PHYSICAL REVIEW D **81,** 023526 (2010)

Redshift space correlations and scale-dependent stochastic biasing of density peaks

Vincent Desjacques[1,*] and Ravi K. Sheth[2,†]

[1]*Institute for Theoretical Physics, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[2]*Center for Particle Cosmology, University of Pennsylvania, 209 South 33rd Street, Philadelphia, Pennsylvania 19104, USA*
(Received 29 September 2009; published 22 January 2010)

We calculate the redshift space correlation function and the power spectrum of density peaks of a Gaussian random field. Our derivation, which is valid on linear scales $k \lesssim 0.1\ h\mathrm{Mpc}^{-1}$, is based on the peak biasing relation given by Desjacques [Phys. Rev. D, **78**, 103503 (2008)]. In linear theory, the redshift space power spectrum is $P^s_{\rm pk}(k,\mu) = \exp(-f^2\sigma^2_{\rm vel}k^2\mu^2)[b_{\rm pk}(k) + b_{\rm vel}(k)f\mu^2]^2P_\delta(k)$, where μ is the angle with respect to the line of sight, $\sigma_{\rm vel}$ is the one-dimensional velocity dispersion, f is the growth rate, and $b_{\rm pk}(k)$ and $b_{\rm vel}(k)$ are k-dependent linear spatial and velocity bias factors. For peaks, the value of $\sigma_{\rm vel}$ depends upon the functional form of $b_{\rm vel}$. When the k dependence is absent from the square brackets and $b_{\rm vel}$ is set to unity, the resulting expression is assumed to describe models where the bias is linear and deterministic, but the velocities are unbiased. The peak model is remarkable because it has unbiased velocities in this same sense—peak motions are driven by dark matter flows—but, in order to achieve this, $b_{\rm vel}$ must be k dependent. We speculate that this is true in general: k dependence of the spatial bias will lead to k dependence of $b_{\rm vel}$ even if the biased tracers flow with the dark matter. Because of the k dependence of the linear bias parameters, standard manipulations applied to the peak model will lead to k-dependent estimates of the growth factor that could erroneously be interpreted as a signature of modified dark energy or gravity. We use the Fisher formalism to show that the constraint on the growth rate f is degraded by a factor of 2 if one allows for a k-dependent velocity bias of the peak type. Our analysis also demonstrates that the Gaussian smoothing term is part and parcel of linear theory. We discuss a simple estimate of nonlinear evolution and illustrate the effect of the peak bias on the redshift space multipoles. For $k \lesssim 0.1\ h\mathrm{Mpc}^{-1}$, the peak bias is deterministic but k dependent, so the configuration-space bias is stochastic and scale dependent, both in real and redshift space. We provide expressions for this stochasticity and its evolution.

DOI: 10.1103/PhysRevD.81.023526 PACS numbers: 98.80.−k, 95.35.+d, 98.65.Dx, 98.80.Es

I. INTRODUCTION

While velocities are directly measured through their Doppler (red)shifts, accurate measurements of cosmological distances are only available for nearby cosmic objects, and even at these small scales they are plagued with observational biases. Therefore, most observational data are described in terms of redshifts; e.g. three-dimensional (3D) galaxy surveys provide the angular positions and redshifts of galaxies. Redshifts differ from distances by the peculiar velocities (deviations from pure Hubble flow) along the line of sight. These generate systematic differences between the spatial distribution of data in redshift and distance (or real) space which are commonly referred to as redshift distortions [1]. Kaiser [2] first derived an expression which describes the effect of linear peculiar motions on 3D power spectra. References [3,4] provide two very different derivations of this same expression. Whereas the original derivation made no assumption about the form of the density and velocity fields, the other two assume they are Gaussian distributed.

The Kaiser formula has been used to interpret observations of the redshift space clustering of galaxies. The angular dependence of the redshift distortions can be used to measure the logarithmic derivative $f = d\ln D/d\ln a$ or growth rate [5] at multiple redshifts and thus potentially constrain many of the dark energy or modified gravity models (e.g. [6]; for a review of these scenarios, see [7]). Essentially all analyses to date assume that (i) galaxies are biased tracers of the underlying matter field, (ii) the bias is linear, local, and deterministic [2,3,8–10], and (iii) the velocities of the tracers are *unbiased*. In fact, except on the largest scales, the relation between the dark matter and galaxy fields is almost certainly nonlinear, nonlocal, and scale dependent [11]. Our main goal in the present study is to explore what complexities one might expect on smaller scales, where the bias relation is more complicated and where the velocities may also be biased. We do so by investigating the impact of redshift distortions on the correlation function of density maxima in a Gaussian density field.

We have chosen to study density peaks because the statistics of Gaussian random density [12] and velocity fields [13] in a cosmological context, and of the peak distribution in particular, has already received considerable

*dvince@physik.uzh.ch
†shethrk@physics.upenn.edu

attention [14–20]. Some of these results have been used in studies of the nonspherical formation of large-scale structures [21–24]. Others, especially from peaks theory, have been used to interpret the abundance and clustering of rich clusters [25–28]. Density peaks define a well-behaved point process which can account for the discrete nature of dark matter halos and galaxies. On asymptotically large scales, peaks are linearly biased tracers of the dark matter field, and this bias is scale independent [14,17,26]. However, these conclusions are based on a configuration-space argument known as the peak-background split. Extending the description of peak bias to smaller scales is more easily accomplished by working in Fourier space. It has been shown that peaks are linearly biased with respect to the mass, but this bias is k dependent [11,29].

The first part of this paper demonstrates that, in the large-scale limit, the configuration and Fourier-based approaches yield consistent results. This is important, because the first (and only other) study of the redshift space clustering of peaks, Ref. [20], reported that in redshift space peaks behave very differently from the deterministic, linear, and scale-independent biased tracers investigated in [2,3,8,10]. Since the linear bias assumption that is extensively advocated to convert large-scale redshift space measurements into information about the background cosmology [9,30], the fact that peaks might behave very differently is potentially very worrisome. In addition, peak velocities exhibit a k-dependent bias even though peaks locally flow with the dark matter [29]. This is remarkable given that one commonly refers to such flows as having *unbiased* velocities. We explain the origin of this effect and argue that it should be a generic feature of any k-dependent spatial bias model. Again, however, peaks are remarkable because, in the high peak limit where their spatial bias is expected to be linear and scale independent, their velocity bias remains k dependent.

The second part shows that, at the linear order, redshift space distortions for peaks can be recast in a way that retains the simplicity of the original Kaiser formulas [2] while generalizing them to tracers whose linear bias is k dependent. Because the present derivation is based on a model which is supposed to be accurate at smaller scales, we can identify an important term which does not appear in [2]. Furthermore, our analysis reaches very different conclusions from that of [20]. Our peaks-based formula for redshift space distortions, which includes k-dependent linear bias factors for both the density and the velocity fields, has a rich structure. We hope it will serve as a guide for what one might expect in the case of more realistic (nonlinear, nonlocal, scale-dependent) bias prescriptions.

In the last part of this study, we use the Fisher formalism to quantify the extent to which any k-dependent velocity bias of the peak type would degrade the uncertainties on the growth rate f. We also demonstrate the stochastic nature of the peak bias and discuss its evolution with redshift. The peak biasing is interesting because, although it is deterministic in Fourier space, it is stochastic in real space. A final section summarizes our findings and speculates on some implications of the peak model.

Throughout the paper we work in the "distant observer" limit, where the line of sight is oriented along the z direction. In all illustrative examples, we assume a flat ΛCDM cosmology with $\Omega_m = 0.279$, $\Omega_b = 0.0462$, $h = 0.7$, $n_s = 0.96$, and a present-day normalization $\sigma_8 = 0.81$ [31]. It will also be convenient to work with scaled velocities $v_i \equiv \mathrm{v}_i/(aHf)$, where v_i is the (proper) peculiar velocity, $H \equiv d\ln a/dt$, and $f \equiv d\ln D/d\ln a$ with $D(z)$ the linear theory growth factor. At $z = 0.5$ this is $aHf \approx 61\ \mathrm{km\,s^{-1}}\ h\mathrm{Mpc}^{-1}$. As a result, v_i has dimensions of length.

II. PROPERTIES OF DENSITY PEAKS

We begin by reviewing some general properties of peaks in Gaussian random fields. We then discuss the biasing relation which is used in the calculation of the redshift space correlation of density maxima.

A. Spectral moments

The statistical properties of density peaks depend not only on the underlying density field, but also on its first and second derivatives. We are, therefore, interested in the linear (Gaussian) density field $\delta(\boldsymbol{x})$ and its first and second derivatives, $\partial_i\delta(\boldsymbol{x})$ and $\partial_i\partial_j\delta(\boldsymbol{x})$. In this regard, it is convenient to introduce the normalized variables $\nu = \delta(\boldsymbol{x})/\sigma_0$ and $u = -\nabla^2\delta(\boldsymbol{x})/\sigma_2$, where the σ_n are the spectral moments of the matter power spectrum,

$$\sigma_n^2 \equiv \frac{1}{2\pi^2}\int_0^\infty dk k^{2(n+1)} P_\delta(k,z)\hat{W}(k,R_S)^2. \qquad (1)$$

Here, $P_\delta(k,z)$ denotes the dimensionless power spectrum of the linear density field at redshift z, and $\hat{W}$ is a spherically symmetric smoothing kernel of length R_S (a Gaussian filter will be adopted throughout this paper) introduced to ensure convergence of all spectral moments. We will use the notation $P_{\delta_S}(k,z)$ to denote $P_\delta(k,z)\hat{W}(k,R_S)^2$. The ratio σ_0/σ_1 is proportional to the typical separation between zero-crossings of the density field [17]. For subsequent use, we also define the spectral parameters

$$\gamma_n = \frac{\sigma_n^2}{\sigma_{n-1}\sigma_{n+1}} \qquad (2)$$

which reflect the range over which $k^{2n+1}P_{\delta_S}(k,z)$ is large.

We will also need the quantities analogous to σ_n^2 but for nonzero lag:

$$\xi_\ell^{(n)}(r) = \frac{1}{2\pi^2}\int_0^\infty dk k^{2(n+1)} P_{\delta_S}(k,z) j_\ell(kr), \qquad (3)$$

where $j_\ell(x)$ are spherical Bessel functions. As ℓ gets larger, these harmonic transforms become increasingly sensitive to small-scale power.

Finally, we note that the autocorrelations and cross correlations of the fields v_i, δ, $\partial_i\delta$, and $\partial_i\partial_j\delta$ can generally be decomposed into components with definite transformation properties under rotations. Reference [29] gives explicit expressions for the isotropic and homogeneous linear density field.

B. Smoothing scale and peak height

The peak height ν and the filtering radius R_S could, in principle, be treated as two independent variables. However, in order to make as much connection with dark matter halos (and, to a lesser extent, galaxies) as possible, we assume that density maxima with height $\nu = \delta_{\rm sc}(z)/\sigma_0(R_S)$ identified in the primeval density field smoothed at scale R_S are related to dark matter halos of mass M_S collapsing at redshift z, where $\delta_{\rm sc}(z)$ is the critical density for collapse at z in the spherical model [32,33]. For the sake of illustration, we will present results at $z = 0.5$. In the background cosmology that we assume, the linear critical density for (spherical) collapse at $z = 0.5$ is $\delta_{\rm sc} \approx 1.681$. The Gaussian smoothing scale at which $\nu = 1$ is $R_\star \approx 1.3\ h^{-1}$Mpc, so the characteristic mass scale is $M_\star \approx 6.5 \times 10^{11}{\rm M}_\odot/h$.

While there is a direct correspondence between massive halos in the evolved density field and the largest maxima of the initial density field, the extent to which galaxy-sized halos trace the initial density maxima is unclear. Therefore, we will only consider mass scales M_S significantly larger than the characteristic mass for clustering, $M_\star$, for which the peak model is expected to work best. We will present results at redshift $z = 0.5$ for two (Gaussian) filtering lengths, $R_S = 2.5\ h^{-1}$Mpc and $R_S = 4\ h^{-1}$Mpc; these correspond to masses $M_S = 1.9 \times 10^{13}{\rm M}_\odot/h$ and $7.8 \times 10^{13}{\rm M}_\odot/h$, which roughly match the mean redshift and typical mass of halos harboring luminous red galaxies in the Sloan Digital Sky Survey [34–36]. This makes $\sigma_0/\sigma_1 \approx 3.2\ h^{-1}$Mpc and $4.9\ h^{-1}$Mpc. To help set scales in the discussion which follows, the associated values of (ν, b_ν, b_ζ) are $(2.1, 1.0, 16.4\ h^2{\rm Mpc}^{-2})$ and $(2.8, 2.8, 43.0\ h^2{\rm Mpc}^{-2})$. The three-dimensional velocity dispersion of these peaks is $\sigma_{-1}^2(1-\gamma_0^2)$: for our two smoothing scales, this corresponds to $(7.12\ h^{-1}{\rm Mpc})^2$ and $(6.66\ h^{-1}{\rm Mpc})^2$ [recall that our velocities are in units of aHf, so they have dimensions of (length)^2].

C. Biasing

The large-scale asymptotics $r \to \infty$ of the two-point correlation $\xi_{\rm pk}(r)$ and line-of-sight mean streaming $[v_{12}\cdot\hat{\boldsymbol{r}}](r)$ for *discrete* local maxima of height ν can be thought of as arising from the *continuous*, nonlinear bias relation [29]

$$\begin{aligned}\delta n_{\rm pk}(\boldsymbol{x}) &= b_\nu\delta_S(\boldsymbol{x}) - b_\zeta\nabla^2\delta_S(\boldsymbol{x}),\\ \boldsymbol{v}_{\rm pk}(\boldsymbol{x}) &= \boldsymbol{v}_S(\boldsymbol{x}) - \frac{\sigma_0^2}{\sigma_1^2}\nabla\delta_S(\boldsymbol{x}),\end{aligned} \tag{4}$$

where $\boldsymbol{v}_S$ is the dark matter velocity smoothed at scale R_S (so as to retain only the large-scale, coherent motion of the peak), and the bias parameters b_ν and b_ζ are

$$\begin{aligned}b_\nu &= \frac{1}{\sigma_0}\left(\frac{\nu-\gamma_1\bar{u}}{1-\gamma_1^2}\right),\\ b_\zeta &= \frac{1}{\sigma_2}\left(\frac{\bar{u}-\gamma_1\nu}{1-\gamma_1^2}\right) = \frac{\sigma_0^2}{\sigma_1^2}\frac{(\nu-\sigma_0 b_\nu)}{\sigma_0}.\end{aligned} \tag{5}$$

Here, $\bar{u}$ denotes the mean curvature of the peaks. Furthermore, b_ν is dimensionless, whereas b_ζ has units of (length)^2. Note that b_ν is precisely the amplification factor found by [17] which neglected derivatives of the density correlation function (i.e. their analysis assumes $b_\zeta \equiv 0$). We emphasize that Eq. (4) is the only bias relation that can account for the first order peak correlation and mean streaming.

Strictly speaking, the bias relation (4) is nonlocal because of the smoothing. In configuration space, the peak bias $b_{\rm pk}$ at first order could thus be defined as the convolution

$$(b_{\rm pk}\otimes\delta)(\boldsymbol{x}) \equiv (b_\nu - b_\zeta\nabla^2)\delta_S(\boldsymbol{x}). \tag{6}$$

In Fourier space, this becomes

$$b_{\rm pk}(\boldsymbol{k}) \equiv (b_\nu + b_\zeta k^2)\hat{W}(k, R_S), \tag{7}$$

so it has the same functional form as Eq. (57) in Ref. [11] where density extrema were considered. Our coefficients thus agree with those of [11] only in the limit $\nu \gg 1$, in which nearly all extrema are local maxima.

This bias relation is distinct from either linear [14] or nonlinear [37–39] biasing transformations of the density field for which $b_\zeta = 0$. Note, in particular, that Eq. (7) shows that *local* bias schemes can generate k-dependent bias factors if the bias relation involves differential operators. Furthermore, when $\nu \gg 1$, then $\bar{u} \to \gamma_1\nu$, so that $\sigma_0 b_\nu \to \nu$ and $\sigma_2 b_\zeta \to 0$ [40]. This is clearly seen in Fig. 1 where the biasing factors are plotted as a function of the peak height. Thus, the spatial bias of the highest peaks is expected to become scale independent, approaching the local deterministic relation of linearly biased tracers for which there is no k-dependent bias. However, notice that the k dependence in the velocity bias remains. We will return to this point shortly.

D. Relation to peak-background split

There is another route for estimating large-scale bias of peaks [26] which utilizes the peak-background split argument [17,41–43]. This approach is *very* different from ours, because it is based on configuration-space counts-in-cells

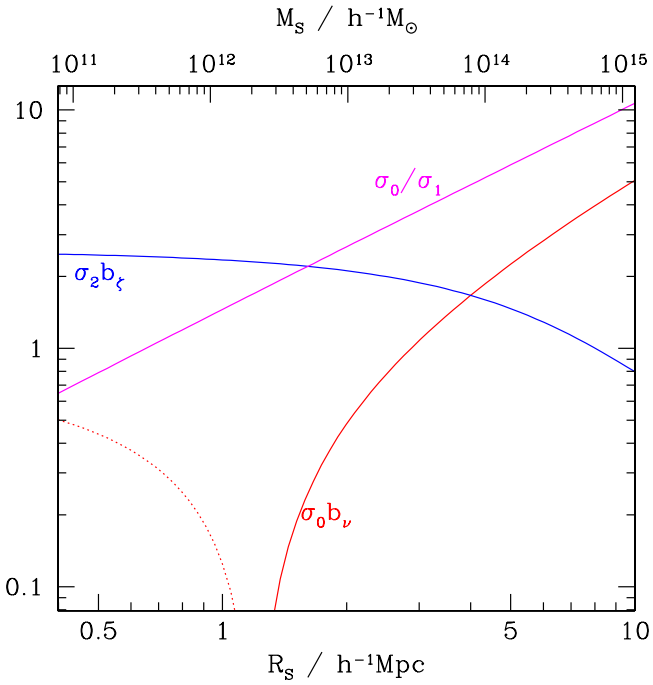

FIG. 1 (color online). Bias parameters $\sigma_0 b_\nu$, $\sigma_2 b_\zeta$ and the ratio of spectral moments σ_0/σ_1 as a function of the filtering scale. Results are shown for density maxima of height $\nu = \delta_{\rm sc}/\sigma_0$ at redshift $z = 0.5$. Dotted curves denote negative values. The linear spatial bias of peaks becomes scale independent in the limit $\nu \to \infty$. However, the k dependence of the velocity bias, which is controlled by σ_0/σ_1, remains and even increases with the peak height.

statistics. In particular, it makes no mention of the bias in Fourier space.

