Nonlinear contributions to angular power spectra

Mona Jalilvand,[*] Basundhara Ghosh,[†] Elisabetta Majerotto,[‡] Benjamin Bose,[§] Ruth Durrer,[‖] and Martin Kunz[¶]
Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland

(Received 18 December 2019; accepted 24 January 2020; published 24 February 2020)

Future galaxy clustering surveys will probe small scales where nonlinearities become important. Since the number of modes accessible on intermediate to small scales is very high, having a precise model at these scales is important especially in the context of discriminating alternative cosmological models from the standard one. In the mildly nonlinear regime, such models typically differ from each other, and galaxy clustering data will become very precise on these scales in the near future. As the observable quantity is the angular power spectrum in redshift space, it is important to study the effects of nonlinear density and redshift space distortion (RSD) in the angular power spectrum. We compute nonlinear contributions to the angular power spectrum using a flat-sky approximation, and compare the results of different perturbative and nonperturbative approaches. We find that the Taruya-Nishimichi-Saito (TNS) perturbative approach is significantly closer to the comoving Lagrangian acceleration approximation than Eulerian or partially resummed-Lagrangian one-loop approximations, effective field theory of large scale structure or a halofit-inspired model. However, none of these prescriptions agree with each other in the nonlinear regime. A surprising and new result of the present analysis is that for narrow redshift bins, $\Delta z \lesssim 0.01$, the angular power spectrum acquires nonlinear contributions on all scales, right down to $\ell = 2$, and is hence not a reliable tool at this time. To overcome this problem, we need to model nonlinear RSD terms, for example as TNS does, but for a matter power spectrum that remains reasonably accurate well into the deeply nonlinear regime, such as halofit.

DOI: 10.1103/PhysRevD.101.043530

I. INTRODUCTION

After the tremendous success of cosmic microwave background observations [1], presently major efforts in cosmology are going into the observation and modeling of the distribution of galaxies [2–10]. As this data set is three dimensional, it is potentially much richer and may allow us to study the evolution of cosmic structure formation.

However, on small scales the fluctuations in the matter density can become large at the present time. Therefore, first order cosmological perturbation theory is not sufficient to describe structure formation on these scales and numerical N-body simulations, in principle including also hydrodynamic effects, are needed. This is a very complicated process and usually many phenomenological parameters have to be used to describe the highly nonlinear hydrodynamic processes which are affected by star formation, active galactic nuclei feedback and more [11–13].

On intermediate scales, higher order perturbation theory and phenomenological modeling of the galaxy power spectrum can be used [14–18]. This is the topic of the present work. In the past, people have mainly looked at the power spectrum in Fourier space [4,8]. Within linear perturbation theory this is approximated by the so-called Kaiser formula [14], which includes redshift space distortions (RSD), i.e., the fact that the observed redshift is affected by peculiar velocities which are in turn correlated with matter overdensities,

$$P(k,\mu,\bar{z}) = D_1^2(\bar{z})[b(\bar{z}) + f(\bar{z})\mu^2]^2 P_m(k), \qquad (1)$$

where $\mu = \hat{\boldsymbol{k}} \cdot \boldsymbol{n}$ is the cosine of the angle between the unit vector in direction $\boldsymbol{k}$, $\hat{\boldsymbol{k}}$, and the observation direction $\boldsymbol{n}$, which is a unit vector. Here $\bar{z}$ is a mean redshift of the survey under consideration, $P_m(k)$ is the linear matter density power spectrum today (in real space), $D_1(\bar{z})$ is the linear growth factor normalized to $D_1(0) = 1$, $b(\bar{z})$ is the galaxy bias and

$$f(\bar{z}) = -\frac{D_1'}{D_1}(1+\bar{z}) = \frac{d\ln D_1}{d\ln(a)} \qquad (2)$$

is the growth rate, where the prime denotes the derivative with respect to the redshift $\bar{z}$. This formula has been generalized in the literature to include nonlinearities in the matter power spectrum, usually by replacing

[*]mona.jalilvand@unige.ch
[†]basundhara.ghosh@unige.ch
[‡]elisabetta.majerotto@unige.ch
[§]benjamin.bose@unige.ch
[‖]ruth.durrer@unige.ch
[¶]martin.kunz@unige.ch

$D_1^2(\bar{z})P_m(k)$ by a one-loop or two-loop power spectrum [15] or by a phenomenological approximation like halofit [16,19]. Workers in the field have also corrected the "Kaiser relation" $f(\bar{z})(\hat{\boldsymbol{k}}\cdot\boldsymbol{n})^2$ for the peculiar velocity with a nonlinear and phenomenological description [20]. With the increasing precision of the data available from galaxy surveys such as Euclid [21–23], WFIRST [24,25], 4MOST [26] and DESI [27,28], and with the upcoming HI surveys (e.g., [29–31]) that have a very high redshift resolution, it is important to model the theoretical galaxy power spectrum as accurately as possible. Even at scales as large as those of baryon acoustic oscillations, we need to go beyond linear perturbation theory [8,32].

Equation (1) is a good approximation to cosmological observations only if we have a small, far away galaxy survey in a fixed direction $\boldsymbol{n}$ at nearly fixed redshift $\bar{z}$. A true galaxy survey lives on our background light cone and the radial distance between galaxies is related to their redshift difference. The correlation function therefore is truly a function of two directions, $\boldsymbol{n}_1$, $\boldsymbol{n}_2$ and two redshifts, z_1, z_2. Assuming statistical isotropy it depends only on $\cos\theta = \boldsymbol{n}_1\cdot\boldsymbol{n}_2$, z_1 and z_2. A harmonic transform in $\cos\theta$ yields the spherical power spectrum $C_\ell(z_1,z_2)$. This has been derived at first order in perturbation theory in [33,34]. Apart from density and RSD, the complete formula includes several relativistic effects like the integrated Sachs Wolfe effect, the Shapiro time delay, the gravitational potential at the source and gravitational lensing convergence (also termed "magnification bias"). Apart from the last term, all relativistic contributions are relevant only on very large scales corresponding to $\ell \lesssim 10$. The gravitational lensing contribution is relevant in wide redshift bins, at relatively high redshifts, $z \gtrsim 1$, or in widely separated redshift bins [33,35,36]. For the redshift bin widths used in this work, we discuss briefly in the Appendix E the importance of lensing in angular power spectra, relative to the RSD contribution.

Here, we consider spectroscopic surveys which have a very precise redshift distribution and we neglect lensing. We want to determine the effect of loop corrections in Eulerian and Lagrangian perturbation theory as well as other phenomenological approaches to the nonlinear matter power spectrum. We study how these corrections affect the observable angular power spectrum, C_ℓ, when considering density and redshift space distortions, and we compare them with results from the more accurate comoving Lagrangian acceleration (COLA) [37–40] simulations. The density and RSD contributions are dominant at relatively low redshifts and for spectroscopic surveys like Euclid; furthermore, it is these terms which are most affected by nonlinearities. The main point of this paper is not to make precise forecasts for which certainly the lensing term should not be neglected, but to study the effect of nonlinear corrections in the C_ℓ's coming from clustering and RSD.

In the next section, we derive a flat-sky approximation for density and RSD which is surprisingly accurate even at low ℓ. In Sec. III we compute the C_ℓ's from the different approximations for nonlinear power spectrum in redshift space (described in Appendix A) and compare them with the linear and halofit results. This section contains our main findings. We also compare our theoretical predictions to measurements made from a set of COLA N-body simulations. In Sec. IV we discuss our findings and conclude.

II. THE FLAT-SKY APPROXIMATION

In this section, we discuss a flat-sky approximation [41–43] which we use to compute the angular power spectrum $C_\ell(z_1,z_2)$ for galaxy number counts from the three-dimensional power spectrum, where z_1 and z_2 are two (relatively close) redshifts. We start from the correlation function in configuration space which in principle depends on two spatial positions and two redshifts, $\xi(\boldsymbol{x}_1,z_1;\boldsymbol{x}_2,z_2)$, where $(\boldsymbol{x}_1,z_1)$ and $(\boldsymbol{x}_2,z_2)$ are constrained to lie on our background light cone. We assume that the redshifts are relatively close so that the time evolution between z_1 and z_2 can be neglected. Then the correlation function depends only on $\mathbf{r} = \boldsymbol{x}_2 - \boldsymbol{x}_1$ and $\bar{z} = (z_1+z_2)/2$ (see Fig. 1). This correlation function in real space, $\xi(\mathbf{r},\bar{z})$, is the Fourier transform of the power spectrum

$$\xi(\mathbf{r},\bar{z}) = \frac{1}{(2\pi)^3}\int d^3\mathbf{k}P(\mathbf{k},\bar{z})e^{-i\mathbf{k}\cdot\mathbf{r}}. \tag{3}$$

Let us now consider the flat-sky approximation, which amounts to assuming that the direction from the observer to the points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is nearly equal, $\mathbf{n}_1 \simeq \mathbf{n}_2 = \mathbf{n}$; i.e., the survey covers a relatively small patch of the sky in a fixed direction $\boldsymbol{n}$. This is the situation for which Eq. (1) can be used as an approximation for the power spectrum. In this case we can also decompose the separation vector $\mathbf{r}$ into components perpendicular and parallel to the line-of-sight direction $\mathbf{n}$, as shown in Fig. 1, so we have

$$\begin{aligned}\mathbf{r} &= \mathbf{r}_\perp + r_\parallel\mathbf{n},\\ r_\parallel &= r\nu \simeq \chi(z_2) - \chi(z_1) \simeq \frac{\Delta z}{H(\bar{z})},\end{aligned} \tag{4}$$

where $\chi(z)$ is the comoving distance to redshift z, $\nu = \hat{\mathbf{r}}\cdot\hat{\mathbf{n}}$ as shown in Fig. 1. Similarly in k-space we define

$$\begin{aligned}\mathbf{k} &= \mathbf{k}_\perp + k_\parallel\mathbf{n},\\ k_\parallel &= k\mu = k\hat{\mathbf{k}}\cdot\mathbf{n}.\end{aligned} \tag{5}$$

We introduce the dimensionless two-dimensional (2D) vector $\boldsymbol{\ell}$ by $\mathbf{k}_\perp \equiv \boldsymbol{\ell}/\chi(\bar{z})$. Therefore by using Eqs. (4) and (5) we can rewrite Eq. (3) as

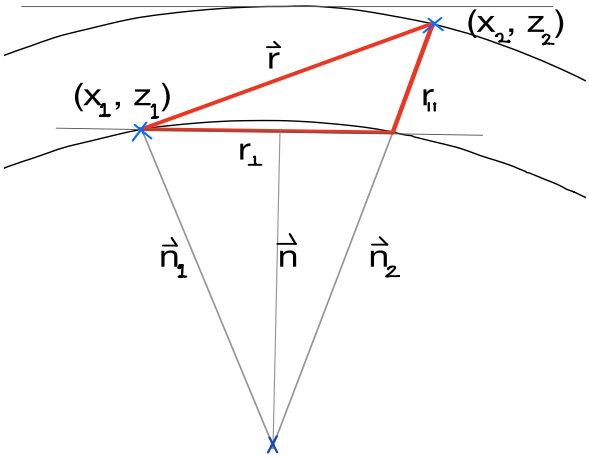

FIG. 1. We show the positions $(\boldsymbol{x}_1, z_1)$ and $(\boldsymbol{x}_2, z_2)$ on the background light cone of an observer situated at X and their flat-sky approximations.

$$\xi(\mathbf{r}, \bar{z}) = \frac{1}{(2\pi)^3} \int \frac{d^2\boldsymbol{\ell}}{\chi^2(\bar{z})} dk_\parallel P(\mathbf{k}, \bar{z}) e^{-i\left(\frac{\boldsymbol{\ell}\cdot\mathbf{r}_\perp}{\chi(\bar{z})} + k_\parallel \frac{(z_2 - z_1)}{H(\bar{z})}\right)}. \quad (6)$$

On the other hand, we know how to compute $\xi(\mathbf{r}, z)$ from the angular power spectrum. In the flat-sky approximation this yields (see e.g., [44])

$$\xi(\mathbf{r}_\perp, z_1, z_2) = \frac{1}{(2\pi)^2} \int d^2\boldsymbol{\ell} C_\ell(z_1, z_2) e^{-i\boldsymbol{\ell}\cdot\mathbf{r}_\perp/\chi(\bar{z})}. \quad (7)$$

