoto, M. Nakamichi, A. Kamino, B. A. Bassett, and H. Nishioka, *PASJ* **58**, 93 (2006), arXiv:astro-ph/0505115. [VI](#)
- [66] B. Reid, S. Ho, N. Padmanabhan, W. J. Percival, J. Tinker, R. Tojeiro, M. White, D. J. Eisenstein, C. Maraston, A. J. Ross, et al., *MNRAS* **455**, 1553 (2016), 1509.06529. [VII](#)
- [67] K. Bundy, A. Leauthaud, S. Saito, A. Bolton, Y.-T. Lin, C. Maraston, R. C. Nichol, D. P. Schneider, D. Thomas, and D. A. Wake, *ApJS* **221**, 15 (2015), 1509.01276. [VII](#)
- [68] A. Leauthaud, K. Bundy, S. Saito, J. Tinker, C. Maraston, R. Tojeiro, S. Huang, J. R. Brownstein, D. P. Schneider, and D. Thomas, *MNRAS* **457**, 4021 (2016), 1507.04752. [VII](#)
- [69] H. Gil-Marín, W. J. Percival, J. R. Brownstein, C.-H. Chuang, J. N. Grieb, S. Ho, F. Shu Kitaura, C. Maraston, F. Prada, S. Rodríguez-Torres, et al., *MNRAS* (2016), 1509.06386. [VII](#)

- [70] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres, and M. D. Olmstead, ArXiv e-prints (2016), 1606.00439. [VII](#)
- [71] D. Bianchi, H. Gil-Marín, R. Ruggeri, and W. J. Percival, MNRAS **453**, L11 (2015), 1505.05341. [VII A](#)
- [72] R. Scoccimarro, Phys. Rev. D **92**, 083532 (2015), 1506.02729. [VII A](#)
- [73] M. J. Wilson, J. A. Peacock, A. N. Taylor, and S. de la Torre, ArXiv e-prints (2015), 1511.07799. [VIIB](#), [VIIB](#), [VIIB](#)
- [74] P. Norberg, C. M. Baugh, E. Gaztañaga, and D. J. Croton, MNRAS **396**, 19 (2009), 0810.1885. [VIIC](#)
- [75] F.-S. Kitaura, S. Rodríguez-Torres, C.-H. Chuang, C. Zhao, F. Prada, H. Gil-Marín, H. Guo, G. Yepes, A. Klypin, C. G. Scóccola, et al., MNRAS **456**, 4156 (2016), 1509.06400. [VIIC](#)
- [76] S. A. Rodríguez-Torres, C.-H. Chuang, F. Prada, H. Guo, A. Klypin, P. Behroozi, C. H. Hahn, J. Comparat, G. Yepes, A. D. Montero-Dorta, et al., MNRAS **460**, 1173 (2016), 1509.06404. [VIIC](#)
- [77] S. Saito, A. Leauthaud, A. P. Hearin, K. Bundy, A. R. Zentner, P. S. Behroozi, B. A. Reid, M. Sinha, J. Coupon, J. L. Tinker, et al., MNRAS **460**, 1457 (2016), 1509.00482v1. [VIIC](#)
- [78] M. Takada and W. Hu, ArXiv e-prints (2013), 1302.6994. [VIIC](#)
- [79] Y. Li, W. Hu, and M. Takada, ArXiv e-prints (2014), 1401.0385. [VIIC](#)
- [80] J. Hartlap, P. Simon, and P. Schneider, A&A **464**, 399 (2007), arXiv:astro-ph/0608064. [VIIC](#)
- [81] W. J. Percival, A. J. Ross, A. G. Sánchez, L. Samushia, A. Burden, R. Crittenden, A. J. Cuesta, M. V. Magana, M. Manera, F. Beutler, et al., MNRAS **439**, 2531 (2014), 1312.4841. [VIIC](#)
- [82] E. Sellentin and A. F. Heavens, ArXiv e-prints (2015), 1511.05969. [VIIC](#)
- [83] T. Matsubara, A. S. Szalay, and A. C. Pope, ApJ **606**, 1 (2004), arXiv:astro-ph/0401251. [VIIE](#)
- [84] M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro, M. S. Vogeley, D. H. Weinberg, I. Zehavi, A. Berlind, et al., ApJ **606**, 702 (2004), arXiv:astro-ph/0310725.
- [85] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, M. R. Blanton, J. A. Frieman, M. Fukugita, J. E. Gunn, A. J. S. Hamilton, G. R. Knapp, et al., Phys. Rev. D **74**, 123507 (2006), arXiv:astro-ph/0608632. [VIIE](#)