The large-scale bias predicted by this approach is [26]

$$b_{\rm pkbs} \equiv -\frac{1}{\sigma_0\nu}\frac{\partial \ln[\bar{n}_{\rm pk}(\nu)]}{\partial \ln\nu} \tag{8}$$

where

$$\bar{n}_{\rm pk}(\nu) = \frac{1}{(2\pi)^2 R_1^3} e^{-\nu^2/2} G_0(\gamma_1, \gamma_1\nu) \tag{9}$$

is the differential averaged number density of peaks in the range ν to $\nu + d\nu$ [17]. Here $R_1 = \sqrt{3}\sigma_1/\sigma_2 \propto R_S$ characterizes the typical radius of density maxima, and G_0 is given by setting $n = 0$ in our Eq. (A10). Therefore,

$$b_{\rm pkbs} = \frac{\nu^2 + g_1}{\sigma_0\nu}, \tag{10}$$

where

$$g_1 \equiv -\frac{\partial \ln G_0(\gamma_1, y)}{\partial \ln y}\bigg|_{y=\gamma_1\nu}. \tag{11}$$

Performing the derivative yields

$$g_1 = -\gamma_1\nu\frac{G_1(\gamma_1, \gamma_1\nu)/G_0(\gamma_1, \gamma_1\nu) - \gamma_1\nu}{1-\gamma_1^2}, \tag{12}$$

where G_1 is given by Eq. (A10) with $n = 1$. However, $G_1/G_0 \equiv \bar{u}$ [see the discussion immediately following Eq. (A10)], so

$$g_1 = -\gamma_1\nu\left(\frac{\bar{u} - \gamma_1\nu}{1-\gamma_1^2}\right) = -\gamma_1\nu\sigma_2 b_\zeta. \tag{13}$$

The last equality follows from the definition of b_ζ [Eq. (5)]. Equations (2) and (5) eventually imply that

$$-\gamma_1\nu\sigma_2 b_\zeta = -\nu(\nu - b_\nu\sigma_0), \tag{14}$$

so

$$b_{\rm pkbs} = \frac{\nu^2 - \nu(\nu - b_\nu\sigma_0)}{\sigma_0\nu} = b_\nu. \tag{15}$$

This demonstrates that the large-scale, constant, deterministic bias factor returned by the peak-background split approach is exactly the same as in our approach, when one considers scales large enough such that the k dependence associated with the b_ζ term can be ignored.

This is very reassuring for two reasons. First, recall that our expressions for b_ν and b_ζ only agree with those given in [11] in the limit $\nu \gg 1$ (in which extrema are almost certainly peaks). The analysis above shows that our b_ν is the appropriate generalization to lower ν. And second, the peak-background split approximation has been shown to provide an excellent description of large-scale peak bias in simulations [26,44]. Since our expressions reproduce this limit, we have confidence that our approach will provide a good approximation on the smaller scales where the peak-background split fails (i.e., where the bias $b_{\rm pk}$ becomes scale dependent).

E. Power spectra and correlation functions

Using the bias relations (4), it is straightforward to show that the real-space cross-power and autopower spectra are

$$P_{{\rm pk},\delta}(k) = (b_\nu + b_\zeta k^2) P_\delta(k)\hat{W}(k, R_S), \tag{16}$$

$$P_{\rm pk}(k) = (b_\nu + b_\zeta k^2)^2 P_{\delta_S}(k). \tag{17}$$

We have omitted the explicit redshift and ν dependence for brevity. The corresponding relations for the correlation functions are

$$\xi_{{\rm pk},\delta}(r) = b_\nu\xi_0^{(0)}(r) + b_\zeta\xi_0^{(1)}(r), \tag{18}$$

$$\begin{aligned}\xi_{\rm pk}(r) &= b_\nu^2\xi_0^{(0)}(r) + 2b_\nu b_\zeta\xi_0^{(1)}(r) + b_\zeta^2\xi_0^{(2)}(r) \\ &\equiv b_\xi^2(r)\xi_0^{(0)}(r),\end{aligned} \tag{19}$$

where the final expression defines the (scale-dependent) peak bias factor in configuration space. As shown in [29] [for $\xi_{\rm pk}(r)$] and in Appendix A [for $\xi_{{\rm pk},\delta}(r)$], these expres-

sions agree with those obtained from a rather lengthy derivation based on the peak constraint, which involves joint probability distributions of the density field and its derivatives. It is worth noticing that, while expressions (17) and (19) for the autopower spectrum and correlation are only valid at first order in the correlation functions $\xi_\ell^{(n)}$, the cross-power spectrum (16) and correlation (18) are exact to all orders.

F. Velocities

In what follows, we will be interested in redshift space quantities, for which the velocity field also matters. The bias of peak velocities is particularly simple in Fourier space. Taking the divergence of Eq. (4), we find

$$\theta_{\rm pk}(\boldsymbol{x}) \equiv \nabla\cdot\boldsymbol{v}_{\rm pk}(\boldsymbol{x}) = \nabla\cdot\boldsymbol{v}_S(\boldsymbol{x}) - \frac{\sigma_0^2}{\sigma_1^2}\nabla^2\delta_S(\boldsymbol{x}). \quad (20)$$

The linear continuity equation stipulates that $\theta_S(\boldsymbol{x}) \equiv \nabla\cdot\boldsymbol{v}_S(\boldsymbol{x}) = -\delta_S(\boldsymbol{x})$, so the result of Fourier transforming the expression above implies that

$$\theta_{\rm pk}(\boldsymbol{k}) = \left(1 - \frac{\sigma_0^2}{\sigma_1^2}k^2\right)\hat{W}(k,R_S)\theta(\boldsymbol{k}) \equiv b_{\rm vel}(\boldsymbol{k})\theta(\boldsymbol{k}). \quad (21)$$

This defines the peak-velocity bias factor, $b_{\rm vel}(\boldsymbol{k})$, which depends on k but not on ν. As seen in Fig. 1, the ratio σ_0/σ_1 increases monotonically with the filtering scale such that, even in the limit $R_S \to \infty$ ($\nu \to \infty$) where the spatial bias is linear [$b_\xi(r) \approx b_\nu$], the peak velocities remain k dependent. In general, the linear bias approximation $\delta n_{\rm pk} = b_\nu\delta_S$ with unbiased velocities $\boldsymbol{v}_{\rm pk} = \boldsymbol{v}_S$ will provide a good description of the large-scale properties of density peaks only when $k \ll \min[\sqrt{b_\nu/b_\zeta}, \sigma_1/\sigma_0]$. For density peaks of height $\nu \gtrsim 1$, the square root approximately is $(\sigma_1/\sigma_0)(\nu/\sqrt{3})$. The above condition thus becomes $k \ll \sigma_1/\sigma_0$. For the density maxima considered here, this implies that the k-independent linear bias approximation will be accurate for $k \ll 0.1\ h{\rm Mpc}^{-1}$.

The three-dimensional velocity dispersion of peaks is known to be smaller than that of the dark matter [17,45–47]:

$$\sigma_{\rm vpk}^2 = \sigma_{-1}^2(1-\gamma_0^2). \quad (22)$$

Notice that Eq. (21) for the peak-velocity bias yields the same number,

$$\sigma_{\rm vpk}^2 = \frac{1}{2\pi^2}\int_0^\infty dk P_{\delta_S}(k) b_{\rm vel}^2(k), \quad (23)$$

as it should, but that

$$\sigma_{\rm vpk}^2 = \frac{1}{2\pi^2}\int_0^\infty dk P_{\delta_S}(k) b_{\rm vel}(k) \quad (24)$$

also. If we regard the integral over one power of $b_{\rm vel}$, say $\langle b_{\rm vel}\rangle$, as the peak-dark matter velocity variance at the same point (when smoothed on the scale of the peak), then the fact that $\langle b_{\rm vel}\rangle = \langle b_{\rm vel}^2\rangle$ indicates that, at the position of the peak, the velocities of the peak and the mass are the same. This can also be seen in the average bias relation, Eq. (4): at the position of the peak the gradient of the density vanishes (by definition), and so $\boldsymbol{v}_{\rm pk}(\boldsymbol{x}_{\rm pk}) = \boldsymbol{v}(\boldsymbol{x}_{\rm pk})$. The peak-velocity dispersion is lower than that of the mass because large-scale flows are more likely to be directed towards peaks than to be oriented randomly. This illustrates an important point: peaks are biased tracers which move with the dark matter flows—so although there is no physical bias in the velocities, there is a statistical bias which arises from the spatial bias. In the case of peaks, the spatial bias implies that $b_{\rm vel}$ is k dependent, and this introduces k dependence into a number of peak-velocity statistics [we provide an explicit calculation of this in Eq. (36) below]. This is almost certainly true in general: k dependence of the spatial bias will lead to k dependence of $b_{\rm vel}$ even if the tracers flow with the dark matter. Note, however, that this is not a necessary condition since, for the highest peaks, the velocity bias remains scale dependent even though the spatial bias has no k dependence.

The equality $\langle b_{\rm vel}\rangle = \langle b_{\rm vel}^2\rangle$ does not uniquely constrain the velocity bias. For instance, the choice $b_{\rm vel}(k) = 1 - (\sigma_{-1/2}^2/\sigma_0^2)k$ also has $\langle b_{\rm vel}\rangle = \langle b_{\rm vel}^2\rangle$. However, if we think of the velocity bias as a real-space operator $b_{\rm vel}(\boldsymbol{x})$ that maps a vector (velocity) field onto another vector field, then for homogeneous and isotropic random fields $b_{\rm vel}(\boldsymbol{x})$ *must* transform as a scalar under rotations. Hence, it must be built from powers of the Laplacian ∇^2, and this brings down a factor of k^2 upon a Fourier transformation. Therefore, we generically expect the lowest order k dependence to scale as $b_{\rm vel}(k) \equiv 1 - R_{\rm vel}^2k^2$ (for some constant $R_{\rm vel}$), at least for tracers whose spatial bias relation can be expressed as a local mapping of the (smoothed) density and its derivatives.

Before concluding, we emphasize that $b_{\rm pk}(k)$ and $b_{\rm vel}(k)$ are *first order* bias parameters. We expect contributions from higher order spatial and velocity bias parameters to become more important as k increases, but calculating them is beyond the scope of this paper.

III. REDSHIFT SPACE CLUSTERING OF DENSITY MAXIMA

We derive three estimates of the redshift space clustering of peaks. The first generalizes the formulation of [2] based on linear theory of gravitational instability; it furnishes a simple estimate of the power spectrum. The second extends the probabilistic interpretation of [3,5]; it provides an expression for the correlation function. Reference [48] has emphasized that, within the context of linear theory, this description of the correlation function is exact whereas that of [2] is only approximate. Analytic approximations based on a probabilistic treatment lead to terms which, upon

Fourier transforming to obtain the power spectrum, are lacking in the approach of [2]. This has recently been emphasized by [10]. Finally, our third estimate shows that if one Fourier transforms at an earlier stage in the analysis, one obtains a slightly more intuitive expression for the redshift space power. We examine in detail the implications of this approach for density peaks, despite the fact that much of this was already done by [20], for the reasons stated in the Introduction.

A. Simple estimate of redshift space clustering

The redshift space coordinate (also in h^{-1}Mpc since velocities are in unit of length) is given by $\boldsymbol{s} = (s_\parallel, \mathbf{s}_\perp)$,

$$\boldsymbol{s} = \boldsymbol{x} + f[\boldsymbol{v}(\boldsymbol{x}) \cdot \hat{\boldsymbol{z}}]\hat{\boldsymbol{z}}, \tag{25}$$

where $\hat{\boldsymbol{z}}$ is the unit vector along the line of sight. Therefore, at the lowest order, the redshift space density contrast is related to that in real space by

$$\delta^s(k, \mu) = \delta(k) + f\mu^2\theta(k), \tag{26}$$

where μ is the cosine of the angle with the line of sight [2]. For peaks, this becomes

$$\begin{aligned}\delta n^s_{\rm pk}(k, \mu) &= \delta n_{\rm pk}(\boldsymbol{k}) + f\mu^2\theta_{\rm pk}(\boldsymbol{k}) \\ &= b_{\rm pk}(\boldsymbol{k})\delta(\boldsymbol{k}) + f\mu^2 b_{\rm vel}(\boldsymbol{k})\delta(\boldsymbol{k}) \\ &= \left[1 + \frac{b_{\rm vel}(k)}{b_{\rm pk}(k)} f\mu^2\right]\delta n_{\rm pk}(k)\end{aligned} \tag{27}$$

upon insertion of the peak bias relation (4). Note that $f\mu^2$ is now multiplied by a k-dependent factor.

Using the former relation, the calculation of the redshift space power spectra at leading order is straightforward and yields

$$\begin{aligned}P^{s0}_{{\rm pk},\delta}(k, \mu) = (b_{\rm pk}(k) + [b_{\rm vel}(k) + b_{\rm pk}(k)]f\mu^2 \\ + b_{\rm vel}(k)f^2\mu^4)P_\delta(k),\end{aligned} \tag{28}$$

$$P^{s0}_{\rm pk}(k, \mu) = (b_\nu + b_\zeta k^2 + b_{\rm vel}(k)f\mu^2)^2 P_\delta(k) \tag{29}$$

(the reason for introducing the superscript 0 will become clear shortly). However, the corresponding expressions for the redshift space correlations are lengthy; we provide them later in this section.

It is conventional to write the redshift space power spectrum in terms of the real-space one,

$$P^{s0}_{\rm pk}(k, \mu) = \left[1 + \frac{b_{\rm vel}(k)}{b_{\rm pk}(k)} f\mu^2\right]^2 P_{\rm pk}(k). \tag{30}$$

Parameter constraints are then derived from the angular dependence of $P^s_{\rm pk}$, under the assumption of linear scale-independent bias, for which $b_\zeta = 0$ and $b_{\rm vel} = 1$, so the term which multiplies μ^2 is f/b_ν, and b_ν is assumed to be a constant. Our analysis shows that, for peaks, this prefactor is k dependent, and it depends on peak height. The window functions cancel out, leaving us with

$$\begin{aligned}\frac{b_{\rm vel}(k)}{b_{\rm pk}(k)} &\approx \frac{1}{b_\nu}\left[1 - k^2\left(\frac{\nu}{\sigma_0 b_\nu}\right)\frac{\sigma_0^2}{\sigma_1^2}\right] \qquad (k \ll 1) \\ &\approx -\left(\frac{\nu}{\sigma_0} - b_\nu\right)^{-1}\left[1 - \frac{1}{k^2}\left(\frac{\nu}{\sigma_0 b_\nu} - 1\right)^{-1}\frac{\sigma_1^2}{\sigma_0^2}\right] \\ &\quad (k \gg 1).\end{aligned} \tag{31}$$

Hence, unless care is taken, this will lead to constraints which depend on k even in the limit $\nu \gg 1$ where $b_\nu \to \nu/\sigma_0$.

B. Probabilistic treatment

In linear theory, the redshift space two-point correlation function ξ^s is related to that in real space by a convolution of the two-point correlation function in real space, $\xi(r)$, with the probability distribution for velocities along the line of sight [3,5]:

$$1 + \xi^s(s_\parallel, \mathbf{s}_\perp) = \int \frac{dyK(y)}{\sqrt{2\pi}f\sigma_{12}(r)} \exp\left[-\frac{(s_\parallel - y)^2}{f^2\sigma^2_{12}(r)}\right], \tag{32}$$

where

$$\begin{aligned}K(y) = 1 + \xi(r) + \left(\frac{y}{r}\right)\frac{v_{12}(r)}{\sigma_{12}(r)}\left(\frac{s_\parallel - y}{f\sigma_{12}(r)}\right) - \frac{1}{4}\left(\frac{y}{r}\right)^2\frac{v^2_{12}(r)}{\sigma^2_{12}(r)} \\ \times\left[1 - \left(\frac{s_\parallel - y}{f\sigma_{12}(r)}\right)^2\right].\end{aligned} \tag{33}$$

Here, $v_{12}(r)$ and $\sigma_{12}(r)$ are the mean and dispersion of the pairwise velocity distribution of pairs separated by r in real space (note that $r^2 = y^2 + \mathbf{s}^2_\perp$). As emphasized by [48], within the context of linear theory and the plane-parallel approximation, this expression is exact. This formulation is usually referred to as the "streaming" model. It should be noted that random pairs in real space are mapped to real space differently at different separations r because the pairwise velocity distribution depends on scale [10].

Equation (32) can be generalized to give $1 + \xi^s_{\rm pk}$, the redshift space correlation function of peaks, simply by replacing v_{12} and σ_{12} with the expressions appropriate for peaks [20]. At first order, these are

$$\begin{aligned}v_{12}(r, \mu) = [1 + \xi_{\rm pk}]^{-1} \times \left[2b_\nu\left(\frac{\sigma_0^2}{\sigma_1^2}\xi_1^{(1/2)} - \xi_1^{(-1/2)}\right)\right. \\ \left. + 2b_\zeta\left(\frac{\sigma_0^2}{\sigma_1^2}\xi_1^{(3/2)} - \xi_1^{(1/2)}\right)\right]L_1(\mu),\end{aligned} \tag{34}$$

$$\sigma_{12}^2(r,\mu) = \left[\frac{2}{3}(1-\gamma_0^2)\sigma_{-1}^2 + \frac{2}{3}\frac{\sigma_0^2}{\sigma_1^2}\left(2\xi_0^{(0)} - \frac{\sigma_0^2}{\sigma_1^2}\xi_0^{(1)}\right) - \frac{2}{3}\xi_0^{(-1)}\right] - \frac{4}{3}\left[\frac{\sigma_0^2}{\sigma_1^2}\left(2\xi_2^{(0)} - \frac{\sigma_0^2}{\sigma_1^2}\xi_2^{(1)}\right) - \xi_2^{(-1)}\right] \times L_2(\mu), \quad (35)$$

where $\mu = \hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{z}}$ is the cosine of the angle between the line of separation and the line of sight, and the $L_\ell(\mu)$ are Legendre polynomials [49].

Appendix A demonstrates that Eq. (34) exactly reproduces the result of a lengthy derivation based on the peak constraint. Note, however, that it can be derived simply from setting

$$v_{12}(r,\mu) \equiv \frac{\langle(1+\delta n_{\rm pk,1})(1+\delta n_{\rm pk,2})(\boldsymbol{v}_{\rm pk,1}-\boldsymbol{v}_{\rm pk,2})\cdot\hat{\boldsymbol{z}}\rangle}{\langle(1+\delta n_{\rm pk,1})(1+\delta n_{\rm pk,2})\rangle}, \quad (36)$$

where the subscripts 1 and 2 indicate positions separated by $\boldsymbol{r}$, and the average is over all peak pairs with separation r. The correspondence with Eq. (34) can be seen by noting that, in k space, the spatial bias from $\delta n_{\rm pk}$ is the sum of two terms, one of which is proportional to k^2 [Eq. (7)], and the velocity bias [Eq. (21)] introduces additional k^2 terms which come with factors of $(\sigma_0/\sigma_1)^2$. Each additional factor of k^2 changes $\xi_1^{(n)}$ to $\xi_1^{(n+1)}$.

The first term on the right-hand side of Eq. (35) is twice the (one-dimensional) velocity dispersion of peaks; recall that it is reduced by a factor of $1-\gamma_0^2$ relative to that of dark matter [see Eq. (22)].

1. Approximating the integral

When $\sigma_{12} \ll s_\parallel$, the Gaussian term in the expression above will be sharply peaked around $y = s_\parallel$. Expanding ξ, v_{12}, and σ_{12} about their redshift space values yields

$$\xi^s \approx \xi - f v_{12}' + \tfrac{1}{2}f^2\sigma_{12}^{2\prime\prime} + \tfrac{1}{2}f^2\xi''\sigma_{12}^2|_\infty, \quad (37)$$

where all quantities in the right-hand side are evaluated at $\boldsymbol{s}$ and primes denote derivatives with respect to $s_\parallel$ (recall that $s^2 = s_\parallel^2 + \mathbf{s}_\perp^2$). Equation (37) describes the large-scale limit of the redshift space correlation function, in which derivatives of the real-space correlation and pairwise moments (i.e. the distortions) are small [10]. When applied to dark matter rather than density peaks, the Fourier transform of the first three terms on the right-hand side yields Kaiser's formula [3,48]. The fourth term arises because the pairwise velocity dispersion does not vanish even in the large-scale limit [10]. We show below that Fourier transforming the analogous terms for peaks gives Eq. (29).