By comparing Eqs. (6) and (7), we find the relation between the angular power spectrum and the three-dimensional power spectrum in Fourier space as

$$C_\ell(z_1, z_2) = \frac{1}{2\pi\chi^2(\bar{z})} \int_{-\infty}^{+\infty} dk_\parallel P(k, \bar{z}) e^{-ik_\parallel (z_2 - z_1)/H(\bar{z})}, \quad (8)$$

for $k = \sqrt{k_\parallel^2 + (\ell/\chi)^2}$. Note that this approximation is not equivalent to the Limber approximation [45], which is often used for weak lensing calculations where $k \simeq (\ell + 1/2)/\chi(z)$ is used instead of an integration of the power spectrum times the Bessel function. In this flat-sky approximation we identify the flat-sky vectors

$$\mathbf{k}_\perp \equiv \boldsymbol{\ell}/\chi(\bar{z}), \quad (9)$$

and integrate over $k_\parallel$. Contrary to Limber's approximation, which is bad for the density and RSD contributions to number counts (see e.g., [46]), this approximation turns out to be excellent for close redshifts $z_1 \simeq z_2$, when compared to the exact definition of $C_\ell(z_1, z_2)$ which, at low ℓ, is given by (see Appendix B of [33], where we have added the bias dependence)

$$\begin{aligned} C_\ell(z_1, z_2) = \frac{2}{\pi} b(z_1) b(z_2) \int dk k^2 P_m(k, z_1, z_2) \\ \times \Big[j_\ell(k\chi(z_1)) j_\ell(k\chi(z_2)) \\ - \frac{f(z_2)}{b(z_2)} j_\ell(k\chi(z_1)) j_\ell''(k\chi(z_2)) \\ - \frac{f(z_1)}{b(z_1)} j_\ell''(k\chi(z_1)) j_\ell(k\chi(z_2)) \\ + \frac{f(z_1)}{b(z_1)} \frac{f(z_2)}{b(z_2)} j_\ell''(k\chi(z_1)) j_\ell''(k\chi(z_2)) \Big]. \end{aligned} \quad (10)$$

Here $P_m(k, z_1, z_2)$ is the matter power spectrum and $b(z_1)$, $b(z_2)$ are the linear tracer biases at z_1 and z_2. For large redshift separations the flat-sky approximation gets worse. This is because this approximation corresponds to replacing the spherical Bessel function by their lowest frequency modes, assuming that $|\chi(z_1) - \chi(z_2)| \ll \chi(z_1)$, $\chi(z_2)$ which is no longer valid when the redshift difference becomes large—for large redshift differences we additionally need to model the decoherence in $P(k, z_1, z_2)$ correctly, e.g., with the fitting function of [47]. In Fig. 2 we compare the angular power spectrum for $z_1 = z_2 = 1$ computed in the flat-sky approximation Eq. (8) with the one computed with the exact formula of Eq. (10). The differences are at most 1%.

III. NONLINEAR CORRECTION TO THE ANGULAR POWER SPECTRUM

To profit optimally from future galaxy redshift surveys (Euclid, DESI, 4MOST, SKA, etc.) [23,26,28,49] we must also be able to model scales where nonlinearities become relevant. Since the angular power spectrum is the true observable, it is important to study directly this quantity. In this section, we discuss the effects of nonlinearities on the angular power spectrum considering four different approaches: one-loop corrections from standard Eulerian perturbation theory which we term SPT, one-loop partially resummed Lagrangian perturbation theory (pr-LPT), corrections from effective field theory of large scale structure (EFT) and the Taruya-Nishimichi-Saito (TNS) model [16]. We compare them on the one hand with COLA simulations [38,40] and on the other hand with the halofit power spectrum from CAMB where we introduce RSD simply via the Kaiser term. More details on the simulations and on these approximations are given in Appendix A where we also discuss the nonlinear correction to the power spectrum in redshift space. The aim of the present work is the study of nonlinear corrections of the $C_\ell(z)$ power spectra. We study their effects at different redshifts and for different widths of the redshift bins considered.

Although the power spectrum $P^s(k, \mu, z)$ in Fourier space, and its counterpart, the correlation function $\xi(r, \mu, z)$, provide some insight into galaxy observations on small

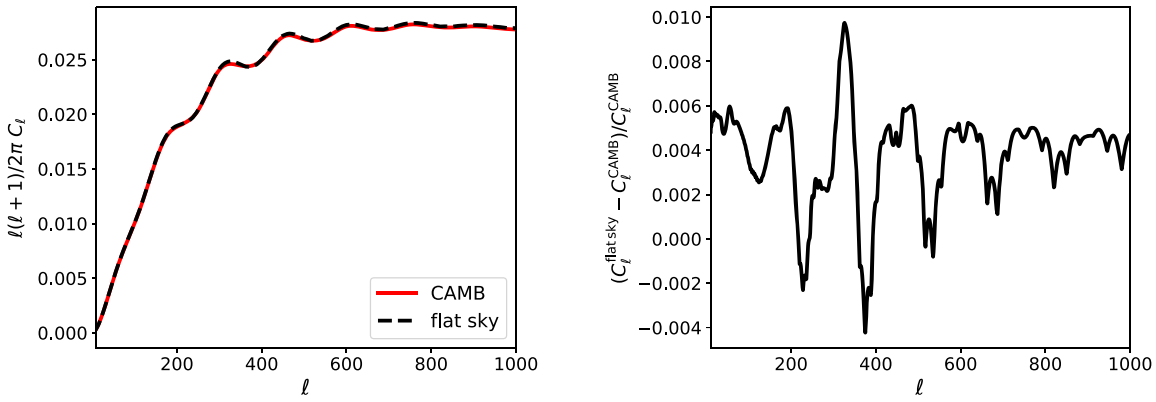

FIG. 2. The left panel shows the comparison between the flat-sky approximation of Eq. (8) and the angular power spectrum computed by code for anisotropies in the microwave background [48] (CAMB) that uses Eq. (10) at $z = 1$, using a top hat window function with $\Delta z = 0.1$. The right panel shows the relative difference between the two.

scales, here we investigate how these nonlinearities project onto the sky, i.e., onto the directly observable angular power spectrum.

In the top panel of Fig. 3 we compare the C_ℓ's from the different nonlinear approximations discussed in Appendix A all at redshift $z = 0.5$ using different bin widths Δz. For $\Delta z = 0.1$ and $\ell \lesssim 150$, which corresponds roughly to the nonlinearity scale at $z = 0.5$, the spectra agree relatively well. Beyond that scale they become very different, and even though in k-space TNS is a better approximation to the numerical results this is no longer true in ℓ space where the CAMB halofit (red line) does best mimic the COLA result (grey line), but also this result is more than 20% off at $\ell = 1000$ from the COLA simulation and a better approximation is certainly needed.

When smaller bin widths are chosen, $\Delta z = 0.01$ for the middle panel and $\Delta z = 0.001$ for the lower panel, the difference between the approximations and the COLA simulations becomes even worse. For these bin widths more small scale power enters the C_ℓ's which not only increases their amplitude but also makes them more sensitive to the treatment of nonlinearities. Let us investigate this important new finding in more detail.

We define the nonlinearity scale through the condition

$$\sigma(R_{\rm NL}, z) = 0.2 \tag{11}$$

that was also used by Euclid [22,50]. Here $\sigma^2(R, z)$ is the usual variance of the mass fluctuation in a sphere of radius R,

$$\sigma^2(R, z) \equiv \frac{1}{2\pi^2}\int_0^\infty \frac{dk}{k}\left(\frac{3j_1(kR)}{kR}\right)^2 k^3|\delta(k, z)|^2, \tag{12}$$

so that $\sigma(R = 8\ {\rm Mpc}/h, z = 0) = \sigma_8$. We then associate a nonlinearity scale in Fourier space through

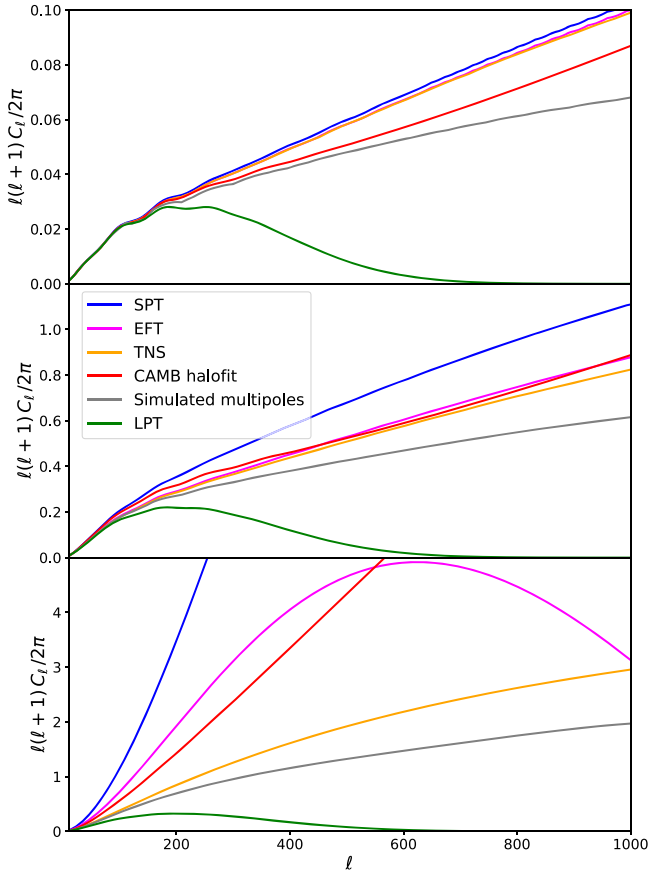

FIG. 3. The plot shows the C_ℓ's in redshift space at $z = 0.5$ computed using the different approaches discussed in the text: $P^{\rm sSPT}_{\rm one-loop}$, Eq. (A2), $P^{\rm sLPT}_{\rm one-loop}$, Eq. (A3), $P^{\rm sEFT}_{\rm one-loop}$, Eq. (A5), and $P^{\rm sTNS}_{\rm one-loop}$, Eq. (A7), as well as the nonlinear C_ℓ computed by CAMB using the halofit model, and the simulated multipoles from COLA. The redshift bin width is $\Delta z = 0.1$ for the top panel, $\Delta z = 0.01$ for the middle panel and $\Delta z = 0.001$ for the bottom panel. None of the models shown here manages to agree with the numerical simulations except on the largest scales and for wide redshift bins.

$$k_{\rm NL}(z) = \frac{2\pi}{R_{\rm NL}(z)}. \tag{13}$$

A given transversal wave number $k_\perp$ at redshift z roughly corresponds to a multipole

$$\ell(k,z) \simeq k_\perp \chi(z). \tag{14}$$

In Fig. 4 we show $\ell(k_\perp, z)$ for three different values of $k_\perp$ as well as $\ell_{\rm NL}(z) = \ell(k_{\rm NL}(z), z)$.

In Fig. 5 we compare linear and nonlinear spectra for different redshifts for the TNS model in Fourier space. For $k < 0.15\ h/\mathrm{Mpc}$ the real space spectra (dashed lines) are closer to the linear result than the redshift space spectra with $\mu = 1$ (solid lines). This indicates that velocities exhibit nonlinearities already on larger scales than the density. Roughly at $k = 0.15\ h/\mathrm{Mpc}$ this trend is reversed. When we enter a more nonlinear regime (after shell crossing), the velocities tend to damp the power in redshift space, so the redshift space spectra are less nonlinear than the real space spectra on these scales. Interestingly, the "crossover" scale of $k = 0.15\ h/\mathrm{Mpc}$ seems to be nearly redshift independent.

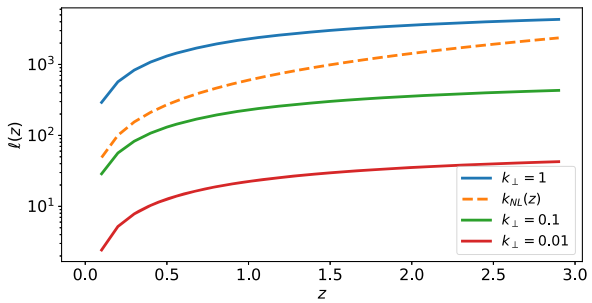

FIG. 4. We show $\ell(k_\perp, z)$ for $k_\perp = (0.01, 0.1, 1)\ h/\mathrm{Mpc}$ as well as $\ell_{\rm NL}(z) = \ell(k_{\rm NL}(z), z)$ as a function of redshift z.

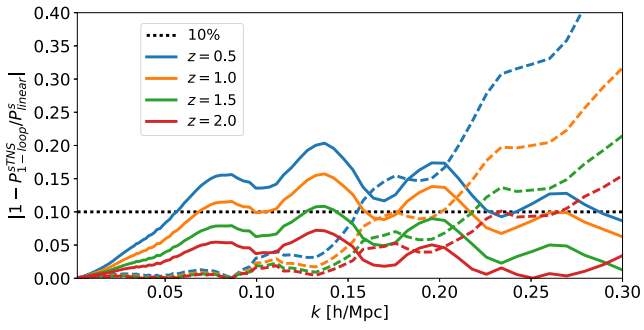

FIG. 5. Relative differences between $P^s(k,\mu)^{\rm NL}$ and $P^s(k,\mu)^{\rm lin}$ for the TNS model for different redshifts. The dashed lines represent the real space case ($\mu = 0$), and the solid ones represent the redshift space case ($\mu = 1$). The horizontal dotted black line is the 10% line. For the real space spectrum nonlinearities become important around $k = 0.1\ h/\mathrm{Mpc}$, while in the $\mu = 1$ spectrum with RSD the nonlinearities appear on much larger scales.

In Fig. 6 we compare linear and nonlinear angular spectra for different redshifts using the TNS model for the nonlinear case. We use here the TNS model and not halofit since, as we explain later in Figs. 10 and 11, for the smaller redshift bins which are sensitive to RSD; TNS follows the simulation results better than CAMB halofit. For $\Delta z = 0.1$, the higher the redshift the higher the value of ℓ below which our model deviates by less than 10% (black dotted line) from the linear result. Furthermore, redshift space distortions are not very visible in ℓ space for $\Delta z = 0.1$ (see the top panel of Fig. 6).

For $\Delta z = 0.01$ and $\Delta z = 0.001$ (middle and low panels of Fig. 6), RSDs are very prominent but now, even for very low ℓ, the linear approximation is no longer sufficient. This is due to the fact that a very precise redshift resolution in the spectrum is sensitive to very small radial modes, hence to very high values of $k_\parallel$ which are affected by nonlinearities. Physically this just means that we are sensitive to nonlinearities if we want high resolution in any direction, radial

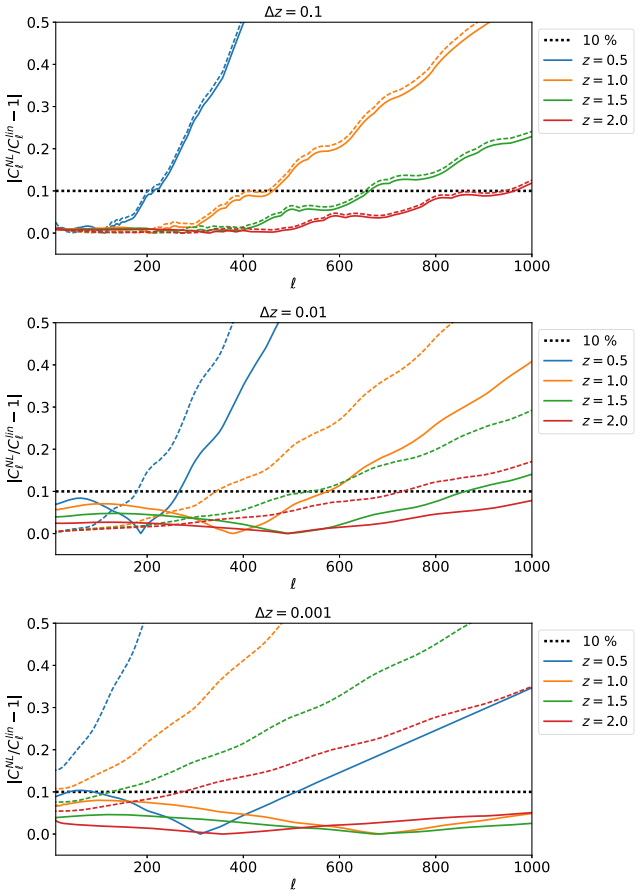

FIG. 6. Relative difference between $C_\ell^{\rm NL}$ and $C_\ell^{\rm lin}$ for the TNS model for different redshifts. The dashed lines represent the real space case (neglecting any RSD terms in the power spectrum), and the solid ones represent the redshift space case. The horizontal dotted black line is the 10% line.

or transversal. This shift of the nonlinearity scale to lower ℓ's for narrow redshift bins is also what we see in the lower panels of Fig. 3.

This is a very important result of the present paper: if we want to resolve RSD in the angular power spectrum we must have sufficiently precise redshift measurements, in which case the C_ℓ's are sensitive to nonlinearities in the radial power spectrum for virtually all ℓ's.

At the highest redshift, $z = 2$ and for the most narrow redshift bin, $\Delta z = 0.001$ this yields simply a nearly constant offset from the linear results by about 5%. For lower redshifts and/or larger bin widths the difference from the linear result grows with ℓ as one naively expects. It is also interesting to note that for the smallest bin width (bottom panel of Fig. 6), the deviation never exceeds 10% for $z \geq 1$ or $z = 0.5$ and $\ell \lesssim 500$. This can be understood by noting that RSDs which are most significant for the smallest bin width damp the nonlinearities.

Mathematically, the fact that nonlinearities at small Δz enter already at low ℓ can be understood very nicely from the flat-sky approximation. Convolving Eq. (8) with a tophat window function of width Δz we find for a mean redshift denoted by $\bar{z}$

$$C_\ell(\bar{z}, \Delta z) = \frac{1}{\pi\chi^2} \int dk_\parallel j_0^2\left(\frac{k_\parallel \Delta z}{2H(\bar{z})}\right) P\left(k_\parallel, \frac{\ell}{\chi}\right). \quad (15)$$

Here the spherical Bessel function $j_0^2(k_\parallel \Delta z/2H)$ acts as a "low pass filter" which filters out modes with $k_\parallel \gg 2H(z)/\Delta z$. For very small Δz the integral therefore extends to high values of $k = \sqrt{k_\parallel^2 + (\ell/\chi)^2}$ for any ℓ, and these modes can become large and nonlinear. In this case nonlinearities affect the result even at the lowest ℓ values. In other words, for linear perturbation theory to apply it is not sufficient that the relevant transverse modes $k_\perp = \ell/\chi(z)$ are well in the linear regime; the relevant radial modes $k_\parallel \leq 2H(z)/\Delta z$ must also be in the linear regime. A crude approximation yields

$$k_{\parallel,\max} \simeq \frac{2\pi H(\bar{z})}{\Delta z} < k_{\rm NL}(\bar{z})$$

or

$$\Delta z \gtrsim (\Delta z)_{\rm min} = \frac{2\pi H(\bar{z})}{k_{\rm NL}(\bar{z})}. \quad (16)$$

We show $(\Delta z)_{\rm min}$ as a function of $\bar{z}$ in Fig. 7. Interestingly, $(\Delta z)_{\rm min}$ actually grows somewhat with growing redshift. This is due to the fact that $H(z)$ grows faster with redshift than $k_{\rm NL}(z)$. The critical width is of the order of $\Delta z \approx 0.01$ to 0.02. For narrower redshift bins (higher redshift resolution) we have to expect that (radial) nonlinearities affect the C_ℓ for all values of ℓ, not only for $\ell > \ell_{\rm NL}$.

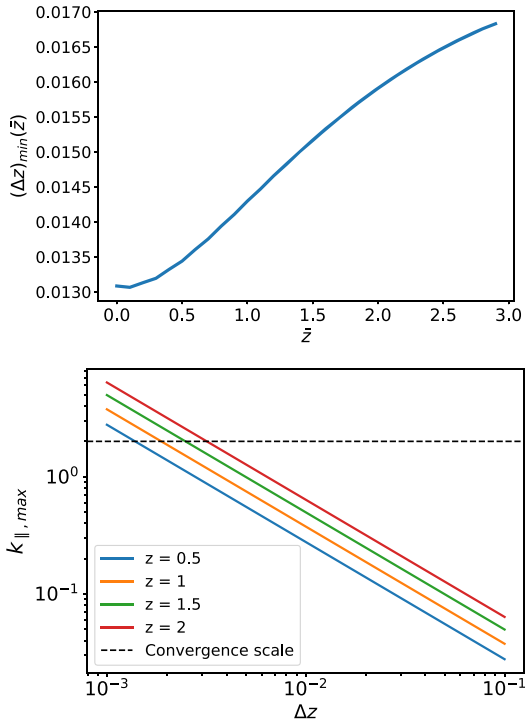

FIG. 7. Top: $(\Delta z)_{\rm min}$ defined in Eq. (16) as a function of $\bar{z}$. For redshift bins narrower than $(\Delta z)_{\rm min}$ we expect radial nonlinearities to affect the C_ℓ's also for low values of ℓ. Bottom: The radial scale $k_{\parallel,\max}$ above which the integrand of the C_ℓ integral is damped by the Bessel function, as a function of bin width Δz, for different redshifts. We also show as a dashed line the convergence scale of the integral in the limit $\Delta z \to 0$.

The radial cutoff scale $k_{\parallel,\max}$, also shown in Fig. 7, lies well below the nonlinear scale for $\Delta z = 0.1$, while for $\Delta z = 0.01$ it is in the range of $k \approx 0.3$ h/Mpc to 0.8 h/Mpc, depending on redshift, already in the nonlinear regime. For narrow redshift bins, $\Delta z = 0.001$, it becomes larger than the "absolute" convergence scale of $k \approx 2$ h/Mpc, for which the C_ℓ integral (15) converges without any damping from the Bessel function, i.e., also for $\Delta z \to 0$ (except for very high ℓ where the effective starting value of the integration, ℓ/χ, is pushed to higher k).

To illustrate clearly the relevance of RSDs we show the difference between the real space spectra (dashed) and redshift space spectra (solid) in the nonlinear predictions in Fig. 8 using the TNS approximation. For the widest redshift bin, $\Delta z = 0.1$, redshift space distortions are not very relevant. For small Δz, however, they significantly reduce the C_ℓ spectrum at high ℓ. As for the power spectrum, on linear scales RSD enhances the power spectrum via the Kaiser effect while on nonlinear scales it reduces it due to the velocity overshoot (shell crossing) which damps the density contribution to power spectrum in redshift space. The crossover between the dashed and the solid line roughly corresponds to the nonlinearity scale at a given redshift. This explains also why the crossover location is nearly independent of the bin width Δz.

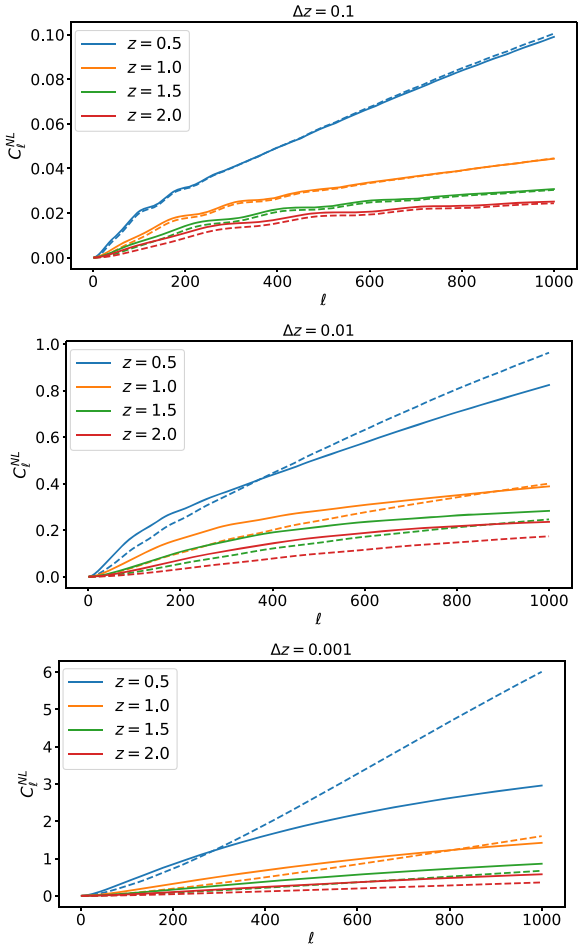

FIG. 8. The comparison between $C_\ell^{\rm NL}$ in real space and redshift space for different mean redshifts and with width $\Delta z = 0.1$, 0.01, 0.001 from top to bottom. *NL* stands for nonlinear, and in this plot we show the case of the TNS model. The dashed lines represent real space while the solid lines represent redshift space.