The derivatives of $\xi_{\rm pk}(\boldsymbol{s})$, $v_{12}(\boldsymbol{s})$, and $\sigma_{12}^2(\boldsymbol{s})$ with respect to the line-of-sight distance $s_\parallel$ can be evaluated using $ds/ds_\parallel = \mu$ and $d\mu/ds_\parallel = (1-\mu^2)s^{-1}$, which follow from the fact that $s^2 = s_\parallel^2 + s_\perp^2$. The following relations are useful:

$$\begin{aligned}\frac{d^2\xi_0^{(n)}}{ds_\parallel^2} &= \frac{2}{3}\xi_2^{(n+1)}L_2(\mu) - \frac{1}{3}\xi_0^{(n+1)},\\ \frac{d}{ds_\parallel}[\xi_1^{(n)}L_1(\mu)] &= -\frac{2}{3}\xi_2^{(n+1/2)}L_2(\mu) + \frac{1}{3}\xi_0^{(n+1/2)} \times \frac{d^2}{ds_\parallel^2}[\xi_2^{(n)}L_2(\mu)]\\ &= \frac{12}{35}\xi_4^{(n+1)}L_4(\mu) - \frac{11}{21}\xi_2^{(n+1)}L_2(\mu) + \frac{2}{15}\xi_0^{(n+1)}. \end{aligned} \quad (38)$$

As a rule, terms in $\xi_\ell^{(n)}$ appear always multiplied by the Legendre polynomial of order ℓ. The lowest even polynomials are $L_0(\mu) = 1$, $L_2(\mu) = (3\mu^2-1)/2$, and $L_4(\mu) = (35\mu^4 - 30\mu^2 + 3)/8$. The calculation of the redshift space correlation of peaks $\xi_{\rm pk}^s(s,\mu)$ is now straightforward. Adding all terms together, we find

$$\begin{aligned}\xi_{\rm pk}^s(s,\mu) = {}& \frac{8}{35}f^2\left(\xi_4^{(0)} - 2\frac{\sigma_0^2}{\sigma_1^2}\xi_4^{(1)} + \frac{\sigma_0^4}{\sigma_1^4}\xi_4^{(2)}\right)L_4(\mu) + \left\{\frac{4}{3}f\left[b_\nu\left(\frac{\sigma_0^2}{\sigma_1^2}\xi_2^{(1)} - \xi_2^{(0)}\right) + b_\zeta\left(\frac{\sigma_0^2}{\sigma_1^2}\xi_2^{(2)} - \xi_2^{(1)}\right)\right]\right.\\ & - \frac{4}{7}f^2\left(\xi_2^{(0)} - 2\frac{\sigma_0^2}{\sigma_1^2}\xi_2^{(1)} + \frac{\sigma_0^4}{\sigma_1^4}\xi_2^{(2)}\right) + \frac{2}{9}f^2(1-\gamma_0^2)\sigma_{-1}^2(b_\nu^2\xi_2^{(1)} + 2b_\nu b_\zeta\xi_2^{(2)} + b_\zeta^2\xi_2^{(3)})\Big\}L_2(\mu) + b_\nu^2\xi_0^{(0)}\\ & + 2b_\nu b_\zeta\xi_0^{(1)} + b_\zeta^2\xi_0^{(2)} - \frac{2}{3}f\left[b_\nu\left(\frac{\sigma_0^2}{\sigma_1^2}\xi_0^{(1)} - \xi_0^{(0)}\right) + b_\zeta\left(\frac{\sigma_0^2}{\sigma_1^2}\xi_0^{(2)} - \xi_0^{(1)}\right)\right] + \frac{1}{5}f^2\left(\xi_0^{(0)} - 2\frac{\sigma_0^2}{\sigma_1^2}\xi_0^{(1)} + \frac{\sigma_0^4}{\sigma_1^4}\xi_0^{(2)}\right)\\ & - \frac{1}{9}f^2(1-\gamma_0^2)\sigma_{-1}^2(b_\nu^2\xi_0^{(1)} + 2b_\nu b_\zeta\xi_0^{(2)} + b_\zeta^2\xi_0^{(3)}). \end{aligned} \quad (39)$$

As can be seen, there are harmonics up to $\xi_\ell^{(3)}(s)$ which arise from the second derivative $\xi''(\boldsymbol{s})$ in Eq. (37). These terms are significant only at distances less than a few smoothing radii and across the baryon acoustic feature where the density correlation $\xi_0^{(0)}$ changes rapidly [29]. Furthermore, terms linear in f arise only from the derivative of the pairwise velocity, $-f v_{12}'(\boldsymbol{s})$.

The redshift space power spectrum $P^s_{\rm pk}(k,\,\mu)$ in this approximation is obtained simply by Fourier transforming Eq. (39). For the sake of completeness,

$$\begin{aligned} P^s_{\rm pk}(k,\,\mu) &= \frac{8}{35}\mathcal{B}^2(k)L_4(\mu)P_{\rm pk}(k) + \left[\frac{4}{3}\mathcal{B}(k) + \frac{4}{7}\mathcal{B}^2(k) - \frac{2}{9}f^2k^2(1-\gamma_0^2)\sigma^2_{-1}\right]L_2(\mu)P_{\rm pk}(k) \\ &\quad + \left[1 + \frac{2}{3}\mathcal{B}(k) + \frac{1}{5}\mathcal{B}^2(k) - \frac{1}{9}f^2k^2(1-\gamma_0^2)\sigma^2_{-1}\right]P_{\rm pk}(k) \\ &= [1 + \mathcal{B}(k)\mu^2]^2P_{\rm pk}(k) - \frac{k^2\mu^2}{3}f^2(1-\gamma_0^2)\sigma^2_{-1}P_{\rm pk}(k), \end{aligned} \tag{40}$$

where the linear redshift distortion parameter

$$\mathcal{B}\,(k) \equiv f\frac{b_{\rm vel}(k)}{b_{\rm pk}(k)} \tag{41}$$

is scale dependent. Recall that $b_{\rm pk}(k)$ and $b_{\rm vel}(k)$ were defined in Eqs. (7) and (21). Thus, except for the second term in the last equality, the above result exactly matches our simple estimate, Eq. (29).

Notice especially that, for linearly biased tracers, redshift space distortions are used to estimate $\beta = f/b$. The analogous quantity for peaks, $\mathcal{B}(k)$, is k dependent. We will consider the implications of this in the next section.

2. A different approximation

A more intuitive approximation to the exact result that is reached upon performing the integral in Eq. (32) can be obtained by Fourier transforming it in the first place. We write

$$\begin{aligned} \exp(-i\boldsymbol{k}\cdot\boldsymbol{s}) &= \exp(-i\boldsymbol{k}_\perp\cdot\mathbf{s}_\perp)\exp(-ik_\parallel y) \\ &\quad\times\exp[-ik_\parallel(s_\parallel - y)] \\ &= \exp(-i\boldsymbol{k}\cdot\boldsymbol{r})\exp[-ik_\parallel(s_\parallel - y)], \end{aligned} \tag{42}$$

and then rearrange the order of the integrals so that the integration over $s_\parallel - y$ is done first. Next, we use the fact that

$$\begin{aligned} \int dt e^{-t^2/2}e^{-ikt} &= e^{-k^2/2}, \\ \int dt t e^{-t^2/2}e^{-ikt} &= -ike^{-k^2/2}, \\ \int dt t^2 e^{-t^2/2}e^{-ikt} &= (1-k^2)e^{-k^2/2}, \end{aligned} \tag{43}$$

to express the result of the integral over $s_\parallel - y$ as $\exp[-f^2k_\parallel^2\sigma_{12}^2(r)/2]$ times other factors. Finally, we recast this term as $\exp[-f^2k_\parallel^2\sigma_{12}^2(\infty)/2]$ times $\exp[-f^2k_\parallel^2(\sigma_{12}^2(r) - \sigma_{12}^2(\infty))/2]$, which for small $k_\parallel$ is approximately $\exp[-f^2k_\parallel^2\sigma_{12}^2(\infty)/2]\times[1 - f^2k_\parallel^2(\sigma_{12}^2(r) - \sigma_{12}^2(\infty))/2]$. Thus, we generically expect the redshift space power spectrum to take the form $\exp[-f^2k_\parallel^2\sigma_{12}^2(\infty)/2]$ times other factors. A little algebra shows that, to leading order, these factors are precisely those given by Eq. (29), giving

$$P^s_{\rm pk}(k,\,\mu) = \exp[-f^2k^2\sigma^2_{\rm vel}\mu^2]P^{s0}_{\rm pk}(k,\,\mu). \tag{44}$$

Here, $P^{s0}_{\rm pk}$ is given by Eq. (30). We have also used the fact that, except in pathological cases, the pairwise dispersion at very large separation is simply twice the one-dimensional velocity dispersion of single particles, $\sigma_{\rm vel}$ ($= \sigma_{\rm vpk}/3$ for peaks), in units of aHf. Our notation is purposely kept general to emphasize that these results apply to any tracers of the linear density field.

Our Eq. (44) corrects a number of important errors in previous analyses [20,50]. In addition, expanding the Gaussian smoothing term shows the origin of the extra terms identified in the previous subsection (those highlighted by [10]). Finally, the form of our expression reflects the fact that the associated correlation function can be written as a convolution of the original expression $\xi^{s0}_{\rm pk}$ [which is the Fourier transform of Eq. (29)] with a Gaussian in the line-of-sight direction:

$$\xi^s_{\rm pk}(\mathbf{s}_\perp,\,s_\parallel) = \int_{-\infty}^{+\infty} ds'_\parallel G[s'_\parallel,\,\sigma_{12}(\infty)]\xi^{s0}_{\rm pk}(\mathbf{s}_\perp,\,s_\parallel + s'_\parallel). \tag{45}$$

Now the meaning of our notation should be clear: the superscript 0 refers to the limit in which the dispersion of the Gaussian smoothing term is vanishingly small. We note that this form for ξ^s was shown to be appropriate for the dark matter, without using any Fourier-space analysis [48]; our analysis demonstrates that it carries through for peaks as well. The interesting subtlety brought by density peaks is that the amplitude of the damping term $\sigma_{\rm vel}$ is related to the form of $b_{\rm vel}$ [Eq. (23)].

The functional form of our Eq. (44) has been studied previously in the context of modeling nonlinear corrections to the redshift space power of linearly biased tracers [9], although there the assumption was that $b_{\rm pk}$ is constant and $b_{\rm vel}$ is unity. We will discuss the effects of nonlinearities shortly. For completeness here, we simply borrow all of that previous analysis to show the effect that the Gaussian smoothing has on the Fourier-space multipoles. These can be written as

$$\frac{\mathcal{P}_0^s(k)}{P_{\rm pk}(k)} = A_0(\kappa) + \frac{2}{3}A_1(\kappa)\mathcal{B}(k) + \frac{1}{5}A_2(\kappa)\mathcal{B}^2(k), \quad (46)$$

$$\frac{\mathcal{P}_2^s(k)}{P_{\rm pk}(k)} = \frac{5}{2}[A_1(\kappa) - A_0(\kappa)] + \left[3A_2(\kappa) - \frac{5}{3}A_1(\kappa)\right]\mathcal{B}(k) + \left[\frac{15}{14}A_3(\kappa) - \frac{1}{2}A_2(\kappa)\right]\mathcal{B}^2(k), \quad (47)$$

and

$$\frac{\mathcal{P}_4^s(k)}{P_{\rm pk}(k)} = \frac{63}{8}A_2(\kappa) - \frac{45}{4}A_1(\kappa) + \frac{27}{8}A_0(\kappa) + \left[\frac{45}{4}A_3(\kappa) - \frac{27}{2}A_2(\kappa) + \frac{9}{4}A_1(\kappa)\right]\mathcal{B}(k) + \left[\frac{35}{8}A_4(\kappa) - \frac{135}{28}A_3(\kappa) + \frac{27}{40}A_2(\kappa)\right]\mathcal{B}^2(k), \quad (48)$$

where $\kappa \equiv fk\sigma_{\rm vel}$ and the coefficients $A_\ell(\kappa)$ are recursively defined as

$$A_0(\kappa) = \frac{\sqrt{\pi}}{2}\frac{\mathrm{erf}(\kappa)}{\kappa} \approx 1 - \frac{\kappa^2}{3},$$
$$A_\ell(\kappa) = \frac{(2\ell+1)}{2\kappa^2}(A_{\ell-1}(\kappa) - \mathrm{e}^{-\kappa^2}) \approx 1 - \frac{(2\ell+1)}{(2\ell+3)}\kappa^2. \quad (49)$$

The final approximations assume $\kappa \ll 1$. Thus, the lowest order corrections to $P^{s0}_{0,\rm pk}/P_{\rm pk}$ and $P^{s0}_{2,\rm pk}/P_{\rm pk}$ are proportional to $-k^2$.

C. Nonlinear evolution

There are four reasons why nonlinear evolution will act to change the expressions above [10,51]: one is related to the change in the bias parameters, and the three others have to do with the effect of peculiar velocities. Gravitational motions are expected to relate a scale-independent, deterministic linear bias parameter in the initial (Lagrangian) field to the evolved (Eulerian) bias according to $b_\nu^{\rm Eul} = 1 + b_\nu$ [41]. Ignoring the fact that b_ζ might also evolve, we follow common practice and assume $b_{\rm pk}^{\rm Eul} \equiv 1 + b_\nu + b_\zeta k^2$, even though we suspect that $b_{\rm pk}^{\rm Eul} \equiv b_{\rm vel} + b_\nu + b_\zeta k^2$. (This issue will be thoroughly explored in a forthcoming paper.) Note that we have omitted the smoothing window for brevity, but it effectively makes little difference at scales $k^{-1} \gg R_S$. Regarding the peak motions, we first assume that virial velocities within peaks or halos will increase $\sigma_{12}(\infty)$; these are responsible for the fingers of God [52,53] seen in galaxy surveys. Second, halo/peak motions may not closely follow linear theory, but this effect is expected to be less dramatic [54]. Finally, the real-space power spectrum will also be modified as a result of the linear theory motions [55–58].

For reasons we describe below, nonlinear effects may be approximated by setting

$$P^s_{\rm pk}(k,\mu) = P^{s0}_{\rm pk}(k,\mu)V_{\rm ql}(k,\mu^2)V_{\rm vir}(k,\mu^2), \quad (50)$$

where $P^{s0}_{\rm pk}$ is given by Eq. (30) with $b_{\rm pk}$ replaced by its nonlinear version. The filters $V_{\rm ql}$ and $V_{\rm vir}$ are supposed to reflect the quasilinear and virial corrections to the *nondamped* linear theory expression, respectively. The exact functional form of $V_{\rm ql}(k,\mu^2)$ depends upon the distribution of pairwise velocities. However, motivated by results from perturbation theory [57–59], we set

$$V_{\rm ql}(k,\mu^2) = \exp[-k^2\sigma^2_{\rm vel}(1-\mu^2) - k^2\sigma^2_{\rm vel}(1+f)^2\mu^2]. \quad (51)$$

This takes into account both the smearing of linear power caused by linear theory displacements and the damping due to the linear pairwise velocity dispersion. In principle, the reduction in linear power should be somewhat mitigated by the addition of nonlinear mode-coupling terms of the sort discussed by [56]. However, we will ignore these terms in what follows. The last multiplicative factor $V_{\rm vir}$ accounts for the damping of redshift space power due to nonlinear virial motions within halos (assumed to be uncorrelated with the large-scale flows). If the mass range is small (i.e. if the peaks cover a small range in ν), then $V_{\rm vir} = \exp(-k^2\mu^2\sigma^2_{\rm vir})$, where $\sigma_{\rm vir}$ depends on the halo or peak mass, should be a good approximation. If the mass range is broad, then an exponential distribution may be more appropriate [60], leading to $V_{\rm nl} = [1 + k^2\sigma^2_{\rm vir}\mu^2]^{-2}$. Removing fingers of God from a survey [61] is equivalent to setting $\sigma_{\rm vir} \to 0$ or $V_{\rm vir} \to 1$.

Here and henceforth, we will assume that $V_{\rm vir}$ is a Gaussian smoothing kernel. This implies that the Fourier-space multipoles $\mathcal{P}^s_\ell(k)$ are given by Eqs. (46)–(48) with

$$\kappa \equiv k\sqrt{\sigma^2_{\rm vel}f(2+f) + \sigma^2_{\rm vir}}, \quad (52)$$

upon making the replacement

$$\frac{\mathcal{P}^s_\ell(k)}{P_{\rm pk}(k)} \to \frac{\mathcal{P}^s_\ell(k)}{P_{\rm pk}(k)e^{-k^2\sigma^2_{\rm vel}}} \quad (53)$$

on the left-hand side. We will now illustrate the effect of the biasing relation Eq. (4) on the two-point correlation through a comparison between density peaks and linearly biased tracers.

D. Comparison between density peaks and linearly biased tracers

For linearly biased tracers, $b_\zeta = 0$, $\gamma_0 = 0$, and all the terms involving σ_0/σ_1 vanish. The pairwise statistics simplify to

$$v_{12}(r,\mu) = \frac{-2b_\nu\xi_1^{(-1/2)}(r)}{1+\xi_{\rm pk}(r)}L_1(\mu), \quad (54)$$

$$\sigma_{12}^2(r,\mu) = \frac{2}{3}\sigma_{-1}^2\left[1 - \frac{\xi_0^{(-1)}}{\sigma_{-1}^2} + 2\frac{\xi_2^{(-1)}}{\sigma_{-1}^2}L_2(\mu)\right]. \quad (55)$$

Setting $\beta = f/b_\nu$, we recover the linear theory prediction of [2] plus a contribution from the large-scale limit of σ_{12}^2 (which underestimates the true effect since we neglect nonlinear corrections to the velocity dispersion),

$$\xi_{\rm L}^s(s,\mu) = \frac{8}{35}f^2\xi_4^{(0)}L_4(\mu) - \left[\left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right)b_\nu^2\xi_2^{(0)} - \frac{2}{9}f^2\sigma_{-1}^2 b_\nu^2\xi_2^{(1)}\right]L_2(\mu) + \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)b_\nu^2\xi_0^{(0)} - \frac{1}{9}f^2\sigma_{-1}^2 b_\nu^2\xi_0^{(1)}. \quad (56)$$

Note that Eq. (56) implicitly assumes that the peculiar velocities of linear tracers match locally that of the matter. To account for the nonlinear evolution, we will also adopt the prescription $b_\nu \to b_\nu^{\rm Eul} = 1 + b_\nu$.

The explicit Legendre decomposition $\xi^s(s,\mu) = \sum \xi_\ell^s(s)L_\ell(\mu)$ of the redshift space correlation function can be read off from Eqs. (39) and (56). For illustration, the multipoles $\xi_\ell^s(s,\nu)$ are plotted in the left and right panels of Fig. 2 for density maxima identified at the smoothing scale $R_S = 2.5$ and 4 h^{-1}Mpc, respectively. These functions are compared to those of linearly biased tracers with the same value of $b_\nu^{\rm Eul}$, namely, $b_\nu^{\rm Eul} = 2.0$ and 3.8, respectively. It is important to note that, in the latter case, the density field is *not* smoothed (in practice, we use $R_S = 0.1$ hMpc^{-1}). Furthermore, we have also neglected the contribution $\sigma_{\rm vir}$ from virialized motions to the velocity dispersion and assumed $\sigma_{\rm vel}^2 \equiv \sigma_{\rm vpk}^2/3$ for the peaks and $\sigma_{-1}^2/3$, with $\sigma_{-1}^2 = (8.11\ h^{-1}{\rm Mpc})^2$, for the linearly biased tracers.