To quantify the importance of RSD terms, we have performed a simple Fisher forecast for several bin widths. We modeled the RSDs with the Kaiser formula applied to halofit, and only kept the cosmic variance contribution to

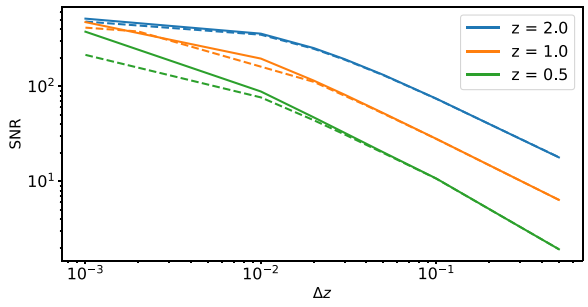

FIG. 9. Signal-to-noise ratio for detecting RSD as a function of Δz for different redshifts. Dashed lines represent SNR using linear C_ℓ's and solid lines represent SNR using nonlinear C_ℓ's. For $z = 1$ and $z = 2$ we sum up to $\ell_{\rm max} = 1000$ and for $z = 0.5$, we sum up to $\ell_{\rm max} = 600$ which is $2\ell_{\rm NL}$.

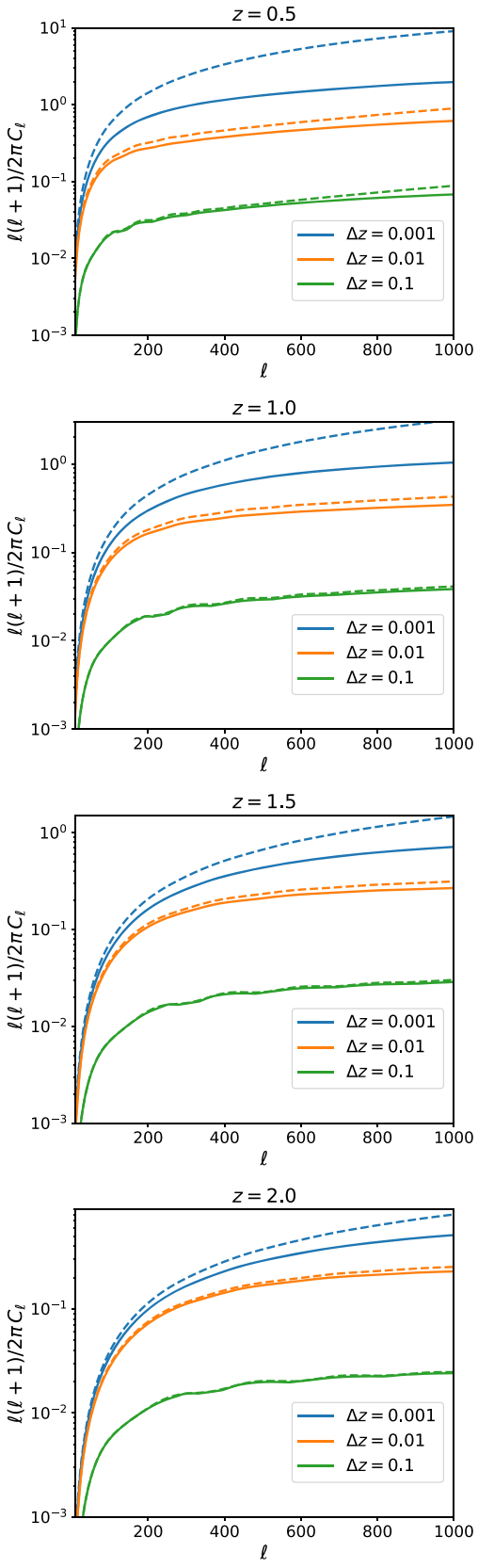

FIG. 10. Comparison between CAMB halofit (dashed) and the COLA simulations (solid) for the redshifts $z = 0.5$, 1.0, 1.5, 2.0 and for different redshift bin widths.

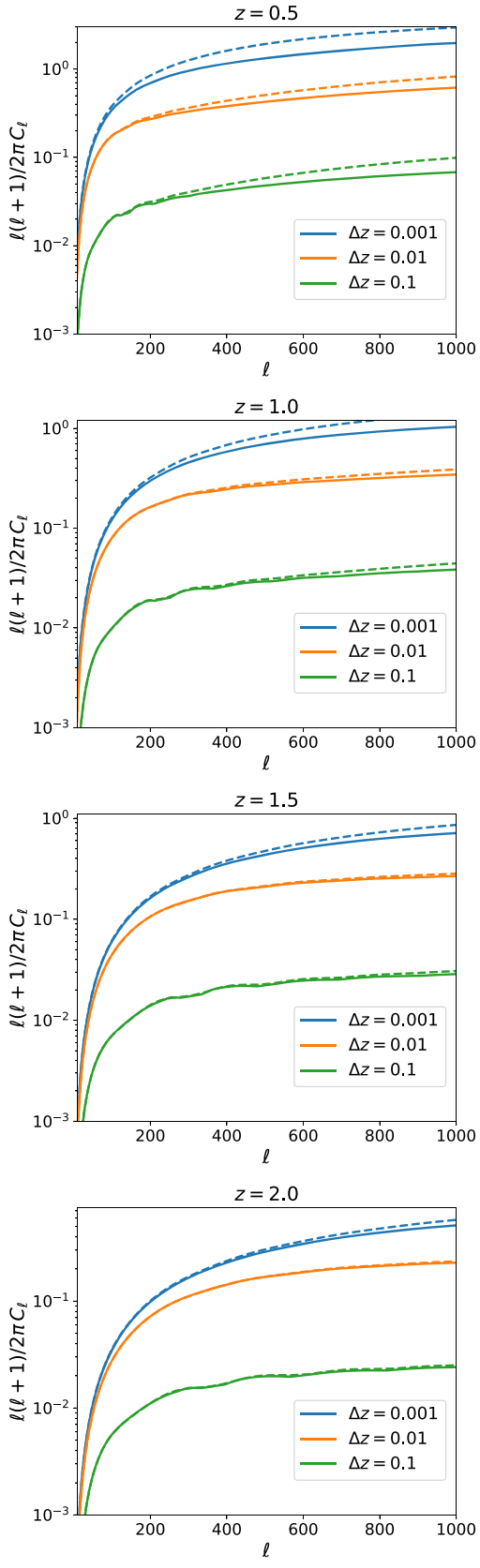

FIG. 11. Comparison between TNS (dashed) and COLA simulations (solid) for the redshifts $z = 0.5$, 1.0, 1.5, 2.0 and for different redshift bin widths.

the noise (neglecting survey-dependent contributions like shot noise and sky fraction). More details about the Fisher analysis are given in Appendix F. The signal-to-noise ratio (SNR) is shown in Fig. 9 as a function of redshift bin width, Δz, for three different redshifts. As we see, the RSD signal drops by an order of magnitude when going from $\Delta z = 0.01$ to $\Delta z = 0.1$, highlighting the importance of using narrow redshift bins for measuring RSD. The RSD signal is however still detectable even for wide bins, and it should therefore be included in the C_ℓ also for $\Delta z = 0.1$.

In Fig. 10 we compare also the result of halofit from CAMB (dashed) with the one from the COLA simulations (solid) for redshift space angular power spectra. For the bin width $\Delta z = 0.1$ we only have a slight enhancement of the CAMB spectrum at $\ell > 400$ for $z = 0.5$; all other spectra are in good agreement. However, for small bin widths $\Delta z \leq 0.01$, and especially for $\Delta z = 0.001$, the insufficient treatment of the RSD in the halofit model where they are taken into account simply by the linear Kaiser formula leads to a significant spurious amplification of the power spectrum already at low values of ℓ. This enhancement is more significant at lower redshifts, where nonlinearities are more relevant, but it is already visible at $z = 2$.

In Fig. 11 we compare the results using the TNS approximation (dashed) with the ones from the COLA simulations (solid) for redshift space angular power spectra. Clearly, the TNS approximation handles redshift space distortions much better that halofit and the spurious excess is reduced and no longer visible for $z = 1.5$ and 2.0. However, we have checked that the relative difference between the simulation result and TNS is larger than cosmic variance for $z = 2$ and $\ell > 200$ as well as $z \leq 1.5$ and $\ell = 100$ for narrow bins, $\Delta z = 0.001$. It becomes larger than cosmic variance at $\ell = 700$, 350 and 200 for $\bar{z} = 1.5$, 1 and 0.5, respectively, for both bin widths, $\Delta z = 0.01$ and $\Delta z = 0.1$. This indicates again that for slim redshift bins, $\Delta z = 0.001$, nonlinearities are relevant already at low ℓ (where however cosmic variance is very large). For wide redshift bins, $\Delta z = 0.1$, where RSD are not significant, the halofit model from CAMB is actually a better approximation than TNS. We have already seen this in Fig. 3.

IV. DISCUSSION AND CONCLUSION

In this paper we have compared different perturbation theory-based schemes, SPT, pr-LPT, the EFT, the TNS model, and halofit to treat nonlinearities in the angular power spectrum. We also compare these predictions with COLA simulations. These simulations are percent level accurate within $k \lesssim 1\ h/\mathrm{Mpc}$ when compared to full N-body measurements of the matter power spectrum [40]. At the level of the redshift space multipoles, the simulations still provide fair accuracy up to $k \sim 0.3\ h/\mathrm{Mpc}$ [32]. All other models considered here are less accurate in the modeling of RSD and so we use the COLA simulations

as our benchmark in accuracy. Since the angular power spectrum is more directly related to observations, we concentrate on this variable. Our main conclusions are summarized here:

(i) We find that the flat-sky approximation is valid at percent level accuracy and so adopt this for all our comparisons of C_ℓ.

(ii) For large bin widths ($\Delta z \sim 0.1$), RSD is much less important and the main contributor to nonlinear information is within the matter power spectrum. At this bin width halofit combined with the Kaiser factor agrees within a few percent with our benchmark model, the COLA simulation up to $\ell \lesssim 400$ at $z = 0.5$ and higher for higher z.

(iii) Small bin widths greatly enhance the impact of nonlinear RSD. Because of this, the TNS model outperforms all other models for $\Delta z = 0.01$ and $\Delta z = 0.001$. Despite this, it is still a poor approximation, being accurate to within a few percent only for $\ell \lesssim 150$ at $z = 0.5$.

(iv) The most surprising result is that for small bin widths, nonlinear RSD information becomes important also for very low ℓ, with significant nonlinear effects ($\sim$10%) being found at $\ell \leq 50$ at $z = 0.5$ for the TNS model.

(v) For large bin widths, the effect of lensing, which is not easy to incorporate in $P(k, \mu)$, but is straightforwardly included in the C_ℓ's, cannot be ignored, and at $z = 0.5$ with $\Delta z = 0.1$, it is already equal in magnitude to the RSD signal at $\ell \leq 150$. This is discussed in Appendix E.

In conclusion, at the level of the angular power spectrum, it becomes very difficult to disentangle nonlinearities and various contributions to the signal. In particular, at low redshift, nonlinear RSD can play a large role at $\ell \leq 150$ for small bin width, while for large bin widths lensing begins to affect the signal. At high redshift ($z > 1$) nonlinear RSD is better controlled, but lensing becomes more important for large bin widths. At $z = 1$ lensing is subdominant to RSD up to $\ell \lesssim 500$ for small and large bin width choices. The TNS model offers a relatively good prescription to model the nonlinear effects of RSD in the angular power spectrum, but is still very limited, especially at low redshift where nonlinearities are enhanced.

While this can be circumvented for wide redshift bins by only considering spectra for linear ℓ's, this becomes impossible for narrow redshift bins. In fact, for $\Delta z \lesssim 0.001$, we need to accurately model the nonlinear spectrum to high k for all values of ℓ. No prescription is currently accurate enough to do this. It appears therefore that, at least for now, the angular spectrum is less well suited to measure RSDs than the correlation function.

We have found that while the TNS approximation is the only one with a reasonably good treatment of velocities, it does not reproduce well the COLA angular power spectra for wide redshift bins, $\Delta z \geq 0.1$. For such wide-bin spectra, RSDs are not important and halofit, which gives the better fit to the real space power spectrum than TNS, is actually preferable. On the other hand, for slim redshift bins, $\Delta z \leq 0.01$ TNS is a much better approximation. For such bin widths, radial nonlinearities are already relevant for very low ℓ's which renders halofit, or even more so the linear power spectrum, simply useless. On the other hand, on scales $\ell > \ell_{\rm NL}(z)$, where also the transverse wave number enters the nonlinear regime, also the TNS approximation which models the real space power spectrum purely becomes insufficient, especially at low redshift, $z \simeq 0.5$.