As recognized in [29], the nonlinear local biasing relation Eq. (4) amplifies the contrast of the (real-space) baryon acoustic signature of density maxima relative to that of linearly biased tracers. A similar enhancement is also observed in the baryonic acoustic signature of dark matter halos in very large cosmological simulations [62,63]. As can be seen in Fig. 2, this amplification is also present in redshift space. In this case, however, both the monopole and the quadrupole of $\xi_{\rm pk}^s$ are affected by the nonlinear peak biasing across the baryon acoustic oscillation (BAO) feature. At distances $s \sim 100$–110 h^{-1}Mpc, the quadrupole of the redshift space peak correlation $\xi_{\rm pk}^s$ is indeed more negative, thereby damping the correlation in the radial direction ($\mu \approx 1$) and increasing it in the perpendicular direction ($\mu \approx 0$).

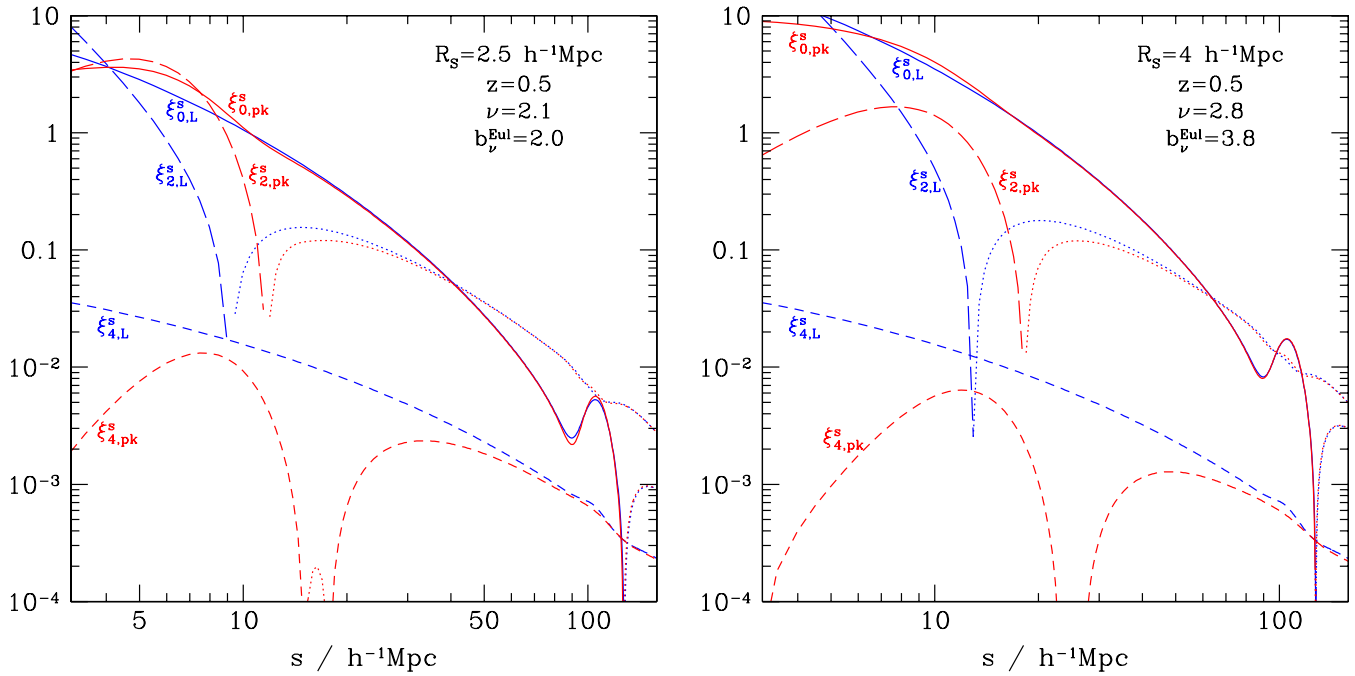

FIG. 2 (color online). A comparison between the redshift space multipoles of the correlation function of density maxima and linearly biased tracers. (Dotted lines denote negative values.) The peaks were identified in the density field when smoothed with a Gaussian filter of characteristic scale $R_S = 2.5$ (left panel) and 4 h^{-1}Mpc (right panel). This corresponds to mass scales $M_S = 1.9 \times 10^{13}$ and $7.8 \times 10^{13}{\rm M}_\odot/h$, respectively. The associated peak height and bias parameters quoted in each panel assume a redshift $z = 0.5$. The linear biased tracers are required to have the same value of $b_\nu^{\rm Eul}$. The peak biasing relation enhances the monopole and the quadrupole around the BAO scale relative to that of linearly biased tracers, and induces significant scale dependence in the hexadecapole at $s \lesssim 100$ h^{-1}Mpc.

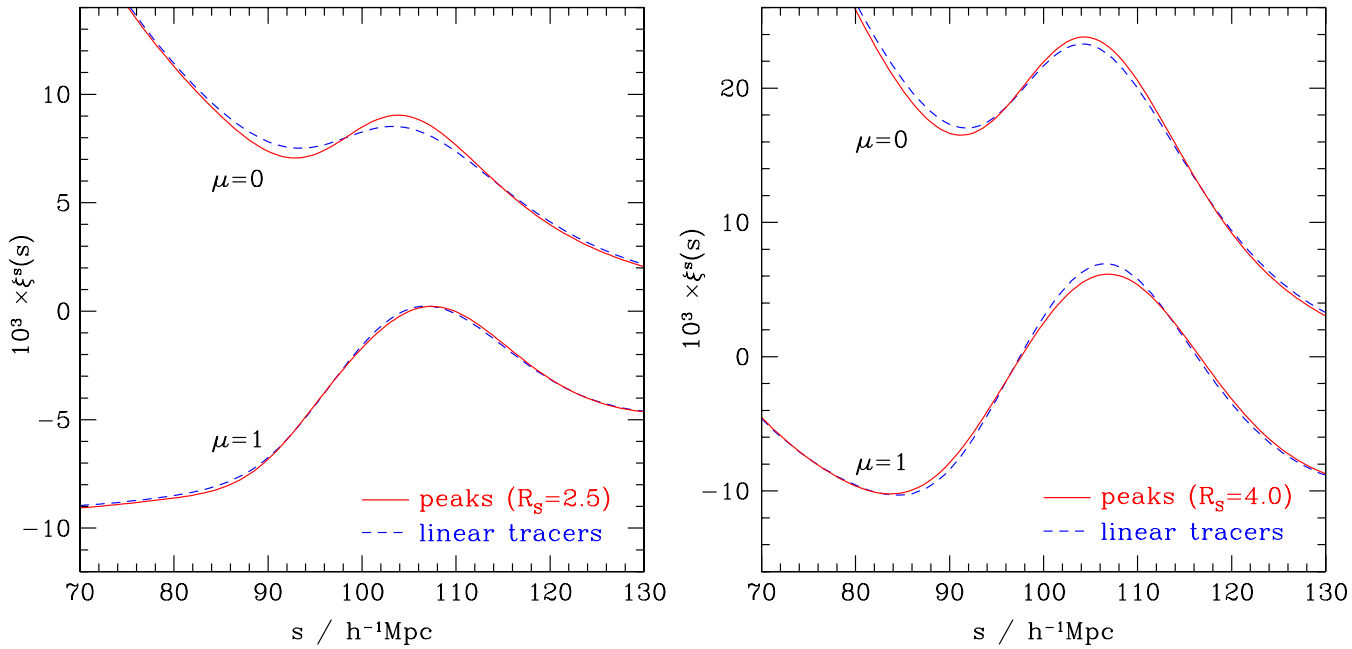

FIG. 3 (color online). Angular dependence of the redshift space correlation for the density maxima (solid curves) and linearly biased tracers (dashed curves) considered in Fig. 2. Results are shown for a separation vector oriented in the direction parallel ($\mu = 1$) and transverse ($\mu = 0$) to the line of sight. Perpendicular to the line of sight, the contrast of the acoustic peak is more pronounced in the correlation of the density peak, whereas in the radial direction, it is comparable to that of linearly biased tracers.

This is more clearly seen in Fig. 3, which compares the redshift space correlation of density peaks and linear tracers in the direction parallel and transverse to the line-of-sight axis. Relative to the baryon acoustic peak of linearly biased tracers, the BAO of density maxima is enhanced in the direction perpendicular to the line of sight while somewhat distorted in the radial direction. The physical origin of this effect presumably is peak-peak exclusion. Namely, while as discussed in [29] the spatial bias of peaks enhances the contrast of the BAO in the real-space correlation, peak-peak exclusion suppresses the infall of peak pairs onto the (slightly overdense) BAO shell at radius $s \approx 105\ h^{-1}$Mpc. In redshift space, this amounts to a reduction of the BAO contrast along the line of sight. The dispersion term $\sigma_{12}(\infty)\xi''$ further smooths the BAO and shifts the position of the local maximum in that direction, but leaves the baryon wiggle unchanged in the transverse direction. The amount of smoothing depends on the exact value of $\sigma^2_{\rm vel}$. Another striking feature of Fig. 3 is the strong suppression of the redshift space correlation and the sharpening of the acoustic peak along the line of sight due to linear coherent infall [2,8].

On scales less than the BAO ring, the contribution of the pairwise velocity dispersion increases with decreasing separation until it reverses the sign of the quadrupole at separation $\sim$10–20 h^{-1}Mpc and stretches structures along the line of sight [10]. Although the velocity dispersion of the density peaks is smaller than that of the linear biased tracers, peak-peak exclusion makes the effect stronger. On those scales, the contribution of the term $2b_\nu^{\rm Eul} b_\zeta \xi_0^{(1)} + b_\zeta^2 \xi_0^{(2)}$ becomes comparable to $(b_\nu^{\rm Eul})^2 \xi_0^{(0)}$ and steepens the profile of the angle-averaged correlation ξ_0^s. As a result, the monopole for the density peaks can be larger by a few tens of percent at separation $s \lesssim 10\ h^{-1}$Mpc relative to that of linearly biased tracers. Note that small-scale halo exclusion is not properly accounted for in our treatment since we consider $\xi_{\rm pk}$ at first order only. Nevertheless, we expect that, while in real space peak-peak exclusion leads to a deficit of pairs at distance $s \lesssim R_S$, in redshift space the suppression may be weaker because peaks tend to move toward each other.

Figure 4 displays the Fourier-space multipoles $\mathcal{P}_\ell(k)$ in units of $P_{\rm pk}(k)$ for the peaks and linear tracers considered above. To emphasize the importance of the exponential damping, results are shown with and without the smearing caused by quasilinear and virialized motions. While for the linearly biased tracers the distortion parameter $\mathcal{B}(k) = f/b_\nu^{\rm Eul}$ is a constant, for peaks $\mathcal{B}(k)$ it is k dependent and, therefore, induces a scale dependence in the multipoles even when the pairwise velocity dispersion is negligible. In this limit ($\sigma_{\rm vel} = 0$), for the linearly biased tracers the ratios $\mathcal{P}_\ell(k)/P_{\rm pk}(k)$ are constant (as in the original Kaiser formula), whereas for peaks they decay rapidly to reach $1 + 2\mathcal{B}(\infty)/3 + \mathcal{B}^2(\infty)/5$, $4\mathcal{B}(\infty)/3 + 4\mathcal{B}^2(\infty)/7$, and $8\mathcal{B}^2(\infty)/35$ when $\ell = 0$, 2, and 4, respectively. Here, $\mathcal{B}(\infty) = -(\nu/\sigma_0 - b_\nu^{\rm Eul})^{-1}$ is the value of $\mathcal{B}(k)$ in the limit $k \to \infty$ [see Eq. (31)]. The difference between peaks

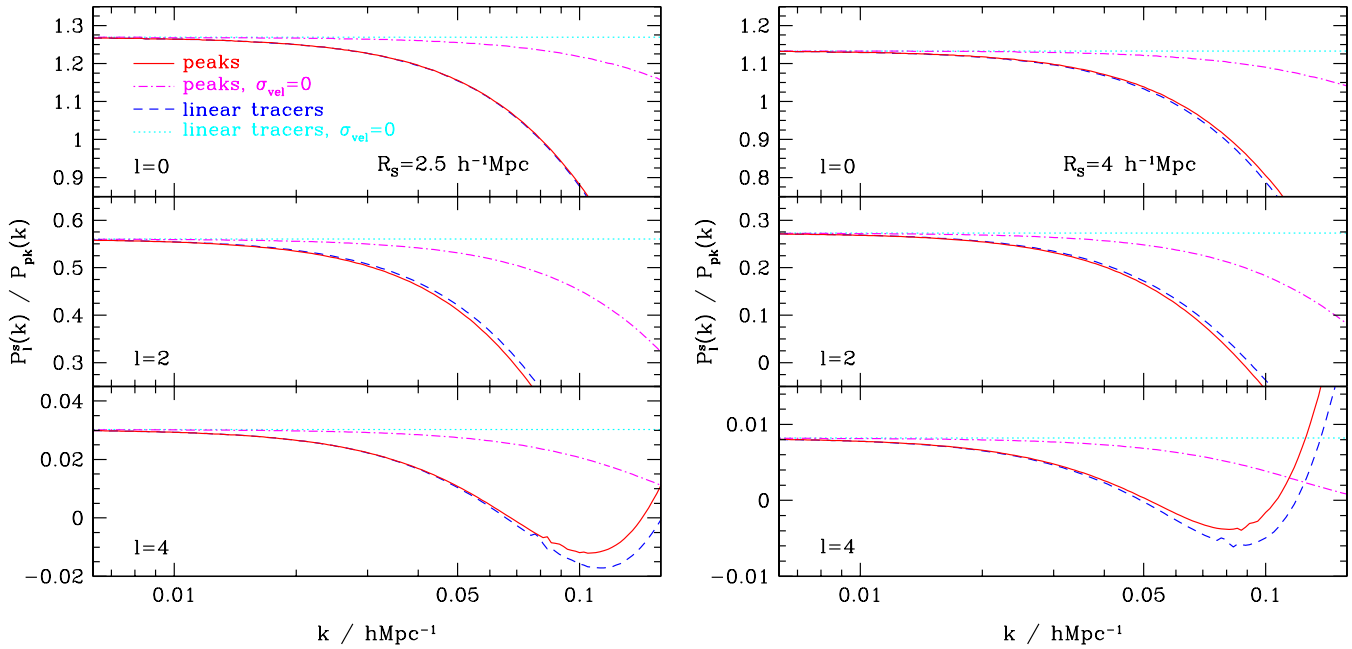

FIG. 4 (color online). Fourier-space multipoles $\mathcal{P}^s_\ell$ in units of $P_{\rm pk}(k)$ as a function of wave number for the density peaks and for the linearly biased tracers with the same value of $b^{\rm Eul}_\nu$. The dotted-dashed (peaks) and dotted (linear tracers) curves are the results without including the velocity damping kernel (assumed to be a Gaussian; see Sec. III C), while the solid (peaks) and dashed (linear tracers) curves represent the multipoles when the Gaussian dispersion is included. For density peaks, $\mathcal{P}^s_\ell(k)$ exhibit a strong k dependence even upon removal of the damping term.

and linear tracers is largest in the hexadecapole and increases with mass scale. For instance, the fractional deviation is 5% at wave numbers $k \approx 0.037\ h{\rm Mpc}^{-1}$ and $\approx 0.027\ h{\rm Mpc}^{-1}$ for the peaks identified at filtering scales $R_S = 2.5$ and 4 h^{-1}Mpc, respectively.

When the Gaussian damping term, Eq. (51), is included, the behavior of the Fourier-space multipoles of density peaks (solid curves) and linear tracers (dashed curves) becomes similar at small scales: they damp to zero like the coefficients $A_\ell(\kappa)$ defined in Eq. (49). Still, significant deviations persist on scales $k \gtrsim 0.01\ h{\rm Mpc}^{-1}$ due to the k dependence of $\mathcal{B}(k)$ and unequal velocity dispersions.

IV. COSMOLOGICAL IMPLICATIONS

A. Estimating the growth rate f

Following [2,10,64], the redshift space power spectrum of density peaks can also be written as

$$P^s_{\rm pk}(k) = [P_{\rm pk}(k) + 2\mu^2 P_{{\rm pk},\theta_{\rm pk}}(k) + \mu^4 P_{\theta_{\rm pk}}(k)]F(k,\mu^2), \tag{57}$$

where $P_{{\rm pk},\theta_{\rm pk}}(k)$ and $P_{\theta_{\rm pk}}(k)$ are the peak-velocity and velocity-velocity power spectra. Here, $P_{\rm pk}$, $P_{{\rm pk},\theta_{\rm pk}}$, and $P_{\theta_{\rm pk}}$ are *linear* spectra and $F(k,\mu^2) = V_{\rm ql}(k,\mu^2)V_{\rm vir}(k,\mu^2)$ describes both the quasilinear damping and the smearing from the small-scale velocity dispersion. As noted in [30,65–67], $P_{\theta_{\rm pk}}(k)$ is independent of the spatial bias and directly measures the matter velocity power spectrum provided there is no velocity bias. Owing to the angular dependence, a measurement of the velocity power spectrum furnishes an estimate of the linear growth rate $f\sigma_8 \propto dD/d\ln a$ that is not affected by the spatial bias.

Although the hexadecapole does not depend upon the spatial bias, it may be noisier than the monopole and dipole, so this has motivated the search for other combinations of P_0 and P_2 which may be more robust [8]. Reference [30] showed that $\mathcal{P}^s_0$ and $\mathcal{P}^s_2$ can be used to derive an estimate of the velocity power spectrum $P_{\theta_{\rm pk}}(k)$ when the density and velocity fields are perfectly correlated, namely, when the cross-correlation coefficient

$$r^2_\theta(k) = \frac{P^2_{{\rm pk},\theta_{\rm pk}}(k)}{P_{\rm pk}(k)P_{\theta_{\rm pk}}(k)} \tag{58}$$

is unity. For example, when smoothing is ignored, then

$$\hat{P} \equiv \frac{245}{48}P^s_0\left(1 + \frac{P^s_{20}}{7} - \sqrt{1 + \frac{2P^s_{20}}{7} - \frac{(P^s_{20})^2}{5}}\right) \tag{59}$$

[30] is proportional to $f^2P_\delta(k)$ when velocities are unbiased. Here, $P^s_{20} \equiv P^s_2/P^s_0$. For peaks $r^2_\theta(k) \equiv 1$ indeed holds at the lowest order, even though the linear spatial and velocity biases $b_{\rm pk}(k)$ and $b_{\rm vel}(k)$ are scale dependent. However, the velocity power spectrum now is $P_{\theta_{\rm pk}}(k) = f^2b^2_{\rm vel}(k)P_\delta(k)$, so there is an extra k dependence associ-

ated with the estimator $\hat{P}$. Since this could be interpreted erroneously as a signature of modified dark energy or gravity, any scale-dependent velocity bias (a scale-independent bias may also be present if the tracers do not move with the matter) will limit the information that can be recovered about the growth factor [30,65,66]. For peaks, the velocity bias is $b_{\rm vel}(k) \leq 1$, and it converges towards unity (i.e. unbiased velocities) in the limit $k \to 0$. At the first order, the deviation from unity is controlled by σ_0/σ_1 [Eq. (21)] so that, at fixed wave number, $b_{\rm vel}(k)$ decreases with increasing mass scale (see Fig. 1). For $M_S = 1.9$ and $7.8 \times 10^{13} {\rm M}_\odot/h$ considered here, $P_{\theta_{\rm pk}}(k)$ is suppressed by $\approx 5\%$ and 9% at wave number $k = 0.05\ h{\rm Mpc}^{-1}$, respectively. The predicted k dependence is smaller than current constraints on the growth rate [65,68]. Furthermore, numerical simulations to date show that the power spectrum of dark matter halo velocities is consistent with $f^2 P_\delta(k)$ within 10% at wave number $k \lesssim 0.1\ h{\rm Mpc}^{-1}$ [30]. Nevertheless, since forthcoming large-scale galaxy surveys will dramatically improve constraints on the growth factor (down to the percent level), it is interesting to assess the extent to which a k-dependent bias would degrade the constraint on the growth rate.

B. Error forecast with a k-dependent velocity bias

To this purpose, we use the Fisher based formalism developed in [66]. For Gaussian random fields, the Fisher matrix for a set of parameters $\{p_i\}$ is [69,70]

$$F_{ij} = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \left(\frac{\partial \ln P}{\partial \ln p_i}\right)\left(\frac{\partial \ln P}{\partial \ln p_j}\right) V_{\rm eff}(\boldsymbol{k}), \qquad (60)$$

where P is the power spectrum and individual wave-mode contributions are weighted by the effective volume [71]

$$V_{\rm eff}(\boldsymbol{k}) = V\left(\frac{\bar{n}P}{1+\bar{n}P}\right)^2 \qquad (61)$$

which depends upon the surveyed volume V and the number density $\bar{n}$ of the tracers (assumed to be homogeneously distributed). To illustrate, we assume the linear, plane-parallel approximation and consider the model

$$P^s(k,\mu) = [b_{\rm pk}(k) + f b_{\rm vel}(k)\mu^2]^2 P_\delta(k), \qquad (62)$$

where $b_{\rm pk}(k) \equiv b_\nu + b_\zeta k^2$ (we drop the superscript Eul for brevity) and $b_{\rm vel}(k) \equiv 1 - R_{\rm vel}^2 k^2$ (for some $R_{\rm vel}$) are motivated by the functional form of the spatial and velocity biases of density peaks (c.f. Secs. II C and II F).

In what follows, we fix the shape and amplitude of the matter power spectrum (i.e. the fractional error on $f\sigma_8$ is equal to that on f) and consider the four-parameter set $\{b_\nu, b_\zeta, R_{\rm vel}, f\}$. Our fiducial model has $(b_\nu, b_\zeta, R_{\rm vel}, f) = (1, 16, 3, 0.46)$. The values of b_ζ and $R_{\rm vel}$ closely correspond to those of density peaks identified at the mass scale $1.9 \times 10^{13} {\rm M}_\odot/h$. Derivatives of the logarithm of the power with respect to the parameters are computed easily:

$$\frac{\partial \ln P}{\partial b_\nu} = \frac{2}{(b_{\rm pk} + f b_{\rm vel}\mu^2)}, \qquad \frac{\partial \ln P}{\partial b_\zeta} = \frac{2k^2}{(b_{\rm pk} + f b_{\rm vel}\mu^2)},$$
$$\frac{\partial \ln P}{\partial R_{\rm vel}} = \frac{-4 f R_{\rm vel}\mu^2 k^2}{(b_{\rm pk} + f b_{\rm vel}\mu^2)}, \qquad \frac{\partial \ln P}{\partial f} = \frac{2 b_{\rm vel}\mu^2}{(b_{\rm pk} + f b_{\rm vel}\mu^2)}. \qquad (63)$$

We integrate over wave numbers from $k_{\rm min} \sim \pi/V^{1/3}$, where V is the volume of the survey, up to a maximum wave number $k = 0.1\ h{\rm Mpc}^{-1}$, above which nonlinear effects are expected to become important [66].

In order to illustrate the effect of including a k-dependent velocity bias into the analysis, we initially set $b_{\rm vel} \equiv 1$ (i.e. ignore $R_{\rm vel}$) and compute the Fisher matrix solely for b_ν, b_ζ, and f. For a survey of volume $V = 10\ h^{-3}{\rm Gpc}^3$ at redshift $z = 0$, we find a fractional marginalized error of $\delta f/f = 1.6\%$ in the limit $\bar{n}P \gg 1$ of negligible shot noise. (In practice, a suitable weighting of galaxies may help approaching this limit [72].) For number densities $\bar{n} = 5 \times 10^{-4}$ and $10^{-4}\ h^3{\rm Mpc}^{-3}$, the constraint weakens to 1.9% and 2.9%, respectively. (These values are consistent with those of [66].) Unsurprisingly, b_ν and b_ζ are strongly anticorrelated (the correlation coefficient is $r \approx -0.8$) because an increase in b_ν can be mostly compensated by a decrease in b_ζ. However, while the correlation between b_ν and f is moderate ($r \lesssim -0.5$), b_ζ and f are weakly degenerate ($r \lesssim -0.05$). In other words, including a k-dependent bias component $b_\zeta k^2$ has little effect on the uncertainty on f. Extending $k_{\rm max}$ beyond $0.1\ h{\rm Mpc}^{-1}$ (where the shot noise again becomes important) can reduce the uncertainty on f (because the fractional error scales as $k_{\rm max}^{-3/2}$), but this is at the price of having to model the smearing due to quasilinear motions and small-scale velocities.

Introducing the parameter $R_{\rm vel}$ substantially increases the uncertainty on f. For the volume V and the average number densities $\bar{n}$ considered above, the fractional marginalized uncertainty on the growth rate becomes $\delta f/f =$ 4%, 4.4%, and 6%, respectively. This can be traced to the strong correlation ($r \approx 0.9$) between $R_{\rm vel}$ and f. The error degradation reflects the fact that we are adding more freedom to the model. It does not depend upon the exact value of $R_{\rm vel}$.

Are the constraints on f obtained using the multitracer method proposed in [73] affected in a similar way? Reference [74] pointed out that several populations of differently biased tracers can achieve a much better determination of the growth rate than a single sample of objects. When power spectra are measured, calculating the Fisher matrix for multiple tracers requires summing over the distinct components of the inverse covariance matrix C_{AB}^{-1}, where A, B label a different pair of tracer populations,

$$F_{ij} = V \sum_{A,B} \int \frac{d^3k}{(2\pi)^3} \left(\frac{\partial P_A}{\partial p_i}\right) C_{AB}^{-1} \left(\frac{\partial P_B}{\partial p_j}\right). \qquad (64)$$

The calculation of the covariance matrix is straightforward if one assumes that the noise term can be treated as an uncorrelated normal variate [66,75]. For completeness, the diagonal and off-diagonal components of the covariance matrix are [66]

$$\langle C_{aaaa}\rangle = 2P_{aa}^2N_a^2, \qquad \langle C_{abab}\rangle = P_{ab}^2 + P_{aa}P_{bb}N_aN_b \tag{65}$$

and

$$\begin{aligned}\langle C_{abcd}\rangle &= 2P_{ab}P_{cd}, \qquad \langle C_{aabc}\rangle = 2P_{ab}P_{ac},\\ \langle C_{abac}\rangle &= P_{ab}P_{ac} + P_{aa}P_{bc}N_a,\\ \langle C_{aaab}\rangle &= P_{ab}P_{aa}N_a, \qquad \langle C_{aabb}\rangle = 2P_{ab}^2,\end{aligned} \tag{66}$$

where P_{ij} is an autopower or cross-power spectrum and $N_a \equiv [1 + 1/(\bar{n}P_{aa})]$. We will consider the simplest case of two types of tracers, since the gains saturate rapidly as the number of samples increases [66]. In this case there are three distinct measured power spectra,

$$P_{ab} = (b_{\rm pk}^{(a)} + fb_{\rm vel}^{(a)}\mu^2)(b_{\rm pk}^{(b)} + fb_{\rm vel}^{(b)}\mu^2)P_\delta(k), \tag{67}$$

where a, $b = 1$, 2 and the bias factors are $b_{\rm pk}^{(a)} = b_\nu^{(a)} + b_\zeta^{(a)}k^2$, $b_{\rm vel}^{(a)} = 1 - (R_{\rm vel}^{(a)})^2k^2$. The power spectra are thus described by seven parameters, and their derivatives are

$$\frac{\partial P_{ab}}{\partial b_\nu^{(c)}} = [(b_{\rm pk}^{(b)} + fb_{\rm vel}^{(b)}\mu^2)\delta_{ac}^K + (b_{\rm pk}^{(a)} + fb_{\rm vel}^{(a)}\mu^2)\delta_{bc}^K]P_\delta(k), \tag{68}$$

$$\frac{\partial P_{ab}}{\partial b_\zeta^{(c)}} = k^2[(b_{\rm pk}^{(b)} + fb_{\rm vel}^{(b)}\mu^2)\delta_{ac}^K + (b_{\rm pk}^{(a)} + fb_{\rm vel}^{(a)}\mu^2)\delta_{bc}^K] \times P_\delta(k), \tag{69}$$

$$\frac{\partial P_{ab}}{\partial R_{\rm vel}^{(c)}} = -2fk^2\mu^2[(b_{\rm pk}^{(b)} + fb_{\rm vel}^{(b)}\mu^2)R_{\rm vel}^{(a)}\delta_{ac}^K + (b_{\rm pk}^{(a)} + fb_{\rm vel}^{(a)}\mu^2)R_{\rm vel}^{(b)}\delta_{bc}^K]P_\delta(k), \tag{70}$$

$$\frac{\partial P_{ab}}{\partial f} = \mu^2[(b_{\rm pk}^{(b)} + fb_{\rm vel}^{(b)}\mu^2)b_{\rm vel}^{(a)} + (b_{\rm pk}^{(a)} + fb_{\rm vel}^{(a)}\mu^2)b_{\rm vel}^{(b)}]P_\delta(k). \tag{71}$$

Here, δ_{ab}^K is the Kronecker delta.

Figure 5 shows the fractional marginalized error on f obtained by combining two different biased samples of the same survey volume. The constraints are shown as a function of the abundance of the second tracers with and without including a k-dependent velocity bias (solid and dotted curves, respectively). We set $b_\nu^{(2)} = 1.4$, 2, and 4 to facilitate the comparison with Fig. 3 of [66]. Although we have assumed the fiducial values $R_{\rm vel}^{(1)} = R_{\rm vel}^{(2)} = 3\ h^{-1}$Mpc for simplicity, we may expect from the analysis done in the previous section that $R_{\rm vel}$ has some mass or bias dependence. E.g., $R_{\rm vel}^{(2)} > R_{\rm vel}^{(1)}$ when $b_\nu^{(2)} \gg b_\nu^{(1)}$. The marginalized error on f is, however, weakly dependent on the fiducial value of $R_{\rm vel}^{(a)}$. Note the considerable improvement in the constraint on f [in agreement with the findings of [66,74]]. The smallest error is achieved with a large number density $\bar{n}_2$ and large relative bias $b_\nu^{(2)}/b_\nu^{(1)}$. (We have used values of $b_\nu^{(2)}$ to simplify comparison with [66].) However, including a k-dependent velocity bias degrades the uncertainty on the growth rate roughly by a factor of 2 when $\bar{n}_2 \gtrsim 10^{-2}\ h^3$Mpc^{-3}, like in the single tracer case. The error degradation becomes increasingly severe as one goes to lower number densities.

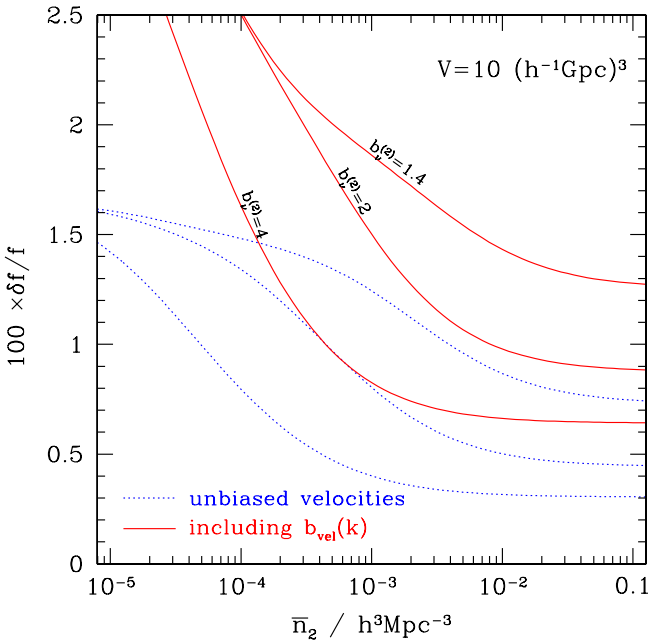

FIG. 5 (color online). The fractional marginalized error $\delta f/f$ for a survey volume $V = 10\ h^{-3}$Gpc^3 at $z = 0$ obtained with two tracer populations : a high density, unbiased sample with $\bar{n}_1 = 10^{-2}\ h^3$Mpc^{-3} and $b_\nu^{(1)} = 1$ and a second population with varying number density $\bar{n}_2$ and bias $b_\nu^{(2)}$. We choose $b_\nu^{(2)} = 1.4$, 2, and 4 (curves from top to bottom). The constraints are shown as a function of $\bar{n}_2$, assuming $b_{\rm vel}^{(1)} = b_{\rm vel}^{(2)} \equiv 1$ (dotted curves) and a k-dependent velocity bias with $R_{\rm vel}^{(1)} = R_{\rm vel}^{(2)} = 3\ h^{-1}$Mpc (solid curves).

Although these constraints are only indicative (we have ignored the influence of cosmological parameters on the constraint [76]), our analysis demonstrates that allowing for a k-dependent velocity bias (with the specific functional form predicted by the peak model) has a large impact on the determination of the growth factor and, therefore, may possibly hamper our ability to distinguish between different dark energy or gravity scenarios [77,78]. Therefore, despite the lack of current evidence for a k-dependent velocity bias [30], it seems prudent to study this possibility further with large cosmological simulations. We hope the

peak model can serve as a useful baseline with which to compare the simulations.

V. STOCHASTICITY

A. Cross-correlation coefficient

The biasing equation (4) derived from the large-scale properties of peak correlation functions is a mean bias relation that does not contain any information about stochasticity. Therefore, it is unsurprisingly deterministic like the local bias model considered by [38], the main difference residing in the fact that peak biasing involves derivatives of the density field. Still, because of the discrete nature of density peaks, one can expect that the peak overdensity $\delta n_{\rm pk}$ at location $\boldsymbol{x}$ is generally a random function of the underlying matter density (and its derivatives) in some neighborhood of that point. We note that stochastic models of the form $\delta n_{\rm pk}(\boldsymbol{x}) = X[\delta_S(\boldsymbol{x})]$ have been studied in [79–81], for instance.

Computing the probability of X given δ, etc., is beyond the scope of this paper (because it requires the full hierarchy of correlation functions). Still, it is instructive to compute the cross-correlation coefficient to gain further understanding of the peak biasing model. The cross-correlation coefficient is defined as

$$r_c^2(k) = \frac{P_{{\rm pk},\delta}^2(k)}{P_{\rm pk}(k)P_\delta(k)}, \qquad r_\xi^2(r) = \frac{\xi_{{\rm pk},\delta}^2(r)}{\xi_{\rm pk}(r)\xi_\delta(r)} \tag{72}$$

in Fourier and configuration space, respectively. Ignoring the damping term, we find $r_c(k) = 1$ for peaks even though the ratio $P_{{\rm pk},\delta}/P_\delta$ depends on k. Thus, a k-dependent bias at the linear order does not yield stochasticity in Fourier space. On the other hand,

$$r_\xi^2(r) = \frac{(b_\nu + b_\zeta \xi_0^{(1)}/\xi_0^{(0)})^2}{b_\nu^2 + 2b_\nu b_\zeta \xi_0^{(1)}/\xi_0^{(0)} + b_\zeta^2 \xi_0^{(2)}/\xi_0^{(0)}}. \tag{73}$$

Therefore, although the bias is deterministic in Fourier space, it is generally stochastic and scale dependent in configuration space. However, when $\nu \gg 1$ then $b_\zeta \to 0$, so $r_\xi \to 1$. Namely, in the high peak limit, the bias becomes linear and deterministic in both Fourier and configuration space.

The real-space cross-correlation coefficient is shown in Fig. 6 as a function of comoving separation for the peaks identified at the smoothing radii $R_S = 2.5$ and 4 h^{-1}Mpc. r_ξ is very close to unity at all separations larger than a few smoothing radii, except around the baryonic bump and the zero-crossing of the correlation function where it can be noticeably larger than unity. These findings seem to run contrary to the common knowledge that $|r_\xi| \leq 1$. However, at separation $r \gtrsim 120\ h^{-1}$Mpc, the fact that the zero-crossings of $\xi_0^{(n)}$ do not generally coincide unavoidably implies $|r_\xi| > 1$, at least over some range of scales. Furthermore, at distance $r \sim 90\ h^{-1}$Mpc, the large

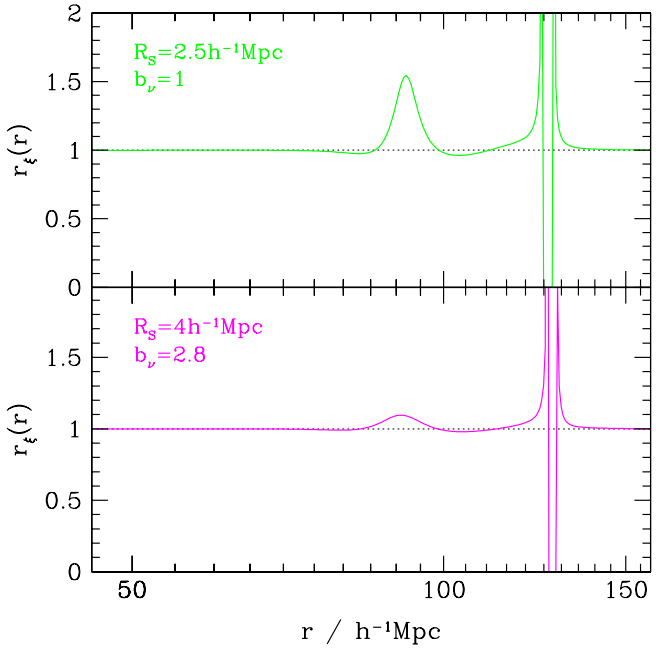

FIG. 6 (color online). The cross-correlation coefficient $r_\xi(r)$ for density peaks identified at the smoothing scales $R_S = 2.5$ and 4 h^{-1}Mpc (upper and lower panels, respectively). There is significant stochasticity only at the zero-crossings of the auto-correlation and cross-correlation functions and across the BAO, where the derivatives of the density correlation, $\xi_0^{(1)}$ and $\xi_0^{(2)}$, are not negligible [29].

values of r_ξ are most plausibly traced to the baryon acoustic feature, which induces large oscillations in $\xi_0^{(1)}$ and $\xi_0^{(2)}$ across the BAO scale $\approx 105\ h^{-1}$Mpc (see Fig. 1 of [29]). At the level of a bias relation $\delta n_{\rm pk} = X[\delta_S, \nabla^2\delta_S, \cdots]$, this suggests that the scatter is strongly sensitive to $\nabla^2\delta_S$.

Figure 7 explores the behavior of the cross-correlation coefficient when the underlying power spectrum is a featureless power law spectrum, $P_\delta(k) \propto k^{n_s}$. Results are presented as a function of the spectral index n_s for a single value of the separation, $r = 100\ h^{-1}$Mpc. At a fixed value of n_s, the stochasticity is unsurprisingly larger for the relatively sparser peaks identified at the scale $R_S = 4\ h^{-1}$Mpc. Most importantly, the amount of stochasticity depends sensitively upon the shape of the matter power spectrum. Overall, r_ξ decreases with larger values of the power-law exponent n_s because the stochasticity rises as the relative amount of small-scale power increases. As can also be seen, r_ξ is slightly larger than unity in the range $-3 < n_s < -2$ and at the points of discontinuity $n_s = 0, 2$ (which are marked as empty symbols). Although the effect is admittedly small and localized in n_s, this demonstrates that the cross-correlation coefficient can exceed unity also when the power spectrum is scale-free. In Appendix B, we investigate the discontinuities in more detail and provide quantitative estimates of the large-scale behavior of the cross-correlation coefficient for a few values of n_s.