From this work it is clear that we are still far away from modeling the angular power spectrum at 1% precision over a reasonable range of ℓ. But we now know better in which direction we have to make progress. We need to model the real space power spectrum similar to halofit but then correct for nonlinear RSD like in the TNS model to obtain a good approximation for the redshift space power spectrum, especially if we want to model the C_ℓ's in narrow redshift bins where they are sensitive to redshift space distortions. This is essential if we wish to safely extract very important cosmological information. We must make sure to model RSDs very precisely, as they can enter the C_ℓ's at small $\ell \ll \ell_{\rm NL}(z)$ depending on the bin width.

ACKNOWLEDGMENTS

It is a pleasure to thank Joyce Byun, Marko Simonovic, Zvonimir Vlah, Alberto Salvio, Farbod Hassani and Antony Lewis for helpful discussions. We acknowledge financial support from the Swiss National Science Foundation. B. B. acknowledges support from the Swiss National Science Foundation (SNSF) Professorship Grant No. 170547.

APPENDIX A: NONLINEAR CORRECTIONS TO THE POWER SPECTRUM IN REDSHIFT SPACE

In this Appendix we give a summary of different nonlinear corrections to the power spectrum that can be found in the literature. More precisely we consider four different approaches: one-loop corrections from standard Newtonian (Eulerian) perturbation theory which we term SPT, pr-LPT, corrections from EFT and the TNS model. Some important references for each of these approaches are [16,51–53], respectively. We perform all perturbative calculations at the one-loop level (see Appendix B for details). We also make use of a set of measurements of the redshift power spectrum from COLA simulations. These represent our most accurate prediction, with which we can compare the perturbative approaches. These simulations are described briefly below. At the end of this section we compare the different approximations to these simulations.

1. COLA

We have run a set of 10 parallel comoving Lagrangian acceleration (PICOLA) simulations [38,40] of box size 1024 Mpc/h with 1024^3 dark matter particles and a starting redshift $z_{\rm ini} = 49$. These are all run under a similar ΛCDM cosmology with Planck parameters [1]: $\Omega_m = 0.315$, $\Omega_b = 0.0493$, $h = 0.674$, $n_s = 0.965$ and $\sigma_8(z = 0) = 0.811$. The simulation redshift space power spectrum multipoles are measured using the distant-observer (or flat-sky) approximation [54] and are then averaged over three line-of-sight directions. We further average over the 10 PICOLA simulations. We measure the first three even multipoles, the monopole, quadrupole and hexadecapole. Using these we can then construct the full anisotropic power spectrum, $P(k,\mu)$ [55],

$$P_{\rm tot}^{\rm sCOLA}(\mathbf{k},z) = 2P_0(k,z) + \frac{2}{5}\mathcal{L}_2(\mu)P_2(k,z) + \frac{2}{9}\mathcal{L}_4(\mu)P_4(k,z) + [\text{higher order multipoles}], \tag{A1}$$

where $\mathcal{L}_i$ is the Legendre polynomial of order "i" and P_i is the ith multipole which is an average over the measurements made from the COLA simulations. Finally, we note that the COLA method is an approximate method and has been shown to deviate from the full N-body approach at smaller scales [56–58]. This issue can be ignored as we simply use these simulations as a benchmark in accuracy with which to compare the less accurate perturbative predictions outlined next. For example, the redshift space monopole for lowly biased halos was shown to be accurate to full N body to within a few percent at $z \leq 1$ up to $k = 0.7$ h/Mpc in [56]. On the other hand, the quadrupole deviates by up to 10% at $z = 1$ at $k = 0.7$ h/Mpc in the same paper. Regarding this issue, we expect the dark matter monopole and quadrupole to perform better than the halo multipoles, and for their accuracy to improve at higher redshifts. Furthermore, we expect the theoretical models discussed in this section to perform significantly worse at these scales.

2. SPT

One-loop contributions to the power spectrum in redshift space (denoted by superscript s) in the context of SPT are already well established and have previously been calculated in the literature (for a review see e.g., [59]). Here as a reference we point to Eq. (15) of [51],

$$\begin{aligned} P_{\rm tot}^{\rm sSPT}(\mathbf{k},z) &\equiv P_{\rm lin}^s + P_{\rm one-loop}^{\rm sSPT} = P_{\rm lin}^s + P_{22}^s + P_{13}^s \\ &= (1+\beta\mu^2)^2 b^2 P_{\rm lin}(k,z) + 2\int \frac{d^3\mathbf{q}}{(2\pi)^3} P_{\rm lin}(q,z) P_{\rm lin}(|\mathbf{k}-\mathbf{q}|,z)[F_2^S(\mathbf{q},\mathbf{k}-\mathbf{q})]^2 \\ &\quad + 6(1+\beta\mu^2) b P_{\rm lin}(k,z) \int \frac{d^3\mathbf{q}}{(2\pi)^3} P_{\rm lin}(q,z) F_3^S(\mathbf{q},-\mathbf{q},\mathbf{k}), \end{aligned} \tag{A2}$$

where $P_{\rm lin}(k,z)$ is the linear power spectrum in real space, $\beta \equiv f/b$, f being the linear growth rate and b being the linear bias, $\mu = \hat{\mathbf{k}}\cdot\mathbf{n}$, and $F_2^S(\mathbf{q},\mathbf{k}-\mathbf{q})$ and $F_3^S(\mathbf{q},-\mathbf{q},\mathbf{k})$ are the kernels of higher order perturbations. Their expressions are computed from Eq. (13) of Ref. [51] by neglecting higher order biases. The details of the integrations that appear in Eq. (A2) are given in Appendix B for completeness. Furthermore, since we only consider dark matter we set $b = 1$.

3. pr-LPT

Here we consider a specific case of pr-LPT introduced in [52]. The power spectrum using pr-LPT is given in Eq. (63) of Ref. [52],

$$P_{\rm tot}^{\rm sLPT}(\mathbf{k},z) = \exp\{-k^2[1+f(f+2)\mu^2]A\} \times \{P_{\rm tot}^{\rm sSPT}(\mathbf{k},z) + (1+f\mu^2)^2[1+f(f+2)\mu^2]k^2 P_{\rm lin}(k,z)A\}, \tag{A3}$$

where

$$A = \frac{1}{6\pi^2}\int dq P_{\rm lin}(q,z). \tag{A4}$$

The prefactor encodes a damping on small scales from velocity dispersion.

4. EFT

We also consider effective field theory of large scale structure [60–65] where counterterms are added to the SPT power spectrum, for which we refer to Eq. (3.8) of Ref. [53],

$$P_{\rm tot}^{\rm sEFT}(\mathbf{k},z) = P_{\rm tot}^{\rm sSPT}(\mathbf{k},z) - 2\sum_{n=0}^{3} c_{2|\delta_s,2n}\mu^{2n}\frac{k^2}{k_{\rm nl}^2}P_{\rm lin}(k,z), \tag{A5}$$

where

$$c_{2|\delta_s,6} = f^3 c_{2|\delta_s,0} - f^2 c_{2|\delta_s,2} + f c_{2|\delta_s,4}. \quad \text{(A6)}$$

We do not apply a resummation scheme as is commonly done in the literature. The effect of resummation has been shown to have a low impact on the fitting to COLA data conducted in [53]. The values of the counterterm coefficients are determined by fitting to the COLA simulations. This follows a similar procedure to [32]. We refer the reader to this work for justifications and details of this procedure. This is briefly described in Appendix D where also the numerical values of the fitting parameters are given.

5. TNS

The last model we consider is the one-loop TNS model. This model was introduced in [16] and is one of the best approaches to perturbation theory known at present, having been applied in the recent BOSS galaxy clustering analysis [4,8]. It has also been thoroughly validated against simulations and has stood up to other perturbative models [9,18,20,32,57,66–71]. The model is given by [16]

$$P^{\rm sTNS}_{\rm tot}(k,z) = \frac{1}{1+(k^2\mu^2\sigma_v^2)/2}[P^{\delta\delta}_{\rm one-loop}(k,z) + 2\mu^2 P^{\delta\theta}_{\rm one-loop}(k,z) + \mu^4 P^{\theta\theta}_{\rm one-loop}(k,z) + A(k,\mu,z) + B(k,\mu,z) + C(k,\mu,z)]. \quad \text{(A7)}$$

The terms in brackets are all constructed within SPT, with $\delta\delta$, $\delta\theta$ and $\theta\theta$ denoting density-density, density-velocity and velocity-velocity one-loop power spectra. The perturbative correction terms A, B and C are nonlinear corrections coming from the RSD modeling while the prefactor is added for phenomenological modeling of the fingers-of-God effect. Within this prefactor, σ_v is a free parameter that is fit to the COLA simulations (see Appendix D). We refer the reader to [16,32] for a detailed description of the components A, B and C of the model, but we give some basic expressions in Appendix C. We stick to one-loop TNS in this study; the reason is that although the two-loop version is shown to do marginally better than the one loop at higher redshifts (look at [72] for example), the addition of loops in standard PT is not guaranteed to improve predictions and our aim is to quantify how nonlinearities (here expressed through the one-loop and fingers-of-God damping) translate to the angular spectra, and in particular, how redshift bin width acts as a filter for these redshift anisotropy nonlinearities, highlighted in Fig. 3.

6. Comparisons

We compare Eq. (A2) (SPT, blue), Eq. (A3) (pr-LPT, green), Eq. (A5) (EFT, magenta) and Eq. (A7) (TNS, orange) with Eq. (A1) (COLA reconstructed 2D spectrum, grey dots) in Fig. 12. We also compare Eq. (1) with $P_m(k)$ given by linear theory (linear Kaiser, dashed black), nonlinear halofit power spectrum [19] (halofit, red) and the matter power spectrum as measured from the simulations (black dots). These comparisons are done at $z = 0.5$ and are shown in the left and right column of the Fig. 12 for $\mu = 0$ (transversal direction) and $\mu = 1$ (radial direction), respectively. We expect that the grey dots marking the reconstructed COLA 2D spectrum of Eq. (A1) provide the most accurate modeling for the full spectrum. This is our benchmark for accuracy.

We also show the one-loop contributions to $P_{\rm lin}$, P_{13} and P_{22}, of Eq. (A2) in the upper panel of Fig. 12. They start to become important at $k \sim 0.1\ h/\mathrm{Mpc}$ for $\mu = 0$, which is well known from the literature, and on smaller scales for $\mu = 1$ that we also discuss later. Furthermore, P_{13} and P_{22} have opposite signs and their amplitudes are individually much larger than their sum, which is an indication for the well-known bad convergence properties of SPT [73].

The one-loop SPT power spectrum (blue line) at $z = 0.5$ is shown in the middle panel of Fig. 12 for $\mu = 0$ on the left (in this plot the blue line is covered by the orange line) and for $\mu = 1$ on the right. One sees clearly that SPT has too much power at small scales and fits the COLA simulations (grey dots) in a satisfactory way only for $k \lesssim 0.1\ h/\mathrm{Mpc}$.

The black dots are the COLA matter power spectrum multiplied by the Kaiser factor $(1+\beta\mu^2)^2$. They are accurate until about $k = 0.13\ h/\mathrm{Mpc}$. The keen reader may ask why the grey dots and black dots do not overlap in the left, middle and lower panels of Fig. 12 at small scales. This could be due to inaccuracies in the COLA velocities used in computing the multipoles as well as the exclusion of higher order multipoles in Eq. (A1).

Next we consider pr-LPT (green curves). It is clear from Fig. 12 that the damping introduced in the pr-LPT model is much too strong. Nevertheless, this correction can fit the power spectrum roughly until $k < 0.13\ h/\mathrm{Mpc}$ which is already better than the SPT fit.