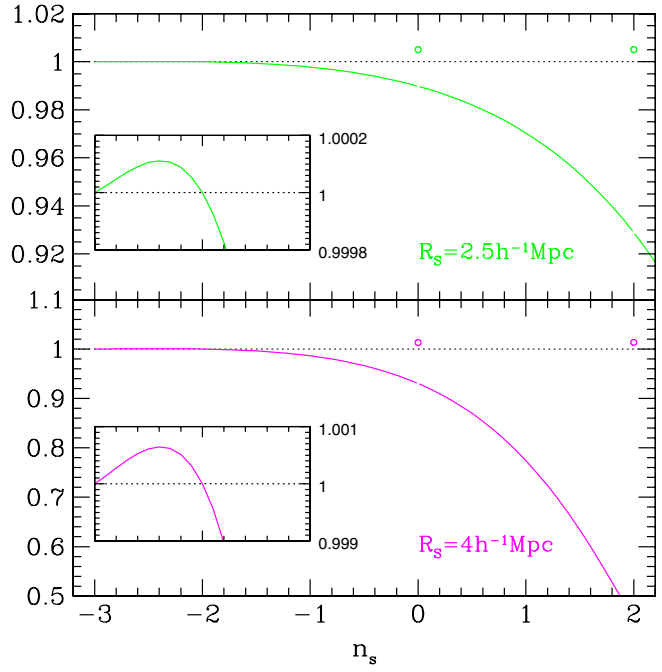

FIG. 7 (color online). Cross-correlation coefficient $r_\xi(r)$ for power-law power spectra $P_\delta(k) \propto k^{n_s}$ as a function of the spectral index n_s. Results are shown at a single separation $r = 100\ h^{-1}$Mpc, for density peaks identified at the smoothing scales $R_S = 2.5$ (top panel) and 4 h^{-1}Mpc (bottom panel). The insert is an enlarged view of r_ξ in the range $-3 < n_s < -1$. Notice the discontinuities at $n_s = 0$ and 2, at which the cross-correlation coefficient is slightly larger than unity.

B. Evolution of stochastic bias

As discussed in Sec. III C, gravitational evolution maps a scale-independent, deterministic linear bias factor in the initial conditions onto a similar quantity in the evolved distribution [41]. The scaling $b_\nu^{\rm Eul} = 1 + b_\nu$ also works for a k-dependent deterministic bias. More precisely,

$$b_{\rm pk}^{\rm Eul}(k, z) - 1 = b_{\rm pk}(k, z) = \frac{b_{\rm pk}^{\rm Eul}(k, z_0) - 1}{D(z)/D(z_0)} = \frac{b_{\rm pk}(k, z_0)}{D(z)/D(z_0)}, \quad (74)$$

where $D(z)$ is the linear theory growth factor [82,83]. This is easily understood if one recognizes that a peak of height $b_z\delta_z$, where δ_z is the linearly evolved field at z, could also have been written as having height $b_0\delta_0$, where δ_0 is the field evolved to z_0. The relation $b_z\delta_z = b_0\delta_0$ implies $b_0 = b_z(\delta_z/\delta_0) = b_z D(z)/D(z_0)$, from which the above expression is derived. Alternatively, notice that $b_0 \propto 1/\sigma_0(z_0) \propto D(z)/D(z_0)/\sigma_0(z) \propto b_z D(z)/D(z_0)$, so the factor $D(z)/D(z_0)$ is simply converting from one choice of fiducial time to another.

In configuration space, it has been argued that for linear stochastic bias,

$$b_\xi^{\rm Eul}(z) r_\xi(z) - 1 = \frac{b_\xi^{\rm Eul}(z_0) r_\xi(z_0) - 1}{D(z)/D(z_0)} \quad (75)$$

(e.g. [84]). The corresponding expression for the evolution of $b_\xi^{\rm Eul}(r, z)$ itself is

$$b_\xi^{\rm Eul}(z)^2 D^2(z) = b_\xi^{\rm Eul}(z_0)^2 D^2(z_0) + [D(z_0) - D(z)]^2 - 2[D(z_0) - D(z)] D(z_0) b_\xi^{\rm Eul}(z_0) r_\xi(z_0). \quad (76)$$

This is a good model of $b_\xi^{\rm Eul} r_\xi$ for density peaks, provided we interpret the denominator of Eq. (73) as b_ξ^2. Moreover, the numerator of this equation is similar to (the square of) a Eulerian bias factor minus 1: $([1 + b_\nu + b_\zeta \xi_0^{(1)}/\xi_0^{(0)}] - 1)^2$. This quantity clearly scales with the growth factor like its Fourier-space analog. Hence, the real-space evolution of the stochastic bias of density peaks is simple in spite of the additional scale dependence. Notice that both $b_{\rm pk}^{\rm Eul}(k)$ and $b_\xi^{\rm Eul} r_\xi$ tend to unity at late times (even though they might not effectively reach this limit because the growth factors freeze out in ΛCDM-like models). In fact b_ξ does as well, but this is not as easy to see from our expressions.

C. Connection to previous work

We mentioned earlier that peak bias and its evolution have been studied in simulations by [26,44]. These authors found that the peak-background split argument [Eq. (15)] provides a good description of the large-scale bias of the peaks extracted from their simulations. They also found that Eq. (74) is in reasonable agreement with the evolution of this large-scale bias. However, on smaller scales, the real-space bias was found to be scale dependent and stochastic [44], two features which a peak-background split based analysis does not model. Nevertheless, Eqs. (74) and (75) were found to provide a good description of the evolution. The above analysis shows why. It would be interesting to see if our approach correctly predicts the scale dependence of the bias. We defer this issue to a future work.

VI. DISCUSSION AND CONCLUSIONS

We have presented an extensive analysis of the Gaussian peak model. Density peaks are biased tracers of the underlying matter density field—on large scales this bias is scale independent—and we studied the limit in which this bias just starts to exhibit a scale dependence, both in the spatial and in the velocity fields [Eq. (5)]. In almost all cases we presented a relatively straightforward analysis in the main text, which was sometimes backed up with detailed calculations in the appendixes.

We showed that, in the large-scale, scale-independent limit, our expressions reduce to those of the peak-background split (Sec. II D), but, in general, the scale dependence in our model implies a much richer structure.

For example, even though the peaks flow with the underlying field, their velocities appear to be biased (Sec. II F). In addition, we showed how this k-dependent bias propagates into the analysis of redshift space distortions (Sec. III). We derived an exact formula for the linear theory redshift space correlation function [Eq. (32)], and then argued that it should be well approximated by a simpler expression which has considerable intuitive appeal [Eq. (44)].

Our formula shows that (i) redshift space distortions of peaks can be modeled using the same formula (and physics) as in Ref. [2], except that various terms now become k dependent, and (ii) there is in addition a Gaussian smoothing term, which reflects the dispersion of particle velocities in linear theory. Thus, our formula has the same form as the phenomenological relation that is commonly used to model nonlinear effects, and which has been shown to provide increased accuracy when comparing theory with simulations. However, here we demonstrated explicitly that this functional form is also part and parcel of linear theory. Kaiser's relation [2] assumes that the smoothing term is unity (the $k \ll 1$ limit) whereas Scoccimarro's formula [10], which is derived from the full Gaussian random field expression, is equivalent to expanding the Gaussian smoothing term and retaining only the monopole and quadrupole. Our result implies that linear theory can account for some of the effects that such a phenomenological model would otherwise ascribe to nonlinear evolution.

We provided a crude treatment of nonlinear effects (Sec. III C), which, though not properly accounting for the nonlinear evolution of the matter density and velocity fields [85] nor for the mode-coupling contribution induced by nonlinear gravitational clustering [86], illustrates how the new smoothing term, and the k dependence of the (spatial and velocity) bias factors, impacts cosmological constraints from galaxy redshift surveys (Sec. IV). Our analysis showed that allowing for a k-dependent velocity bias degrades constraints on the growth rate f by at least a factor of 2. Large cosmological simulations will be needed to ascertain whether dark matter halos hosting the surveyed galaxies also exhibit a k-dependent velocity bias. If they do, then improving the determination of f will lie in our ability to model this bias.

We also used the peaks bias model to investigate the stochasticity of the bias and its evolution (Sec. V). We provided explicit expressions for the evolution of the scale-dependent peak bias and stochasticity, and argued that they helped us to understand recent measurements of these quantities in numerical simulations.

As regards the evolution of peak bias, it is interesting to consider the peak model in light of recent work on possible modifications of gravity. In standard gravity, the linear theory growth factor is scale independent. Therefore, a peak retains its height when the initial density field is linearly evolved. However, in modified gravity models, the linear growth factor is k dependent. As a result, peaks in the initial field may not correspond to peaks in the linearly evolved field because the shape of the power spectra for the two Gaussian fields is different. This can also be seen directly by studying the (linear theory) motions of peaks. In such models, the bias of objects which coherently flow with the matter evolves just as it does in standard gravity [83]. Thus, whereas objects initially placed at maxima of the density field will still move in accordance with the (modified) matter flows, gravitational motions will bring them to positions which are no longer local maxima.

The question then arises as to whether it is the initial peaks or those in the evolved field which bear a closer resemblance to the galaxies and clusters we see today. Presumably it is the peaks which have managed to survive from the initial time to the present which are the ones of most interest—the ones which are transients are probably less interesting. In theories with k-dependent linear growth, only the peaks with exactly the right large-scale surroundings (determined by the k dependence of linear theory) will survive at later times; this raises the possibility that the correlation between galaxy clusters and their environments can constrain theories of large-scale modifications to gravity. We have not pursued this further, but note that this is consistent with recent analyses of dark matter halos [87].

To conclude, it is worth mentioning that halo-based approaches, which provide a reasonably good description of the weakly nonlinear clustering of simulated dark matter halos and galaxies [88], are now commonly used to extract cosmological information from redshift surveys (see [89] for a review). Although the dark matter halos are the local density maxima of the evolved matter distribution, there is no easy correspondence with the initial density maxima. This is the reason why the peak model has somewhat fallen out of favor. We believe our work has shown that many insights can be gained from a study of density peaks [24,29], particularly with regard to a number of effects—including scale dependence and stochasticity of the spatial and velocity biases—which matter in the age of precision cosmology.

ACKNOWLEDGMENTS

We thank the organizers of the Benasque cosmology meeting in August 2008, where this work was initiated. We are grateful to Eiichiro Komatsu and Roman Scoccimarro for their careful reading and comments on the manuscript, and acknowledge useful discussions with Adam Amara, Stéphane Colombi, Roman Scoccimarro, Uroš Seljak, and Robert Smith. R. K. S. thanks J. Bagla and the Harish Chandra Research Institute for support during the later stages of this work, and S. Mei and J. Bartlett, as well as the APC, Paris 7 Diderot, where he was a Visiting Professor when this work was written up. V. D. is supported by the Swiss National Foundation under

Contract No. 200021-116696/1. R. K. S. is supported by NSF Grant No. AST-0908241.

APPENDIX A: CHECKING THE CONSISTENCY OF THE PEAK BIASING RELATION

In this appendix, we sketch the derivation of the cross-correlation between peaks and the underlying density field, $\xi_{\mathrm{pk},\delta}(r)$, and the averaged peak pairwise velocity, $v_{12}(r,\mu)$, which have not been derived previously. We compute both quantities using the peak constraint and demonstrate that the results are consistent with those inferred (after a trivial calculation) from the peak biasing relation (4). See [29] for complementary details about the calculation.

1. Cross correlation between peaks and the density field

Let $\eta_i = \partial_i\delta(\boldsymbol{x})/\sigma_1$ and $\zeta_{ij} = \partial_i\partial_j\delta(\boldsymbol{x})/\sigma_2$ be the normalized first and second derivatives of the density field. Furthermore, let Λ be the diagonal matrix of entries $\mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)$, where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ is the nonincreasing sequence of eigenvalues of the symmetric matrix $-\zeta$. The cross correlation $\xi_{\mathrm{pk},\delta}(r) = \langle\delta n_{\mathrm{pk}}(\boldsymbol{x}_1)\nu(\boldsymbol{x}_2)\rangle$ ($r = |\boldsymbol{x}_2 - \boldsymbol{x}_1|$) follows from the Kac-Rice formula [90],

$$\xi_{\mathrm{pk},\delta}(r) = \frac{3^{3/2}\sigma_0}{\bar{n}_{\mathrm{pk}}(\nu)R_1^3}\langle|\det\zeta(\boldsymbol{x}_1)|\delta^3[\eta(\boldsymbol{x}_1)]\theta(\lambda_3)\nu(\boldsymbol{x}_2)\rangle, \tag{A1}$$

where $\bar{n}_{\mathrm{pk}}$ is the differential number density of peaks of height ν as given in Eq. (9), R_1 is the typical radius of peaks, and λ_3 is the smallest eigenvalue of $-\zeta$ at $\boldsymbol{x}_1$. We have omitted the ν dependence for brevity. The Heaviside step function arises because we are interested in counting maxima solely, for which ζ_{ij} is negative definite at the extrema position. Unlike [20] which expresses the covariance matrix in a coordinate system where the two density maxima lie on the z axis, we write the expectation value in the right-hand side of Eq. (A1) as an integral over the angular average, joint probability distribution function $P(\boldsymbol{y},\nu_2;r)$. Here, $\boldsymbol{y}^\top = (\eta_i,\nu,\zeta_A)$ is a ten-dimensional vector whose components ζ_A, $A = 1,\ldots,6$ symbolize the entries $ij = 11, 22, 33, 12, 13, 23$ of ζ_{ij}.

To evaluate the 11-dimensional covariance matrix $\mathrm{C}(r)$, it is convenient to split the degrees of freedom associated with the tensor ζ into the scalar $u = -\mathrm{tr}\zeta = \sum_i\lambda_i$ (so that u is positive when $\lambda_3 > 0$) and the traceless matrix $\tilde{\zeta} = \zeta - 1/3(\mathrm{tr}\zeta)\mathrm{I}$, where I is the 3×3 identity matrix. Let $\tilde{\zeta}_A$ designate the 5 degrees of freedom of $\tilde{\zeta}$. We see that $\mathrm{C}(r) = (1/4\pi)\int d\Omega_{\hat{\boldsymbol{r}}}\mathrm{C}(\boldsymbol{r})$ has the block diagonal decomposition $\mathrm{C} = \mathrm{diag}(\mathrm{C}_1,\mathrm{C}_2,\mathrm{C}_3)$. Consequently, $P(\boldsymbol{y}_1,\nu_2;r)$ can be expressed as a product of three joint probability distributions

$$P(\boldsymbol{y},\nu_2;r) = P(\nu_1,u_1,\nu_2;r)P(\eta_1)P(\tilde{\zeta}_1) \tag{A2}$$

where, for shorthand convenience, subscripts denote the variables evaluated at different Lagrangian positions. The one-point probability distributions $P(\eta_1)$ and $P(\tilde{\zeta}_1)$ are

$$P(\eta_1) = \left(\frac{3}{2\pi}\right)^{3/2}\exp\left(-\frac{3\eta_1^2}{2}\right),$$
$$P(\tilde{\zeta}_1) = \frac{15^3}{(2\pi)^{5/2}2\sqrt{5}}\exp\left[-\frac{15}{4}\,\mathrm{tr}(\tilde{\zeta}_1^2)\right], \tag{A3}$$

while the joint density $P(\nu_1,u_1,\nu_2;r)$ has a covariance matrix

$$\mathrm{C}_1 = \begin{pmatrix} 1 & \xi_0^{(0)}/\sigma_0^2 & \gamma_1\xi_0^{(1)}/\sigma_1^2 \\ \xi_0^{(0)}/\sigma_0^2 & 1 & \gamma_1 \\ \gamma_1\xi_0^{(1)}/\sigma_1^2 & \gamma_1 & 1 \end{pmatrix}. \tag{A4}$$

Upon inversion of C_1, the quadratic form Q_1 that appears in the probability density $P(\nu_1,u_1,\nu_2;r)$ reads as

$$Q_1(\nu_1,u_1,\nu_2) = \frac{\nu_1^2 + u_1^2 - 2\gamma_1\nu_1u_1}{2(1-\gamma_1^2)} + \frac{(\nu_2 - A_1)^2}{2\Delta_\xi}, \tag{A5}$$

where

$$\Delta_\xi = 1 - \frac{1}{\sigma_0^4}\frac{(\xi_0^{(0)} - \frac{\sigma_0}{\sigma_1}\xi_0^{(1)})^2}{1-\gamma_1^2}, \tag{A6}$$

$$A_1 = \frac{1}{\sigma_0}\left[\frac{1}{\sigma_0}\left(\frac{\nu_1 - \gamma_1u_1}{1-\gamma_1^2}\right)\xi_0^{(0)} + \frac{1}{\sigma_2}\left(\frac{u_1 - \gamma_1\nu_1}{1-\gamma_1^2}\right)\xi_0^{(1)}\right]. \tag{A7}$$

The calculation now proceeds along lines similar to [29].

We choose a coordinate system whose axes are aligned with the principal frame of ζ_1, and introduce the asymmetry parameters

$$v = (\lambda_1 - \lambda_3)/2, \qquad w = (\lambda_1 - 2\lambda_2 + \lambda_3)/2. \tag{A8}$$

Our choice of ordering imposes the constraints $v \geq 0$ and $-z \leq w \leq v$, while the peak constraint enforces $(u + w) \geq 3v$. Upon integration over the angular variables that define the orientation of the orthonormal triad of ζ_1 and the variables v and w, the cross correlation $\xi_{\mathrm{pk},\delta}(r)$ of peaks of height ν is given by

$$\xi_{\mathrm{pk},\delta}(r) = \sigma_0G_0(\gamma_1,\gamma_1\nu_1)^{-1}\int_{-\infty}^{+\infty}d\nu_2\nu_2\frac{e^{-(\nu_2-A_1)^2/2\Delta_\xi}}{\sqrt{2\pi\Delta_\xi}}$$
$$\times\int_0^\infty du_1 f(u_1)\frac{e^{-(u_1-\gamma_1\nu_1)^2/2(1-\gamma_1^2)}}{\sqrt{2\pi(1-\gamma_1^2)}}, \tag{A9}$$

where the auxiliary function $f(u)$ is defined as in Eq. (A15) of [17], and

$$G_n(\gamma,\omega) = \int_0^\infty dx x^n f(x)\frac{e^{-(x-\omega)^2/2(1-\gamma^2)}}{\sqrt{2\pi(1-\gamma^2)}} \tag{A10}$$

are moments of the peak curvature. In particular, $\bar{u}(\nu) = G_1(\gamma_1, \gamma_1\nu)/G_0(\gamma_1, \gamma_1\nu)$ is the average curvature of peaks of height ν. Finally, the integral over ν_2 is performed and we arrive at the desired result:

$$\xi_{\mathrm{pk},\delta}(r) = \frac{1}{\sigma_0}\frac{(\nu - \gamma_1\bar{u})}{(1-\gamma_1^2)}\xi_0^{(0)}(r) + \frac{1}{\sigma_2}\frac{(\bar{u} - \gamma_1\nu)}{(1-\gamma_1^2)}\xi_0^{(1)}(r). \tag{A11}$$

This agrees with Eq. (18), which was obtained with much less effort from the peak biasing relation (4). It is worth noticing that, while the derivation based on the peak biasing relation is formally exact at the first order only, this appendix shows that Eq. (A11) is exact to all orders.