The EFT power spectrum is plotted as the magenta line (in Fig. 12 this line is covered by the orange line). Somewhat surprisingly, this fit is only a little but not significantly better than pr-LPT for the angular scales considered. It represents a reasonable approximation until $k \simeq 0.15\ h/\mathrm{Mpc}$. One key reason for the poor fit at $\mu = 1$ is the lack of damping within the SPT spectrum which the EFT counterterms cannot suppress efficiently. The

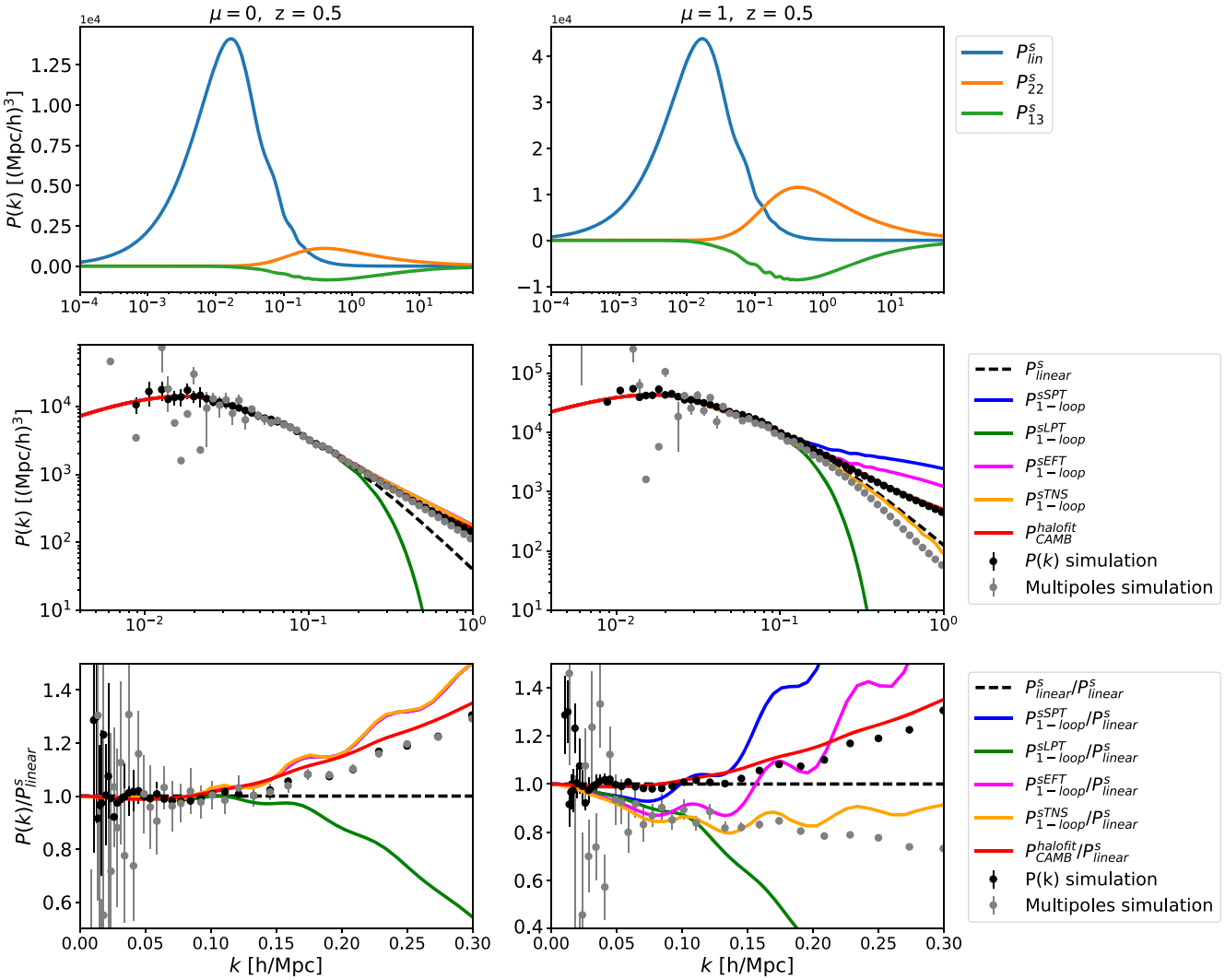

FIG. 12. The figures on the left are for the case with $\mu = 0$ in $P^s(k,\mu)$ (effectively the real space power spectrum) and the figures on the right are for $\mu = 1$. The upper panel shows the linear spectrum P^s_{11} (blue) along with the one-loop contributions P^s_{22} (orange) and P^s_{13} (green). The middle panel shows the comparison between $P^{\rm sSPT}_{\rm one-loop}$, $P^{\rm sLPT}_{\rm one-loop}$, and $P^{\rm sEFT}_{\rm one-loop}$ and $P^{\rm sTNS}_{\rm one-loop}$ defined, respectively, in Eqs. (A2), (A3), (A5) and (A7). In the lower panel the ratios of the corresponding nonlinear spectra and the linear one are shown. The black dots show the monopole of the N-body simulations times the Kaiser term while the grey dots also include the quadrupole and the hexadecapole available from COLA.

inclusion of resummation is expected to improve the fit (see for example [32]) but we leave this to future work.

Lastly, the TNS model is shown in orange. Clearly, this model represents the best fit to the full reconstructed simulated power spectrum for $\mu = 1$ (compare the orange line and the grey dots in Fig. 12, the right lower panel). It can be used roughly until $k \simeq 0.2\ h/\mathrm{Mpc}$. This is somewhat disappointing, as we aspire to achieve a good fit until $k \simeq 1\ h/\mathrm{Mpc}$—to reach convergence in the C_ℓ integral. For narrow redshift bins we find that we need to go even to $2\ h/\mathrm{Mpc}$. On even smaller scales, corrections from baryonic physics that are not present in the simulations used here can at any rate no longer be ignored.

We also note that for $\mu = 0$ the SPT, EFT and TNS power spectra are identical; i.e., in the left column of Fig. 12, the blue, magenta and orange lines overlay. These spectra only differ in their treatment of redshift space distortions which are absent in the transversal direction, $\mu = 0$.

In Fig. 12, we also show the comparison of the COLA measurements with the halofit model given in Eq. (1) (red curve). This approximation is excellent for $\mu = 0$ when fitted to the COLA matter power spectrum, but for $\mu = 1$ it does not correctly model the redshift space distortions. Hence, the higher order RSD and the nonlinearity in the continuity equation which is not taken into account in this formula is very relevant. This is also clear from comparing the black dots, obtained from the matter power spectrum of the COLA simulations by multiplication with the Kaiser term, and the grey dots which represent the full sum of the simulated multipoles. It is also interesting to note that while

the matter power spectrum of the simulations on small scales is larger than the linear power spectrum, adding all the multipoles actually reduces the power spectrum in redshift space on small scales when compared to the linear power spectrum. While the pr-LPT approximation exaggerates this reduction of power, all other approximations either cannot model it at all or (in the case of TNS) underestimate this effect. This is most visible in radial direction, $\mu = 1$. In the transversal direction, $\mu = 0$, the nonlinear corrections from SPT, EFT and TNS all overshoot significantly while pr-LPT is still too small. Here halofit provides the best approximation; see Fig. 12. In the radial direction, $\mu = 1$, only TNS manages to provide a reasonable fit for $k \gtrsim 0.1\ h/\mathrm{Mpc}$, but for $k \gtrsim 0.2\ h/\mathrm{Mpc}$ it also starts to overestimate the power significantly so that there is effectively no good analytical prescription available to model the redshift space power spectrum into the nonlinear regime.

APPENDIX B: DERIVATION OF THE ONE-LOOP TERMS

Following the notations in [51] for Eq. (A2), we have

$$\begin{aligned} P^s_{\mathrm{tot}}(\mathbf{k}) &\equiv P^s_{11} + P^s_{22} + P^s_{13} \\ &= (1+\beta\mu^2)^2 b_1^2 P_{11}(k) + 2\int \frac{d^3\mathbf{q}}{(2\pi)^3} P_{11}(q) P_{11}(|\mathbf{k}-\mathbf{q}|)[F_2^S(\mathbf{q},\mathbf{k}-\mathbf{q})]^2 \\ &\quad + 6(1+\beta\mu^2) b_1 P_{11}(k) \int \frac{d^3\mathbf{q}}{(2\pi)^3} P_{11}(q) F_3^S(\mathbf{q},-\mathbf{q},\mathbf{k}), \end{aligned} \tag{B1}$$

where $\beta \equiv f/b_1$ and b_1 denotes the linear bias. The symmetrized expression for $F_2^S(\mathbf{k}_1,\mathbf{k}_2)$ and the unsymmetrized one for $F_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ are shown in Eq. (13) of [51]. We symmetrize $F_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ and find $F_3^S(\mathbf{q},-\mathbf{q},\mathbf{k})$, neglecting higher order biases.

As can be seen in Eq. (13) of [51], the expressions for F_2^S and F_3^S are given in terms of J_2^S, J_3^S, K_2^S and K_3^S, which can be computed from the general nth order expression as found in the literature [see for example Eqs. (10a) and (10b) of [74]]. While Eqs. (B2) and (B3) given below are easily available in literature, for obtaining Eq. (B4) and Eq. (B5), we have used the expression for $n = 3$ and symmetrized it. However, our results did not match very accurately with the symmetrized expression obtained from Eq. (11) of [51], and therefore we explicitly write them below in Eqs. (B4) and (B5). We find these relations to be as follows:

$$J_2^S(\mathbf{q}_1,\mathbf{q}_2) = \frac{5}{7} + \frac{1}{2}\frac{\mathbf{q}_1.\mathbf{q}_2}{q_1 q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right) + \frac{2}{7}\frac{(\mathbf{q}_1.\mathbf{q}_2)^2}{q_1^2 q_2^2} \tag{B2}$$

$$K_2^S(\mathbf{q}_1,\mathbf{q}_2) = \frac{3}{7} + \frac{1}{2}\frac{\mathbf{q}_1.\mathbf{q}_2}{q_1 q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right) + \frac{4}{7}\frac{(\mathbf{q}_1.\mathbf{q}_2)^2}{q_1^2 q_2^2} \tag{B3}$$

$$\begin{aligned} J_3^S(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \frac{1}{3}\mathrm{Sym}\Bigg[& 7\frac{\mathbf{q}.\mathbf{q}_1}{q_1^2} J_2^S(\mathbf{q}_2,\mathbf{q}_3) + \frac{q^2 \mathbf{q}_1.(\mathbf{q}_2+\mathbf{q}_3)}{q_1^2|\mathbf{q}_2+\mathbf{q}_3|2} K_2^S(\mathbf{q}_2,\mathbf{q}_3) \\ & + \left(7\frac{\mathbf{q}.(\mathbf{q}_1+\mathbf{q}_2)}{|\mathbf{q}_1+\mathbf{q}_2|^2} + \frac{q^2(\mathbf{q}_1+\mathbf{q}_2).\mathbf{q}_3}{|\mathbf{q}_1+\mathbf{q}_2|^2 q_3^2}\right) K_2^S(\mathbf{q}_1,\mathbf{q}_2))\Bigg] \end{aligned} \tag{B4}$$

$$\begin{aligned} K_3^S(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \frac{1}{3}\mathrm{Sym}\Bigg[& \frac{\mathbf{q}_1.\mathbf{q}}{q_1^2} J_2^S(\mathbf{q}_2,\mathbf{q}_3) + \frac{q^2 \mathbf{q}_1.(\mathbf{q}_2+\mathbf{q}_3)}{q_1^2|\mathbf{q}_2+\mathbf{q}_3|2} K_2^S(\mathbf{q}_2,\mathbf{q}_3) \\ & + \left(\frac{\mathbf{q}.(\mathbf{q}_1+\mathbf{q}_2)}{|\mathbf{q}_1+\mathbf{q}_2|^2} + \frac{q^2(\mathbf{q}_1+\mathbf{q}_2).\mathbf{q}_3}{|\mathbf{q}_1+\mathbf{q}_2|^2 q_3^2}\right) K_2^S(\mathbf{q}_1,\mathbf{q}_2)\Bigg]. \end{aligned} \tag{B5}$$

Here Sym indicates symmetrization in $\mathbf{q}_1$, $\mathbf{q}_2$ and $\mathbf{q}_3$. One can replace $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$ and $\mathbf{k} = \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$ as required and effectively calculate these kernels.

The final expressions for $F_2^S(\mathbf{q}, k-q)$ and $F_3^S(\mathbf{q},-\mathbf{q},\mathbf{k})$ along with subsequent calculations can be found in a *Mathematica* notebook [75].