The cross correlation $\xi_{\mathrm{pk},\delta}(r)$ has a straightforward interpretation: it is the average density profile around density maxima, i.e. $\langle\delta(\boldsymbol{x}_2)|\text{peak at } \boldsymbol{x}_1\rangle$. As shown by [17], this constrained density profile can be calculated easily and, after some algebra, one indeed finds $\xi_{\mathrm{pk},\delta}(r) = \langle\delta(\boldsymbol{x}_2)|\text{peak at } \boldsymbol{x}_1\rangle$. Note that $\psi(r)$ in Eq. (7.10) of Ref. [17] corresponds to our $\xi_0^{(0)}(r)$. Therefore, their Eq. (7.10) appears to have an additional factor of $1/3$ which multiplies the factors of $\nabla^2\xi_0^{(0)} = -\xi_0^{(1)}$ in our expression—but this is only because they measure r in units of R_1—there is, in fact, no difference.

2. Mean streaming of peak pairs

The calculation of the mean streaming is more involved since we have three more degrees of freedom and an extra angular dependence. Our derivation is based on Ref. [29], which calculated the pairwise velocity dispersion along the line of sight.

We introduce the normalized velocity field $\boldsymbol{\varpi}_i = \mathrm{v}_i/(aHf\sigma_{-1}) = \boldsymbol{v}_i/\sigma_{-1}$. Also, we assume that the line-of-sight axis coincides with the third axis, such that $\Delta\boldsymbol{\varpi}_z$ denotes the difference $\boldsymbol{\varpi}_3(\boldsymbol{x}_2) - \boldsymbol{\varpi}_3(\boldsymbol{x}_1)$. The line-of-sight pairwise velocity weighted over all peak pairs with comoving separation r can be expressed as

$$\begin{aligned}&[1 + \xi_{\mathrm{pk}}(r)]v_{12}(r,\mu)\\ &= \bar{n}_{\mathrm{pk}}^{-2}\sigma_{-1}\frac{1}{2\pi}\int_0^{2\pi} d\phi d\boldsymbol{y}_1 d\boldsymbol{y}_2 \Delta\boldsymbol{\varpi}_z n_{\mathrm{pk}}(\boldsymbol{x}_1)n_{\mathrm{pk}}(\boldsymbol{x}_2)\\ &\quad\times P(\boldsymbol{y}_1,\boldsymbol{y}_2;\boldsymbol{r}),\end{aligned} \tag{A12}$$

where μ is the cosine of the angle between $\hat{\boldsymbol{r}} = \boldsymbol{r}/r$ and the third axis, and ϕ is the azimuthal angle in the plane perpendicular to the line of sight. The local peak density $n_{\mathrm{pk}}(\boldsymbol{x})$ is given by $3^{3/2}R_1^{-3}|\det\zeta(\boldsymbol{x})|\delta^3[\eta(\boldsymbol{x})]$, supplemented by the appropriate conditions to select those maxima with a certain threshold height.

In the above expression, the joint probability density $P(\boldsymbol{y}_1,\boldsymbol{y}_2;\boldsymbol{r})$ is now a function of $\boldsymbol{y}^\top = (\boldsymbol{\varpi}_i, \eta_i, \nu, \zeta_A)$, where $\eta_i \equiv 0$ owing to the peak constraint. The corresponding covariance matrix C can be decomposed into four 13×13 block matrices, the zero-point contribution M in the top left and bottom right corners, and the cross-correlation matrix $\mathrm{B}(\boldsymbol{r})$ and its transpose in the bottom left and top right corners, respectively. In the large distance limit $r \gg 1$ where $|\mathrm{B}| \ll \mathrm{M}$, an expansion in the small perturbation M yields, at first order,

$$\begin{aligned}P(\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{r}) &\approx \frac{1}{(2\pi)^{13}|\det\mathrm{C}|^{1/2}}\\ &\quad\times(1 + \boldsymbol{y}_1^\top\mathrm{M}^{-1}\mathrm{B}\mathrm{M}^{-1}\boldsymbol{y}_2)e^{-\bar{Q}(\boldsymbol{y}_1,\boldsymbol{y}_2)},\end{aligned} \tag{A13}$$

where the quadratic form $\bar{Q}(\boldsymbol{y}_1,\boldsymbol{y}_2)$ reads

$$\begin{aligned}2\bar{Q} &= \frac{3\boldsymbol{\varpi}_1^2}{1-\gamma_0^2} + \nu_1^2 + \frac{(\gamma_1\nu_1 + \mathrm{tr}\zeta_1)^2}{1-\gamma_1^2}\\ &\quad+\frac{5}{2}[3\,\mathrm{tr}(\zeta_1^2) - (\mathrm{tr}\zeta_1)^2] + 1 \leftrightarrow 2,\end{aligned} \tag{A14}$$

$\boldsymbol{\varpi}_1$ being the velocity vector at comoving position $\boldsymbol{x}_1$. The inverse M^{-1} and $\mathrm{B}(\boldsymbol{r})$ can be further decomposed into the block matrices

$$\mathrm{M}^{-1} = \begin{pmatrix}\mathrm{P} & \mathrm{R}^\top\\ \mathrm{R} & \mathrm{Q}\end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix}\mathrm{B}_1 & \mathrm{B}_4^\top\\ \mathrm{B}_3 & \mathrm{B}_2\end{pmatrix}, \tag{A15}$$

where

$$\mathrm{P} = \begin{pmatrix}\frac{3}{(1-\gamma_0^2)}\mathrm{I} & \frac{-3\gamma_0}{1-\gamma_0^2}\mathrm{I} & 0_{3\times1}\\ \frac{-3\gamma_0}{1-\gamma_0^2}\mathrm{I} & \frac{3}{1-\gamma_0^2}\mathrm{I} & 0_{3\times1}\\ 0_{1\times3} & 0_{1\times3} & (1-\gamma_1^2)^{-1}\end{pmatrix},$$

$$\mathrm{Q} = \begin{pmatrix}\frac{6-5\gamma_1^2}{1-\gamma_1^2} & -\frac{(3-5\gamma_1^2)}{2(1-\gamma_1^2)} & -\frac{(3-5\gamma_1^2)}{2(1-\gamma_1^2)} & 0 & 0 & 0\\ -\frac{(3-5\gamma_1^2)}{2(1-\gamma_1^2)} & \frac{6-5\gamma_1^2}{1-\gamma_1^2} & -\frac{(3-5\gamma_1^2)}{2(1-\gamma_1^2)} & 0 & 0 & 0\\ -\frac{(3-5\gamma_1^2)}{2(1-\gamma_1^2)} & -\frac{(3-5\gamma_1^2)}{2(1-\gamma_1^2)} & \frac{6-5\gamma_1^2}{1-\gamma_1^2} & 0 & 0 & 0\\ 0 & 0 & 0 & 15 & 0 & 0\\ 0 & 0 & 0 & 0 & 15 & 0\\ 0 & 0 & 0 & 0 & 0 & 15\end{pmatrix}, \tag{A16}$$

and

$$\mathrm{R} = \begin{pmatrix}0_{3\times3} & 0_{3\times3} & \frac{\gamma_1}{1-\gamma_1^2}1_{3\times1}\\ 0_{3\times3} & 0_{3\times3} & 0_{3\times1}\end{pmatrix}. \tag{A17}$$

The explicit expressions for B_i are too long to be given here as they depend upon the correlation functions of $\boldsymbol{\varpi}_i$, η_i, ν, and ζ_A in a rather complicated way. Fortunately, the mean streaming involves only the azimuthal average $\tilde{\mathrm{B}}(r,\mu) = 1/(2\pi)\int d\phi\mathrm{B}(\boldsymbol{r})$, which can generally be expanded as

$$\tilde{\mathrm{B}}(r,\mu) = \sum_{\ell=0}^{4}\tilde{\mathrm{B}}^\ell(r)L_\ell(\mu), \qquad \tilde{\mathrm{B}}^\ell(r) = \begin{pmatrix}\tilde{\mathrm{B}}_1^\ell & \tilde{\mathrm{B}}_4^{\ell\top}\\ \tilde{\mathrm{B}}_3^\ell & \tilde{\mathrm{B}}_2^\ell\end{pmatrix}. \tag{A18}$$

Note that the multipole matrices $\tilde{\mathrm{B}}_3^\ell$ and $\tilde{\mathrm{B}}_4^{\ell\mathsf{T}}$ are not independent of each other since we have $\tilde{\mathrm{B}}_3^\ell = (-1)^\ell \tilde{\mathrm{B}}_4^\ell$. As we will see shortly, all the contributions but that from the $\ell = 1$ multipole (unsurprisingly) cancel out. We will thus detail solely the results for the dipole contribution. After some algebra, we find

$$\tilde{\mathrm{B}}_1^1(r) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\xi_1^{(-1/2)}}{\sigma_0\sigma_{-1}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\xi_1^{(1/2)}}{\sigma_0\sigma_1} \\ 0 & 0 & -\frac{\xi_1^{(-1/2)}}{\sigma_0\sigma_{-1}} & 0 & 0 & -\frac{\xi_1^{(1/2)}}{\sigma_0\sigma_1} & 0 \end{pmatrix},$$

$$\tilde{\mathrm{B}}_4^1(r) = \begin{pmatrix} 0 & 0 & \frac{\xi_1^{(1/2)}}{5\sigma_{-1}\sigma_2} & 0 & 0 & \frac{\xi_1^{(3/2)}}{5\sigma_1\sigma_2} & 0 \\ 0 & 0 & \frac{\xi_1^{(1/2)}}{5\sigma_{-1}\sigma_2} & 0 & 0 & \frac{\xi_1^{(3/2)}}{5\sigma_1\sigma_2} & 0 \\ 0 & 0 & \frac{3\xi_1^{(1/2)}}{5\sigma_{-1}\sigma_2} & 0 & 0 & \frac{3\xi_1^{(3/2)}}{5\sigma_1\sigma_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\xi_1^{(1/2)}}{5\sigma_{-1}\sigma_2} & 0 & 0 & \frac{\xi_1^{(3/2)}}{5\sigma_1\sigma_2} & 0 & 0 & 0 \\ 0 & \frac{\xi_1^{(1/2)}}{5\sigma_{-1}\sigma_2} & 0 & 0 & \frac{\xi_1^{(3/2)}}{5\sigma_1\sigma_2} & 0 & 0 \end{pmatrix}, \tag{A19}$$

whereas the matrix $\tilde{\mathrm{B}}_2^1$ is identically zero. Right and left multiplication by M^{-1} then gives

$$\mathrm{M}^{-1}\tilde{\mathrm{B}}_1^1\mathrm{M}^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3\alpha_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3\alpha_2 \\ 0 & 0 & 3\alpha_1 & 0 & 0 & 3\alpha_2 & 0 \end{pmatrix},$$

$$\mathrm{M}^{-1}\tilde{\mathrm{B}}_4^1\mathrm{M}^{-1} = \begin{pmatrix} 0 & 0 & -3\gamma_1\alpha_1 & 0 & 0 & -3\gamma_1\alpha_2 & 0 \\ 0 & 0 & -3\gamma_1\alpha_1 & 0 & 0 & -3\gamma_1\alpha_2 & 0 \\ 0 & 0 & 9\alpha_3 - 3\gamma_1\alpha_1 & 0 & 0 & 9\alpha_4 - 3\gamma_1\alpha_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9\alpha_3 & 0 & 0 & 9\alpha_4 & 0 & 0 & 0 \\ 0 & 9\alpha_3 & 0 & 0 & 9\alpha_4 & 0 & 0 \end{pmatrix}, \tag{A20}$$

where the functions $\alpha_i(r)$ are identical to those defined in Eq. (A12) of [29]. Namely,

$$\alpha_1(r) = \frac{\frac{\sigma_0^2}{\sigma_2^2}\xi_1^{(3/2)} + \frac{\sigma_0^2}{\sigma_1^2}\xi_1^{(1/2)} - \frac{\sigma_1^2}{\sigma_2^2}\xi_1^{(1/2)} - \xi_1^{(-1/2)}}{\sigma_{-1}\sigma_0(1-\gamma_0^2)(1-\gamma_1^2)}, \qquad \alpha_3(r) = \frac{\xi_1^{(1/2)} - \frac{\sigma_0^2}{\sigma_1^2}\xi_1^{(3/2)}}{\sigma_{-1}\sigma_2(1-\gamma_0^2)},$$
$$\alpha_2(r) = \frac{-\frac{\sigma_1^2}{\sigma_2^2}\xi_1^{(3/2)} + \frac{\sigma_0^2\sigma_1^2}{\sigma_{-1}^2\sigma_2^2}\xi_1^{(1/2)} - \xi_1^{(1/2)} + \frac{\sigma_0^2}{\sigma_{-1}^2}\xi_1^{(-1/2)}}{\sigma_1\sigma_0(1-\gamma_0^2)(1-\gamma_1^2)}, \qquad \alpha_4(r) = \frac{\xi_1^{(3/2)} - \frac{\sigma_0^2}{\sigma_{-1}^2}\xi_1^{(1/2)}}{\sigma_1\sigma_2(1-\gamma_0^2)}. \tag{A21}$$

The rest of the calculation is easily accomplished owing to the separability of the one-point probability distribution $P(\boldsymbol{y})$ into the product $P_{\varpi}(\varpi_i)P_{\nu\zeta}(\nu, \zeta_A)$, where $P_{\nu\zeta}$ is the one-point distribution of the density and its second derivatives (the first derivatives merely contribute a normalization factor), and

$$P_{\varpi}(\varpi_i) = \frac{3^{3/2}}{(2\pi)^{3/2}(1-\gamma_0^2)^{3/2}} \exp\left[-\frac{3\varpi^2}{2(1-\gamma_0^2)}\right] \quad \text{(A22)}$$

is the velocity distribution of density peaks. In particular, the first moment vanishes while the second moment $\langle\varpi_i^2\rangle$ is the one-dimensional velocity dispersion of density maxima,

$$\langle\varpi_i^2\rangle = \tfrac{1}{3}(1-\gamma_0^2). \quad \text{(A23)}$$

The scalar $\boldsymbol{y}_1^{\top}\mathrm{M}^{-1}\tilde{\mathrm{B}}\mathrm{M}^{-1}\boldsymbol{y}_2$ contains terms linear and quadratic in ϖ_i as well as terms independent of the velocity. Upon multiplication by $\Delta\varpi_z$ and integration over the velocities, only quadratic terms survive. We eventually find

$$\begin{aligned}
&\int d^3\boldsymbol{\varpi}_1 d^3\boldsymbol{\varpi}_2 \Delta\varpi_z (\boldsymbol{y}_1^{\top}\mathrm{M}^{-1}\tilde{\mathrm{B}}\mathrm{M}^{-1}\boldsymbol{y}_2) P_{\varpi}(\boldsymbol{\varpi}_1) P_{\varpi}(\boldsymbol{\varpi}_2) \\
&= [\alpha_1(\nu_1+\nu_2) + \gamma_1\alpha_1(\mathrm{tr}\zeta_1 + \mathrm{tr}\zeta_2) \\
&\quad - 3\alpha_3(\zeta_{1,3}+\zeta_{2,3})](1-\gamma_0^2)L_1(\mu) \\
&\quad - \frac{3}{2}[\zeta_{1,1}+\zeta_{2,1}+\zeta_{1,2}+\zeta_{2,2}-2(\zeta_{1,3}+\zeta_{2,3})] \\
&\quad \times \frac{\xi_3^{(1/2)} - \frac{\sigma_0^2}{\sigma_1^2}\xi_3^{(3/2)}}{\sigma_{-1}\sigma_2} L_3(\mu).
\end{aligned} \quad \text{(A24)}$$

Here, $\zeta_{1,A}$ and $\zeta_{2,A}$ designate the component ζ_A of the Hessian ζ at locations $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively. As we can see, although the even multipoles cancel out, a term proportional to $L_3(\mu)$ remains. Also, the dipole receives a contribution from $-3\alpha_3(\zeta_{1,3}+\zeta_{2,3})$ which is not invariant under rotations. However, we have to remember that the principal axes of the tensors $\zeta_1 = \zeta(\boldsymbol{x}_1)$ and $\zeta_2 = \zeta(\boldsymbol{x}_2)$ are not necessarily aligned with those of the coordinate frame. Let us first consider ζ_1. Without loss of generality, we can write $\zeta_1 = -\mathcal{O}\Lambda\mathcal{O}^{\top}$, where $\mathcal{O}$ is an orthogonal matrix and Λ is the diagonal matrix consisting of the three ordered eigenvalues λ_i of $-\zeta_1$. The properties of the trace imply that $\mathrm{tr}\zeta_1 = -\mathrm{tr}\Lambda$, while $\zeta_{1,j} = -\sum_i \lambda_i \mathcal{O}_{ji}^2$. Since the one-point probability density $P(\boldsymbol{y})$ does not depend on $\mathcal{O}$, the integral over the SO(3) manifold that describes the orientation of the orthonormal triad of ζ_1 is immediate,

$$\int d\mathcal{O}\zeta_{1,j} = -\sum_{i=1}^{3}\lambda_i \int d\mathcal{O}\mathcal{O}_{ji}^2 = -\frac{1}{3}\sum_{i=1}^{3}\lambda_i = \frac{1}{3}\mathrm{tr}\zeta_1. \quad \text{(A25)}$$

Similarly, averaging over the orientation of the eigenvectors of ζ_2 yields $\int d\mathcal{O}\zeta_{2,j} = (1/3)\,\mathrm{tr}\zeta_2$. Consequently, the $\ell = 3$ term vanishes and we only need to integrate

$$[\alpha_1(\nu_1+\nu_2) + (\gamma\alpha_1 - \alpha_3)(\mathrm{tr}\zeta_1 + \mathrm{tr}\zeta_2)](1-\gamma_0^2)L_1(\mu) \quad \text{(A26)}$$

over the eigenvalues of ζ_1 and ζ_2 subject to the peak constraint. Substituting the expressions (5) of the bias parameters b_ν and b_ζ, the result can be recast into the form of Eq. (34) when $\nu_1 = \nu_2 = \nu$.

APPENDIX B: THE CROSS-CORRELATION COEFFICIENT FOR POWER-LAW SPECTRA

In this appendix, we examine the large-scale behavior of the cross-correlation coefficient for density peaks, assuming a power-law spectrum of density fluctuations.

The cross-correlation coefficient r_ξ in configuration space can be written as

$$r_\xi^2(r) = \frac{1}{1+\mathcal{R}(r)}, \qquad \mathcal{R}(r) = \frac{\xi_0^{(2)}/\xi_0^{(0)} - [\xi_0^{(1)}/\xi_0^{(0)}]^2}{[b_\nu/b_\zeta + \xi_0^{(1)}/\xi_0^{(0)}]^2}. \quad \text{(B1)}$$

It is larger than unity when $\mathcal{R} < 0$, i.e. when $\xi_0^{(2)}/\xi_0^{(0)} < [\xi_0^{(1)}/\xi_0^{(0)}]^2$. In cold dark matter cosmologies, the correlation functions $\xi_0^{(n)}$ must be calculated numerically because the spectral index is a smooth function of wave number. For a no-wiggle power-law power spectrum $P_\delta(k) \equiv A_s k^{n_s}$ however, they take the exact form

$$\xi_0^{(n)}(r) = \frac{A_s}{4\pi^2} R_S^{-2\alpha}\Gamma(\alpha)\,{}_1F_1(\alpha, \gamma; -z), \quad \text{(B2)}$$

where $\Gamma(\alpha)$ and ${}_1F_1(\alpha, \gamma; -z)$ are the Gamma and confluent hypergeometric functions in the arguments $\alpha = n + 3/2 + n_s/2$, $\gamma = 3/2$, and $z = r^2/(4R_S^2)$; and R_S is the characteristic radius of the window function assumed Gaussian. Since Eq. (B1) only holds at large separation $r \gg 1$, we consider the limit $|z| \to \infty$ to the above expression, in which ${}_1F_1(\alpha, \gamma; -z)$ has the following asymptotic expansion (in a suitable domain of the complex plane [91]),

$$\begin{aligned}
{}_1F_1(\alpha, \gamma; -z) &= \frac{\Gamma(\gamma)}{\Gamma(\alpha)} e^{-z}(-z)^{\alpha-\gamma}\left(1 + \sum_{k=1}^{\infty}(-1)^k \frac{\Gamma(k+\gamma-\alpha)\Gamma(k+1-\alpha)}{k!\Gamma(\gamma-\alpha)\Gamma(1-\alpha)} z^{-k}\right) \\
&\quad + \frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)} z^{-\alpha}\left(1 + \sum_{k=1}^{\infty}\frac{\Gamma(k+\alpha)\Gamma(k+\alpha-\gamma+1)}{k!\Gamma(\alpha)\Gamma(\alpha-\gamma+1)} z^{-k}\right).
\end{aligned} \quad \text{(B3)}$$

In the right half-plane of the variable z [i.e. for $\mathrm{Re}(z) > 0$], the first term in the right-hand side of Eq. (B3) is the subordinate part of the asymptotics, whereas the second is the dominant part. We will now calculate $\mathcal{R}(r)$ for a few integer values of the spectral index covering the range $[-3, 2]$ shown in Fig. 7. As we will see shortly, the cross-correlation coefficient can exceed unity even if the underlying power spectrum $P(k)$ is a featureless power law. The exact amount of stochasticity, however, critically depends upon the shape of the underlying power spectrum.