In the kernels F_2^S and F_3^S we encounter the scalar products $\hat{\boldsymbol{q}}\cdot\boldsymbol{n}\equiv\mu_q=\cos(\gamma)$, $\hat{\boldsymbol{k}}\cdot\boldsymbol{n}\equiv\mu=\cos(\alpha)$, and $\hat{\boldsymbol{k}}\cdot\boldsymbol{q}\equiv x=\cos(\beta)$. We also define $r=|\boldsymbol{q}|/|\boldsymbol{k}|$. We can write μ_q in terms of μ, x, and ϕ_q, where ϕ_q is the angle between the projection of $\boldsymbol{q}$ and $\boldsymbol{n}$ onto the plane perpendicular to $\hat{\mathbf{k}}$ (see Fig. 13),

$$\mu_q=x\mu+\sqrt{(1-x^2)(1-\mu^2)}\cos(\phi_q). \tag{B6}$$

For an arbitrary function $\psi(\mathbf{k},\mathbf{q})$, we can write

$$\int\psi(\mathbf{k},\mathbf{q})d^3q=\int_0^\infty q^2dq\int_{-1}^1dx\int_0^{2\pi}d\phi_q\psi(\boldsymbol{k},\boldsymbol{q}). \tag{B7}$$

Therefore the integration corresponding to $P_{22}^s(k,\mu)$ in Eq. (B1) reduces to

$$P_{22}^s=\frac{2k^3}{(2\pi)^3}\int_0^\infty drr^2P_{11}(r)\int_{-1}^1dxP_{11}\left(k\sqrt{1+r^2-2rx}\right)\times\int_0^{2\pi}d\phi_q[F_2^S(r,\phi_q,x,\mu,b_1,f)]^2, \tag{B8}$$

where we take the integral over ϕ_q analytically, and the result can be found in our *Mathematica* notebook [75]. We write P_{22}^s as a sum over powers of μ, b_1 and f as

$$P_{22}^s=\sum_{\ell=0}^{\ell=4}\sum_{m=0}^{m=2}\sum_{n=0}^{n=4}\mu^{2\ell}b_1^mf^nA_{\ell mn}(k), \tag{B9}$$

and finally by integrating over r and x we find the coefficients $A_{\ell mn}(k)$ numerically.

Next, we explain the computation of $P_{13}^s(k,\mu)$, where the integration is as follows:

$$P_{13}^s=6(1+\beta\mu^2)b_1P_{11}(k)\frac{1}{(2\pi)^3}\int_0^\infty drr^2P_{11}(r)\int_{-1}^1d\mu_q\times\int_0^{2\pi}d\phi_qF_3^S(r,\phi_q,x,\mu,b_1,f). \tag{B10}$$

We integrate over ϕ_q and μ_q analytically, the result of which is contained in our *Mathematica* notebook [75]. Then similar to P_{22}^s, we write P_{13}^s as a sum over powers of μ, b_1 and f as

$$P_{13}^s=\sum_{\ell=0}^{\ell=2}\sum_{m=0}^{m=1}\sum_{n=0}^{n=3}\mu^{2\ell}b_1^mf^nB_{\ell mn}(k), \tag{B11}$$

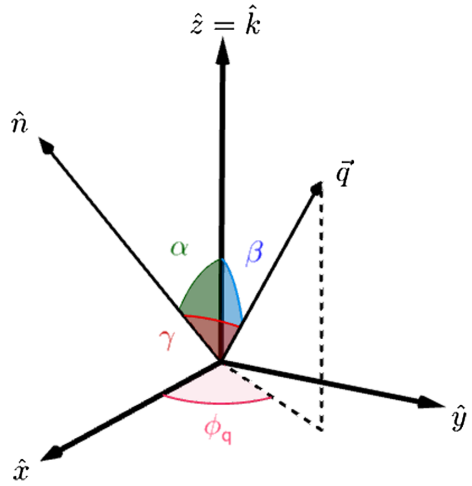

FIG. 13. Vectors and angles involved in the calculations of P_{13} and P_{22}: $\hat{\mathbf{k}}$ is the direction of the wave vector $\mathbf{k}$ of Eq. (A2) and $\mathbf{q}$ is also given in Eq. (A2). $\mathbf{n}$ is the line-of-sight direction.

and by integrating over r numerically, we find the coefficients $B_{\ell mn}(k)$ which are given in our *Mathematica* notebook [75]. We use the minimum and maximum values of wave number that we have from our CLASS output, $k_{\rm min}$, $k_{\rm max}$, for the limits of q.

In order to avoid numerical problems, we employ the following algorithm for computing the integrals in case of P_{13}. We divide the integration range in two large parts.

(i) For $k_{\rm min}<q<k$: If $k<1000k_{\rm min}$, we integrate the original term(s) from $q_{\rm min}=k_{\rm min}$ to k. Or else, we take the sum of the integration of the series expansion around $q=0$ from $q_{\rm min}$ to $q<k/1000$ and integration of the original term(s) from $k/1000$ to $q<k$.

(ii) For $k<q<k_{\rm max}$: If $10k>k_{\rm max}$, we integrate the original term(s) (with the signs corrected for argument of Logarithm) from k to $k_{\rm max}$. Or else, we take the sum of the integration of the original term(s) from k to $10k$ and the integration of the series expansion around $q=\infty$ from $10k$ to $k_{\rm max}$.

APPENDIX C: TNS MODEL, THE *A*, *B* AND *C* CORRECTION TERMS

In this Appendix we present the basic forms of the RSD correction terms appearing in Eq. (A7). These terms are given in terms of the linear power spectrum at k and q multiplied by coefficients A_{mn}, $\tilde{A}_{mn}$, a_{mn} and B_{ab}^n integrated over $r=k/q$ and $x=\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{q}}$ as follows,

$$A(k,\mu)=\sum_{m,n=1}^{3}\mu^{2m}f^n\frac{k^3}{(2\pi)^2}\left[\int dr\int dx(A_{mn}(r,x)P_{\rm lin}(k)+\tilde{A}_{mn}(r,x)P_{\rm lin}(kr,z))\frac{P_{\rm lin}(k\sqrt{1+r^2-2rx},z)}{(1+r^2-2rx)}+P_{\rm lin}(k,z)\int dra_{mn}(r)P_{\rm lin}(kr,z)\right], \tag{C1}$$

$$B(k,\mu) = \sum_{n=1}^{4}\sum_{a,b=1}^{2} \mu^{2n}(-f)^{a+b}\frac{k^3}{(2\pi)^2}\int dr \int dx B^n_{ab}(r,x)\frac{P_{a2}(k\sqrt{1+r^2-2rx},z)P_{b2}(kr,z)}{(1+r^2-2rx)^a}, \tag{C2}$$

$$C(k,\mu) = (k\mu f)^2 \int \frac{d^3p d^3q}{(2\pi)^3}\delta_D(\boldsymbol{k}-\boldsymbol{q}-\boldsymbol{p})\frac{\mu_p^2}{p^2}(1+fx^2)^2 P_{\rm lin}(p,z)P_{\rm lin}(q,z), \tag{C3}$$

where $\mu_p = \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}$. Explicit expressions for A_{mn}, $\tilde{A}_{mn}$, a_{mn} and B^n_{ab} can be found in the Appendices of [16]. The $C(k,\mu)$ term is known to have small oscillatory features and thus it is usually omitted in the literature. We choose to include it in our work.

APPENDIX D: FITTING PROCEDURE FOR EFT AND TNS MODEL

To fit the RSD free parameters of the EFT [Eq. (A5)] and TNS [Eq. (A7)] models to the simulation data we simply minimize the $\chi^2_{\rm red}$,

$$\chi^2_{\rm red}(k_{\rm max}) = \frac{1}{N_{\rm dof}}\sum_{k=k_{\rm min}}^{k_{\rm max}}\sum_{\ell,\ell'=0,2}[P^S_{\ell,{\rm data}}(k) - P^S_{\ell,{\rm model}}(k)] \times {\rm Cov}^{-1}_{\ell,\ell'}(k)[P^S_{\ell',{\rm data}}(k) - P^S_{\ell',{\rm model}}(k)], \tag{D1}$$

where ${\rm Cov}_{\ell,\ell'}$ is the Gaussian covariance matrix between the different multipoles, and $k_{\rm min} = 0.006\ h/{\rm Mpc}$. The number of degrees of freedom $N_{\rm dof}$ is given by $N_{\rm dof} = 2\times N_{\rm bins} - N_{\rm params}$, where $N_{\rm bins}$ is the number of k bins used in the summation and $N_{\rm params}$ is the number of free parameters in the theoretical model. Here, $N_{\rm params} = 2$ for EFT and not 3 because we only fit the first two multipoles [76], and $N_{\rm params} = 1$ for the TNS model.

We increase $k_{\rm max}$ until $\chi^2_{\rm red}(k_{\rm max}) \geq 1$ monotonically. This gives a good indication of where the model does not fit the data so well anymore. We checked that for $k < k_{\rm max}$, chisquared scatters around 1 and it is not systematically less than 1. In the fit we keep cosmology fixed to the COLA simulation's fiducial values and only vary the counterterm coefficients and σ_v.

We use linear theory to model the covariance between the multipoles (see Appendix C of [16] for details). This has been shown to reproduce N-body results up to $k \leq 0.300\ h/{\rm Mpc}$ at $z = 1$. In the covariance matrix we assume a number density of $n = 1\times 10^{-3}\ h^3/{\rm Mpc}^3$ and a survey volume of $V_s = 4\ {\rm Gpc}^3/h^3$ which are similar specifications for a Euclid-like survey [23]. The best fit parameters as well as $k_{\rm max}$ are shown in Table I.

TABLE I. Table showing the maximum $k_{\rm max}[h/{\rm Mpc}]$ used in Eq. (D1) and best fit model parameters for TNS and EFT models found by a least χ^2 fit to the COLA data.

Model	TNS				EFT			
z	0.5	1	1.5	2	0.5	1	1.5	2
$k_{\rm max}$	0.16	0.21	0.27	0.35	0.16	0.21	0.27	0.311
σ_v	7.35	6.26	5.12	4.19	⋯	⋯	⋯	⋯
$c_{2\|\delta_s,0}/k^2_{\rm nl}$	⋯	⋯	⋯	⋯	0.05	0.00	0.00	0.13
$c_{2\|\delta_s,2}/k^2_{\rm nl}$	⋯	⋯	⋯	⋯	13.57	8.96	5.66	1.52
$c_{2\|\delta_s,4}/k^2_{\rm nl}$	⋯	⋯	⋯	⋯	7.34	8.03	6.86	5.73

APPENDIX E: NEGLECTING THE LENSING TERM

Throughout the paper, we have neglected the lensing contribution to the angular power spectrum. In this Appendix, we show that among the three different redshift bins that we used, namely, $\Delta z = 0.1$, $\Delta z = 0.01$, and $\Delta z = 0.001$, lensing is of the same order as the RSD contribution only for $\Delta z = 0.1$ while for the other two redshift bins, it is negligible.

In Fig. 14, we show the ratio of lensing to the RSD term for different redshifts with $\Delta z = 0.01$ (top panel) and $\Delta z = 0.1$ (bottom panel). For $\Delta z = 0.1$, lensing is not negligible when compared to RSD; however, we have shown that for this window width, RSD is also not very significant. For $\Delta z = 0.01$ (top panel), the lensing terms are at most 1% of the RSD terms. It is also interesting to note that for $\Delta z = 0.1$ and $z = 1$ the lensing signal is very small at $\ell < 400$. This comes from the fact that the lensing signal is the sum of the always negative lensing-density correlation and the positive lensing-lensing term. At low redshift, the density term is larger than lensing so that the signal is dominated by the first term and is therefore negative. At sufficiently high redshift when enough lensing has accumulated, the lensing-lensing term starts to dominate and the signal becomes positive. For $\Delta z = 0.1$ this happens roughly at $z = 1$. For $\Delta z = 0.01$ this happens roughly at $z \sim 0.5$ for the low multipoles, $\ell < 200$ while for higher multipoles the positive lensing-lensing signal dominates. Since the cross-correlation lensing (z_2) density (z_1) is significant only for density fluctuations at redshift over which the lensing term is integrated, $z_1 < z_2$, this contribution is smaller for smaller redshift bins.