In the particular case of a white noise spectrum, $n_s = 0$, the dominant part cancels out owing to the fact that $\Gamma(\gamma - \alpha)$ has simple poles at $\gamma - \alpha = n = 0, 1, 2, \ldots$, i.e. $\Gamma(-n)^{-1} = 0$. Moreover, $\Gamma(k + \gamma - \alpha)/\Gamma(\gamma - \alpha) = (k - 1 + \gamma - \alpha)(k - 2 + \gamma - \alpha) \times \cdots \times (\gamma - \alpha)$ vanishes when $k \geq n + 1$, so the summation in the subordinate part involves solely a few terms. In fact, the asymptotic expansion gives the exact result,

$$\xi_0^{(0)}(z) = \frac{A_s}{8\pi^{3/2}R_S^3} e^{-z},$$
$$\xi_0^{(1)}(z) = \frac{1}{R_S^2}\xi_0^{(0)}(z)\left(\frac{3}{2} - z\right), \tag{B4}$$
$$\xi_0^{(2)}(z) = \frac{1}{R_S^4}\xi_0^{(0)}(z)\left(\frac{15}{4} - 5z + z^2\right),$$

which yields

$$\mathcal{R}(r) = \frac{(3 - \frac{r^2}{R_S^2})}{2[\frac{b_\nu R_S^2}{b_\zeta} + \frac{3}{2} - \frac{r^2}{4R_S^2}]^2}. \tag{B5}$$

As can be seen, $\mathcal{R}(r)$ becomes negative at separation $r > \sqrt{3}R_S$, so the cross-correlation coefficient is greater than unity at large scales. Note, however, that the dominant part is nonzero for any small n_s different from zero. More precisely, upon writing $n_s = \epsilon$ where $0 < |\epsilon| \ll 1$ and momentarily ignoring a factor of $A_s/(4\pi^2)R_S^{-2\alpha}$, we have

$$\xi_0^{(n)}(z) \approx \frac{\Gamma(n + \frac{3}{2})\Gamma(3/2)}{\Gamma(-n - \frac{\epsilon}{2})} z^{-n-(3/2)}. \tag{B6}$$

The sign of $\mathcal{R}(r)$ is equal to that of

$$\xi_0^{(2)}/\xi_0^{(0)} - [\xi_0^{(1)}/\xi_0^{(0)}]^2 \approx \frac{15}{4z^2}\frac{\Gamma(-2 - \frac{\epsilon}{2})}{\Gamma(-\frac{\epsilon}{2})} - \frac{3}{2z^2}\left[\frac{\Gamma(-1 - \frac{\epsilon}{2})}{\Gamma(-\frac{\epsilon}{2})}\right]^2 \geq 0, \tag{B7}$$

which is positive for any small nonzero ϵ. Therefore, the cross-correlation coefficient is discontinuous at $n_s = 0$. The same analysis also shows there is a similar discontinuity at $n_s = 2$. These discontinuity points are marked as empty symbols in Fig. 7.

For $n_s = -2$, the dominant part is nonvanishing only when $n = 0$. Furthermore, for $n = 1$ and 2, the subordinate part only sums a finite number of terms. Explicitly,

$$\xi_0^{(0)}(z) \approx \frac{A_s}{8\pi R_S} z^{-1/2}, \qquad \xi_0^{(1)}(z) = \frac{A_s}{8\pi^{3/2}R_S^3} e^{-z},$$
$$\xi_0^{(2)}(z) = \frac{3A_s}{16\pi^{3/2}R_S^5} e^{-z}\left(1 - \frac{2z}{3}\right). \tag{B8}$$

On inserting these expressions into Eq. (B1), we find

$$\mathcal{R}(r) \approx \frac{3\sqrt{\pi}}{4} \frac{(\frac{r}{R_S} - \frac{r^3}{6R_S^3})e^{-r^2/4R_S^2}}{[\sqrt{\pi}\frac{b_\nu R_S^2}{b_\zeta} + \frac{r}{2R_S}e^{-r^2/4R_S^2}]^2}. \tag{B9}$$

Again, $r_\xi > 1$ at sufficiently large separation $r \gg 1$. Note, however, that $\mathcal{R}$ decays much more rapidly to zero when $n_s = -2$. Furthermore, one can show that $r_\xi < 1$ for $n_s = -2 + \epsilon$, and $r_\xi > 1$ for $n_s = -2 - \epsilon$, where $0 < \epsilon \ll 1$. In other words, there is a jump discontinuity at $n_s = -2$.

When the spectral index is an odd integer, e.g. $n_s = -3$, ± 1, the subordinate, complex-valued part is exponentially suppressed relative to the dominant, real-valued part. For $n_s = -1$, we find

$$\xi_0^{(0)}(z) \approx \frac{A_s}{8\pi^2 R_S^2} z^{-1}\left(1 + \frac{1}{2z}\right), \tag{B10}$$

$$\xi_0^{(1)}(z) \approx -\frac{A_s}{16\pi^2 R_S^4} z^{-2}\left(1 + \frac{3}{z}\right), \tag{B11}$$

$$\xi_0^{(2)}(z) \approx \frac{3A_s}{16\pi^2 R_S^6} z^{-3}\left(1 + \frac{15}{2z}\right) \tag{B12}$$

upon including the first two terms of the dominant part. After some simplification, we arrive at

$$\mathcal{R}(r) \approx \frac{20}{r^4}\left[\frac{b_\nu}{b_\zeta} - \frac{2}{r^2}\right]^{-2}. \tag{B13}$$

Similarly, we obtain

$$\mathcal{R}(r) \approx \frac{216}{r^4}\left[\frac{b_\nu}{b_\zeta} - \frac{12}{r^2}\right]^{-2} \tag{B14}$$

for $n_s = -1$. In both cases, $\mathcal{R} > 0$ so the cross-correlation coefficient is less than unity at large scales. Finally, for $n_s = -3$, the density correlation $\xi_0^{(0)}(r)$ diverges owing to the presence of $\Gamma(\alpha) = \Gamma(n)$. Consequently, the cross-correlation coefficient is unity at all scales.

[1] M. Davis and P. J. E. Peebles, Astrophys. J. **267**, 465 (1983); P. B. Lilje and G. Efstathiou, Mon. Not. R. Astron. Soc. **236**, 851 (1989); J. A. Peacock and S. J. Dodds, Mon. Not. R. Astron. Soc. **267**, 1020 (1994); A. N. Taylor and A. J. S. Hamilton, Mon. Not. R. Astron. Soc. **282**, 767 (1996); W. E. Ballinger, J. A. Peacock, and A. F. Heavens, Mon. Not. R. Astron. Soc. **282**, 877 (1996); J. Loveday, G. Efstathiou, S. J. Maddox, and B. A. Peterson, Astrophys. J. **468**, 1 (1996); A. F. Heavens, S. Matarrese, and L. Verde, Mon. Not. R. Astron. Soc. **301**, 797 (1998); H. Magira, Y. P. Jing, and Y. Suto, Astrophys. J. **528**, 30 (2000); X. Kang, Y. P. Jing, H. J. Mo, and G. Börner, Mon. Not. R. Astron. Soc. **336**, 892 (2002); V. Desjacques and A. Nusser, Mon. Not. R. Astron. Soc. **351**, 1395 (2004); X. Wang and W. Hu, Astrophys. J. **643**, 585 (2006); R. E. Smith, R. K. Sheth, and R. Scoccimarro, Phys. Rev. D **78**, 023523 (2008); J. R. Shaw and A. Lewis, Phys. Rev. D **78**, 103512 (2008).

[2] N. Kaiser, Mon. Not. R. Astron. Soc. **227**, 1 (1987).

[3] K. B. Fisher, Astrophys. J. **448**, 494 (1995).

[4] Y. Ohta, I. Kayo, and A. Taruya, Astrophys. J. **608**, 647 (2004).

[5] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980).

[6] A. Lue, R. Scoccimarro, and G. Starkman, Phys. Rev. D **69**, 124015 (2004); E. V. Linder, Phys. Rev. D **70**, 023511 (2004); L. Knox, Y.-S. Song, and J. A. Tyson, Phys. Rev. D **74**, 023512 (2006); M. Ishak, A. Upadhye, and D. Spergel, Phys. Rev. D **74**, 043513 (2006).

[7] R. Dürrer and R. Maartens, arXiv:00811.4132.

[8] A. J. S. Hamilton, Astrophys. J. Lett. **385**, L5 (1992).

[9] S. Cole, K. B. Fisher, and D. Weinberg, Mon. Not. R. Astron. Soc. **275**, 515 (1995).

[10] R. Scoccimarro, Phys. Rev. D **70**, 083007 (2004).

[11] T. Matsubara, Astrophys. J. **525**, 543 (1999).

[12] A. G. Doroshkevich, Astrofiz. **3**, 175 (1970).

[13] K. Gorski, Astrophys. J. Lett. **332**, L7 (1988).

[14] N. Kaiser, Astrophys. J. **284**, L9 (1984).

[15] J. A. Peacock and A. F. Heavens, Mon. Not. R. Astron. Soc. **217**, 805 (1985).

[16] Y. Hoffman and J. Shaham, Astrophys. J. **297**, 16 (1985).

[17] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).

[18] P. Coles, Mon. Not. R. Astron. Soc. **238**, 319 (1989).

[19] S. L. Lumsden, A. F. Heavens, and J. A. Peacock, Mon. Not. R. Astron. Soc. **238**, 293 (1989).

[20] E. Regös and A. S. Szalay, Mon. Not. R. Astron. Soc. **272**, 447 (1995).

[21] J. R. Bond and S. T. Myers, Astrophys. J. Suppl. Ser. **103**, 1 (1996).

[22] R. K. Sheth, H. J. Mo, and G. Tormen, Mon. Not. R. Astron. Soc. **323**, 1 (2001).

[23] V. Desjacques, Mon. Not. R. Astron. Soc. **388**, 638 (2008).

[24] V. Desjacques and R. E. Smith, Phys. Rev. D **78**, 023527 (2008).

[25] N. Kaiser and M. Davis, Astrophys. J. **297**, 365 (1985).

[26] H. J. Mo, Y. P. Jing, and S. D. M. White, Mon. Not. R. Astron. Soc. **284**, 189 (1997).

[27] R. Cen, Astrophys. J. **509**, 494 (1998).

[28] R. K. Sheth, Ann. N.Y. Acad. Sci. **927**, 1 (2001).

[29] V. Desjacques, Phys. Rev. D **78**, 103503 (2008).

[30] W. J. Percival and M. White, Mon. Not. R. Astron. Soc. **393**, 297 (2009).

[31] E. Komatsu *et al.*, Astrophys. J. Suppl. Ser. **180**, 330 (2009).

[32] J. E. Gunn and J. R. Gott, III, Astrophys. J. **176**, 1 (1972).

[33] W. H. Press and P. Schechter, Astrophys. J. **187**, 425 (1974).

[34] R. Mandelbaum, U. Seljak, R. J. Cool, M. Blanton, C. M. Hirata, and J. Brinkmann, Mon. Not. R. Astron. Soc. **372**, 758 (2006).

[35] G. Kulkarni *et al.*, Mon. Not. R. Astron. Soc. **378**, 1196 (2007).

[36] D. Wake *et al.*, Mon. Not. R. Astron. Soc. **387**, 1045 (2008).

[37] A. S. Szalay, Astrophys. J. **333**, 21 (1988).

[38] J. N. Fry and E. Gaztañaga, Astrophys. J. **413**, 447 (1993).

[39] P. Coles, Mon. Not. R. Astron. Soc. **262**, 1065 (1993).

[40] In terms of the *normalized* (and smoothed) variables $\nu_S = \delta_S/\sigma_0$ and $u_S = -\nabla^2\delta_S/\sigma_2$, the peak number density is $\delta n_{\rm pk} = \sigma_0 b_\nu \nu_S + \sigma_2 b_\zeta u_S$ at the first order. This shows that the relative importance of the b_ν and b_ζ terms is controlled by $\sigma_0 b_\nu$ and $\sigma_2 b_\zeta$.

[41] H. J. Mo and S. D. M. White, Mon. Not. R. Astron. Soc. **282**, 347 (1996).

[42] S. Cole and N. Kaiser, Mon. Not. R. Astron. Soc. **237**, 1127 (1989).

[43] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999).

[44] A. Taruya, H. Magira, Y. P. Jing, and Y. Suto, Publ. Astron. Soc. Jpn. **53**, 155 (2001).

[45] A. S. Szalay and L. G. Jensen, Acta Phys. Hung. **62**, 263 (1987).

[46] J. A. Peacock, S. L. Lumsden, and A. F. Heavens, Mon. Not. R. Astron. Soc. **229**, 469 (1987).

[47] W. J. Percival and B. M. Schäfer, Mon. Not. R. Astron. Soc. **385**, L78 (2008).

[48] S. Bharadwaj, Mon. Not. R. Astron. Soc. **327**, 577 (2001).

[49] Our expression for v_{12} corrects a sign error in Eq. (50) of [29], which propagated to Fig. 8 of that paper.

[50] We believe the expression for the real-space peak power spectrum in Ref. [20], their Eq. (70), should read $\Pi(k) = P(k)(x+y)^2/\sigma_0^2$. Once corrected, this relation is equivalent to our Eq. (19) provided that $\sigma_0 b_\nu = \langle W|\mathcal{C}\rangle/\sqrt{1-\gamma_1^2}$ and $\sigma_2 b_\zeta = \langle X|\mathcal{C}\rangle - \gamma_1\langle W|\mathcal{C}\rangle/\sqrt{1-\gamma_1^2}$ (see their paper for details about their notation). For the redshift space power, they have the same Gaussian damping term as we do, but their expression for $P^{s0}_{\rm pk}$, their Eq. (84), does not reduce to the square of peak density and velocity bias terms. In their Eq. (84), w should be a y, and $x - y$ should be $x + y$ (this is the same error that affected their expression for the real-space power spectrum; it also affects their expression for v_{12}). These errors appear to have propagated to their Fig. 6.

[51] R. K. Sheth, L. Hui, A. Diaferio, and R. Scoccimarro, Mon. Not. R. Astron. Soc. **325**, 1288 (2001).

[52] J. C. Jackson, Mon. Not. R. Astron. Soc. **156**, 1 (1972).

[53] A. J. S. Hamilton, *Linear Redshift Distortions: A Review*, in The Evolving Universe, edited by D. Hamilton (Kluwer Academic Publishers, Dordrecht, Holland, 1998).

[54] R. K. Sheth and A. Diaferio, Mon. Not. R. Astron. Soc. **322**, 901 (2001).
[55] S. Bharadwaj, Astrophys. J. **460**, 28 (1996).
[56] M. Crocce and R. Scoccimarro, Phys. Rev. D **77**, 023533 (2008).
[57] T. Matsubara, Phys. Rev. D **77**, 063530 (2008).
[58] T. Matsubara, Phys. Rev. D **78**, 083519 (2008).
[59] D. J. Eisenstein, H.-J. Seo, and M. White, Astrophys. J. **664**, 660 (2007).
[60] R. K. Sheth, Mon. Not. R. Astron. Soc. **279**, 1310 (1996).
[61] M. Tegmark *et al.*, Phys. Rev. D **74**, 123507 (2006).
[62] J. Kim, C. Park, R. J. Gott, and J. Dubinski, Astrophys. J. **701**, 1547 (2009).
[63] M. Manera, R. K. Sheth, and R. Scoccimarro, arXiv:0906.1314.
[64] J. A. Peacock and S. J. Dodds, Mon. Not. R. Astron. Soc. **267**, 1020 (1994).
[65] Y.-S. Song and W. J. Percival, arXiv:astro-ph/0807.0810.
[66] M. White, Y.-S. Song, and W. J. Percival, Mon. Not. R. Astron. Soc. **397**, 1348 (2009).
[67] A. N. Taylor, W. E. Ballinger, A. F. Heavens, and H. Tadros, Mon. Not. R. Astron. Soc. **327**, 689 (2001).
[68] S. Nesseris and L. Perivolaropoulos, Phys. Rev. D **77**, 023504 (2008).
[69] M. S. Vogeley and A. S. Szalay, Astrophys. J. **465**, 34 (1996).
[70] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, Astrophys. J. **499**, 555 (1998).
[71] H. A. Feldman, N. Kaiser, and J. A. Peacock, Astrophys. J. **426**, 23 (1994).
[72] U. Seljak, N. Hamaus, and V. Desjacques, Phys. Rev. Lett. **103**, 091303 (2009).
[73] U. Seljak, Phys. Rev. Lett. **102**, 021302 (2009).
[74] P. McDonald and U. Seljak, J. Cosmol. Astropart. Phys. 10 (2009) 007.
[75] D. Burkey and A. N. Taylor, Mon. Not. R. Astron. Soc. **347**, 255 (2004).
[76] A. Stril, R. N. Cahn, and E. V. Linder, arXiv:0910.1833.
[77] B. Jain and P. Zhang, Phys. Rev. D **78**, 063503 (2008).
[78] Y.-S. Song and K. Koyama, J. Cosmol. Astropart. Phys. 01 (2008) 048.
[79] R. J. Scherrer and D. H. Weinberg, Astrophys. J. **504**, 607 (1998).
[80] U.-L. Pen, Astrophys. J. **504**, 601 (1998).
[81] A. Dekel and O. Lahav, Astrophys. J. **520**, 24 (1999).
[82] J. N. Fry, Astrophys. J. Lett. **461**, L65 (1996).
[83] L. Hui and K. Parfrey, Phys. Rev. D **77**, 043527 (2008).
[84] M. Tegmark and P. J. E. Peebles, Astrophys. J. Lett. **500**, L79 (1998).
[85] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063520 (2006).
[86] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063519 (2006).
[87] M. Martino, H. F. Stabenau, and R. K. Sheth, Phys. Rev. D **79**, 084013 (2009).
[88] D. Jeong and E. Komatsu, Astrophys. J. **691**, 569 (2009).
[89] A. Cooray and R. K. Sheth, Phys. Rep. **372**, 1 (2002).
[90] M. Kac, Bull. Am. Math. Soc. **49**, 314 (1943); S. O. Rice, *Mathematical Analysis of Random Noise, in Selected Papers on Noise and Stochastic Processes* (Dover, New York, 1954).
[91] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic Press, New York, 2000), 6th ed.