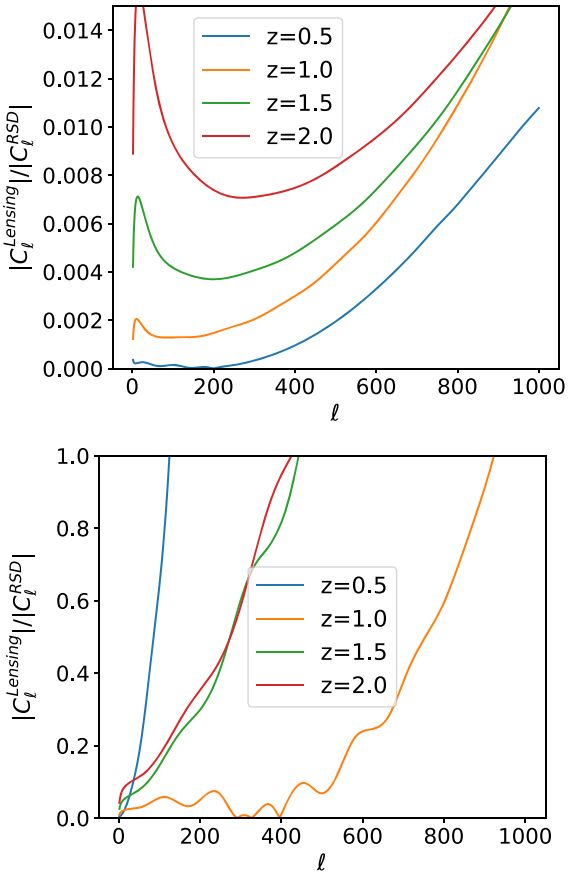

FIG. 14. Ratio of lensing terms to RSD terms for different redshifts with $\Delta z = 0.01$ (top panel) and $\Delta z = 0.1$ (bottom panel).

APPENDIX F: FISHER FORECAST

In this Appendix, we present some details on the Fisher forecast we have performed for RSD detection. We replace each μ^2 term with $A\mu^2$, where A is an artificial amplitude with fiducial value of 1, and our aim is to forecast how precisely we can measure this amplitude. For nonlinear RSD, we simply use the Kaiser formula applied to halofit model; basically replacing $P(k_\parallel, \frac{\ell}{\chi}, \bar{z})$ in Eq. (8) with the Kaiser formula given in Eq. (1), we have

$$\begin{aligned} C_\ell^{\Delta\Delta}(z,z) &= \frac{1}{2\pi\chi^2(\bar{z})}\left[\int_{-\infty}^{+\infty} dk_\parallel (1+2Af\mu^2+A^2f^2\mu^4)P(k_\parallel,\ell/\chi,z)\right] \\ &\equiv C_\ell^{\delta\delta}+2AC_\ell^{\delta\theta}+A^2C_\ell^{\theta\theta}, \end{aligned} \tag{F1}$$

where Δ is the density perturbations in redshift space,

$$\Delta(\boldsymbol{n},z) = \delta(\boldsymbol{n},\mathbf{r}) - \frac{\nabla_z v_z(\mathbf{r})}{aH(z)}, \tag{F2}$$

and $\theta = \nabla_z v_z(\mathbf{r})/aH(z)$, where z is the line-of-sight direction. For the Fisher forecast, we follow a similar approach as the one used in Sec. 4 of [77]. The Fisher matrix element corresponding to the parameters α and β with covariance matrix, $\mathcal{C}$, is given by

$$F_{\alpha\beta} = \sum_\ell \frac{2\ell+1}{2}[(\partial_\alpha \mathcal{C})(\mathcal{C}^{-1})(\partial_\beta \mathcal{C})(\mathcal{C}^{-1})], \tag{F3}$$

which for our case with only one parameter, A, and the covariance matrix being the C_ℓ's, reduces to

$$F_{AA} = \sum_{\ell=2}^{\ell_{\max}} \frac{2\ell+1}{2}\left[\frac{\partial_A C_\ell^{\Delta\Delta}}{C_\ell^{\Delta\Delta}}\right]^2. \tag{F4}$$

The importance of including RSD for galaxy number counts angular power spectrum has been studied in a recent paper [78].

[1] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).
[2] L. Anderson *et al.*, Mon. Not. R. Astron. Soc. **427**, 3435 (2012).
[3] Y.-S. Song, A. Taruya, E. Linder, K. Koyama, C. G. Sabiu, G.-B. Zhao, F. Bernardeau, T. Nishimichi, and T. Okumura, Phys. Rev. D **92**, 043522 (2015).
[4] F. Beutler *et al.* (BOSS Collaboration), Mon. Not. R. Astron. Soc. **466**, 2242 (2017).
[5] C. Blake *et al.*, Mon. Not. R. Astron. Soc. **415**, 2876 (2011).
[6] B. A. Reid *et al.*, Mon. Not. R. Astron. Soc. **426**, 2719 (2012).
[7] E. Macaulay, I. K. Wehus, and H. K. Eriksen, Phys. Rev. Lett. **111**, 161301 (2013).
[8] F. Beutler *et al.* (BOSS Collaboration), Mon. Not. R. Astron. Soc. **443**, 1065 (2014).
[9] H. Gil-Marín *et al.*, Mon. Not. R. Astron. Soc. **460**, 4188 (2016).
[10] F. Simpson, C. Blake, J. A. Peacock, I. Baldry, J. Bland-Hawthorn, A. Heavens, C. Heymans, J. Loveday, and P. Norberg, Phys. Rev. D **93**, 023525 (2016).
[11] M. Rabold and R. Teyssier, Mon. Not. R. Astron. Soc. **467**, 3188 (2017).
[12] A. Schneider, R. Teyssier, J. Stadel, N. E. Chisari, A. M. C. Le Brun, A. Amara, and A. Refregier, J. Cosmol. Astropart. Phys. 03 (2019) 020.
[13] P. Ocvirk *et al.*, arXiv:1811.11192.
[14] N. Kaiser, Mon. Not. R. Astron. Soc. **227**, 1 (1987).
[15] R. Scoccimarro, Phys. Rev. D **70**, 083007 (2004).
[16] A. Taruya, T. Nishimichi, and S. Saito, Phys. Rev. D **82**, 063522 (2010).

[17] T. Okumura, N. Hand, U. Seljak, Z. Vlah, and V. Desjacques, Phys. Rev. D **92**, 103516 (2015).
[18] B. Bose and K. Koyama, J. Cosmol. Astropart. Phys. 08 (2016) 032.
[19] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, Astrophys. J. **761**, 152 (2012).
[20] K. Markovic, B. Bose, and A. Pourtsidou, arXiv:1904.11448.
[21] www.euclid-ec.org.
[22] R. Laureijs *et al.* (EUCLID Collaboration), arXiv:1110.3193.
[23] L. Amendola *et al.*, Living Rev. Relativ. **21**, 2 (2018).
[24] https://wfirst.gsfc.nasa.gov/.
[25] D. Spergel *et al.*, arXiv:1503.03757.
[26] C. J. Walcher, M. Banerji, C. Battistini, C. P. M. Bell, O. Bellido-Tirado, T. Bensby, J. M. Bestenlehner, T. Boller, J. Brynnel, and A. Casey, The Messenger **175**, 12 (2019).
[27] www.desi.lbl.gov.
[28] A. Aghamousa *et al.* (collaboration DESI Collaboration), arXiv:1611.00036.
[29] K. Bandura *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9145**, 914522 (2014).
[30] L. B. Newburgh *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9906**, 99065X (2016).
[31] D. J. Bacon *et al.* (SKA Collaboration), arXiv:1811.02743.
[32] B. Bose, A. Pourtsidou, K. Markovič, and F. Beutler, arXiv:1905.05122.
[33] C. Bonvin and R. Durrer, Phys. Rev. D **84**, 063505 (2011).
[34] A. Challinor and A. Lewis, Phys. Rev. D **84**, 043516 (2011).
[35] F. Montanari and R. Durrer, J. Cosmol. Astropart. Phys. 10 (2015) 070.
[36] W. Cardona, R. Durrer, M. Kunz, and F. Montanari, Phys. Rev. D **94**, 043007 (2016).
[37] S. Tassev, M. Zaldarriaga, and D. Eisenstein, J. Cosmol. Astropart. Phys. 06 (2013) 036.
[38] C. Howlett, M. Manera, and W. J. Percival, Astron. Comput. **12**, 109 (2015).
[39] G. Valogiannis and R. Bean, Phys. Rev. D **95**, 103515 (2017).
[40] H. A. Winther, K. Koyama, M. Manera, B. S. Wright, and G.-B. Zhao, J. Cosmol. Astropart. Phys. 08 (2017) 006.
[41] S. Bharadwaj and S. S. Ali, Mon. Not. R. Astron. Soc. **356**, 1519 (2005).
[42] K. K. Datta, T. R. Choudhury, and S. Bharadwaj, Mon. Not. R. Astron. Soc. **378**, 119 (2007).
[43] E. Castorina and M. White, Mon. Not. R. Astron. Soc. **479**, 741 (2018).
[44] R. Durrer, *The Cosmic Microwave Background* (Cambridge University Press, Cambridge, England, 2008).
[45] D. N. Limber, Astrophys. J. **119**, 655 (1954).
[46] E. Di Dio, R. Durrer, R. Maartens, F. Montanari, and O. Umeh, J. Cosmol. Astropart. Phys. 04 (2019) 053.
[47] N. E. Chisari and A. Pontzen, Phys. Rev. D **100**, 023543 (2019).
[48] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000).
[49] M. Santos *et al.*, *Proc. Sci.*, AASKA14 (2015) 019.
[50] A. Rassat, A. Amara, L. Amendola, F. J. Castander, T. Kitching, M. Kunz, A. Refregier, Y. Wang, and J. Weller, arXiv:0810.0003.
[51] A. F. Heavens, S. Matarrese, and L. Verde, Mon. Not. R. Astron. Soc. **301**, 797 (1998).
[52] T. Matsubara, Phys. Rev. D **77**, 063530 (2008).
[53] L. F. de la Bella, D. Regan, D. Seery, and D. Parkinson, arXiv:1805.12394.
[54] That is, we assume the observer is located at a distance much greater then the box size ($r \gg 1024$ Mpc$/h$), and so all lines of sight are treated as being parallel to the chosen Cartesian axes of the simulation box. Next, we disturb the position of the matter particles using their velocity components (v_x, v_y or v_z).
[55] Note that the hexadecapole at the redshifts considered here is already very small in magnitude and so the exclusion of higher order multipoles will only negligibly affect the form of $P(k,\mu)$.
[56] A. Izard, M. Crocce, and P. Fosalba, Mon. Not. R. Astron. Soc. **459**, 2327 (2016).
[57] B. Bose, K. Koyama, and H. A. Winther, J. Cosmol. Astropart. Phys. 10 (2019) 021.
[58] L. Blot *et al.*, Mon. Not. R. Astron. Soc. **485**, 2806 (2019).
[59] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).
[60] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 07 (**2012**) 051.
[61] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, J. High Energy Phys. 09 (2012) 082.
[62] L. Senatore and M. Zaldarriaga, arXiv:1409.1225.
[63] M. Lewandowski, L. Senatore, F. Prada, C. Zhao, and C.-H. Chuang, Phys. Rev. D **97**, 063526 (2018).
[64] A. Perko, L. Senatore, E. Jennings, and R. H. Wechsler, arXiv:1610.09321.
[65] S. Foreman, H. Perrier, and L. Senatore, J. Cosmol. Astropart. Phys. 05 (2016) 027.
[66] T. Nishimichi and A. Taruya, Phys. Rev. D **84**, 043526 (2011).
[67] A. Taruya, T. Nishimichi, and F. Bernardeau, Phys. Rev. D **87**, 083509 (2013).
[68] T. Ishikawa, T. Totani, T. Nishimichi, R. Takahashi, N. Yoshida, and M. Tonegawa, Mon. Not. R. Astron. Soc. **443**, 3359 (2014).
[69] Y. Zheng and Y.-S. Song, J. Cosmol. Astropart. Phys. 08 (2016) 050.
[70] H. Gil-Marín *et al.*, Mon. Not. R. Astron. Soc. **460**, 4210 (2016).
[71] B. Bose, K. Koyama, W. A. Hellwing, G.-B. Zhao, and H. A. Winther, Phys. Rev. D **96**, 023519 (2017).
[72] D. Blas, M. Garny, and T. Konstandin, J. Cosmol. Astropart. Phys. 01 (2014) 010.
[73] J. Carlson, M. White, and N. Padmanabhan, Phys. Rev. D **80**, 043531 (2009).
[74] B. Jain and E. Bertschinger, Astrophys. J. **431**, 495 (1994).
[75] B. Ghosh, M. Jalilvand, and E. Majerotto, One loop standard perturbation theory calculations, https://doi.org/10.5281/zenodo.3407616 (2019).
[76] The inclusion of the hexadecapole would restrict the determined range we can safely fit to. Further, the monopole and quadrupole contain most of the RSD information so we can omit the hexadecapole from these fits.
[77] M. Jalilvand, E. Majerotto, R. Durrer, and M. Kunz, J. Cosmol. Astropart. Phys. 01 (**2019**) 020.
[78] K. Tanidis and S. Camera, Mon. Not. R. Astron. Soc. **489**, 3385 (2019).