11-23-2010

Modeling Scale-Dependent Bias on the Baryonic Acoustic Scale with the Statistics of Peaks of Gaussian Random Fields

Vincent Desjacques
University of Zurich

Martin Crocce
Institut de Ciencies de L'Espai

Roman Scoccimarro
New York University

Ravi K. Sheth
University of Pennsylvania, shethrk@sas.upenn.edu

Suggested Citation:
Desjacques, V., M. Crocce, R. Scoccimarro and R.K. Sheth. (2010). "Modeling scale-dependent bias on the baryonic acoustic scale with the statistics of peaks of Gaussian random fields." *Physical Review D.* 82, 103529.

http://dx.doi.org/10.1103/PhysRevB.82.103529

Modeling Scale-Dependent Bias on the Baryonic Acoustic Scale with the Statistics of Peaks of Gaussian Random Fields

Abstract

Models of galaxy and halo clustering commonly assume that the tracers can be treated as a continuous field locally biased with respect to the underlying mass distribution. In the peak model pioneered by Bardeen et al. [Astrophys. J. 304, 15 (1986)], one considers instead density maxima of the initial, Gaussian mass density field as an approximation to the formation site of virialized objects. In this paper, the peak model is extended in two ways to improve its predictive accuracy. First, we derive the two-point correlation function of initial density peaks up to second order and demonstrate that a peak-background split approach can be applied to obtain the k-independent and k-dependent peak bias factors at all orders. Second, we explore the gravitational evolution of the peak correlation function within the Zel'dovich approximation. We show that the local (Lagrangian) bias approach emerges as a special case of the peak model, in which all bias parameters are scale independent and there is no statistical velocity bias. We apply our formulas to study how the Lagrangian peak biasing, the diffusion due to large scale flows, and the mode coupling due to nonlocal interactions affect the scale dependence of bias from small separations up to the baryon acoustic oscillation (BAO) scale. For 2σ density peaks collapsing at $z = 0.3$, our model predicts a ~5% residual scale-dependent bias around the acoustic scale that arises mostly from first order Lagrangian peak biasing (as opposed to second order gravity mode coupling). We also search for a scale dependence of bias in the large scale autocorrelation of massive halos extracted from a very large N-body simulation provided by the MICE Collaboration. For halos with mass $M \gtrsim 1014 M_{\odot}/h$, our measurements demonstrate a scale-dependent bias across the BAO feature which is very well reproduced by a prediction based on the peak model.

Disciplines

Physical Sciences and Mathematics | Physics

Comments

Suggested Citation:
Desjacques, V., M. Crocce, R. Scoccimarro and R.K. Sheth. (2010). "Modeling scale-dependent bias on the baryonic acoustic scale with the statistics of peaks of Gaussian random fields." *Physical Review D.* 82, 103529.

http://dx.doi.org/10.1103/PhysRevB.82.103529

Modeling scale-dependent bias on the baryonic acoustic scale with the statistics of peaks of Gaussian random fields

Vincent Desjacques,[1,*] Martin Crocce,[2] Roman Scoccimarro,[3] and Ravi K. Sheth[4]

[1]*Institute for Theoretical Physics, University of Zurich, 8057 Zurich, Switzerland*
[2]*Institut de Ciències de l'Espai, IEEC-CSIC, Campus UAB, Facultat de Ciències, Barcelona 08193, Spain*
[3]*Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, New York 10003, USA*
[4]*Center for Particle Cosmology, University of Pennsylvania, 209 South 33rd Street, Philadelphia, Pennsylvania 19104, USA*

(Received 21 September 2010; published 23 November 2010)

Models of galaxy and halo clustering commonly assume that the tracers can be treated as a continuous field locally biased with respect to the underlying mass distribution. In the peak model pioneered by Bardeen *et al.* [Astrophys. J. **304**, 15 (1986)], one considers instead density maxima of the initial, Gaussian mass density field as an approximation to the formation site of virialized objects. In this paper, the peak model is extended in two ways to improve its predictive accuracy. First, we derive the two-point correlation function of initial density peaks up to second order and demonstrate that a peak-background split approach can be applied to obtain the k-independent and k-dependent peak bias factors at all orders. Second, we explore the gravitational evolution of the peak correlation function within the Zel'dovich approximation. We show that the local (Lagrangian) bias approach emerges as a special case of the peak model, in which all bias parameters are scale independent and there is no statistical velocity bias. We apply our formulas to study how the Lagrangian peak biasing, the diffusion due to large scale flows, and the mode coupling due to nonlocal interactions affect the scale dependence of bias from small separations up to the baryon acoustic oscillation (BAO) scale. For 2σ density peaks collapsing at $z = 0.3$, our model predicts a $\sim 5\%$ residual scale-dependent bias around the acoustic scale that arises mostly from first order Lagrangian peak biasing (as opposed to second order gravity mode coupling). We also search for a scale dependence of bias in the large scale autocorrelation of massive halos extracted from a very large N-body simulation provided by the MICE Collaboration. For halos with mass $M \gtrsim 10^{14}\mathrm{M}_{\odot}/h$, our measurements demonstrate a scale-dependent bias across the BAO feature which is very well reproduced by a prediction based on the peak model.

DOI: 10.1103/PhysRevD.82.103529 PACS numbers: 98.80.−k, 95.35.+d, 98.65.−r, 98.80.Es

I. INTRODUCTION

A considerable amount of effort has already been invested in measuring the large scale distribution of galaxies, especially the galaxy two-point correlation function and power spectrum, to constrain viable cosmological models (e.g., Refs. [1–14]). The amplitude, shape, and baryon acoustic feature in these two-point statistics encode a wealth of cosmological information [15–32]. Ongoing and planned galaxy surveys of the high redshift Universe will furnish measurements of the underlying mass distribution with unprecedented precision and statistics. Alongside this great observational effort, interpreting this vast amount of data will require a much better understanding of the relation between the surveyed galaxies and the mass fluctuations they are thought to trace.

Essentially all models of galaxy clustering assume that galaxies are biased tracers of the mass density fluctuation field. Although this bias is expected to be nonlinear, scale-dependent, and stochastic [33,34], the simpler, linear, scale-independent, deterministic model has proved to be an extremely useful first order approximation [35–37]. However, in order to predict corrections beyond linear order to the galaxy two-point correlation, or even the leading-order contribution to higher-order statistics such as the galaxy three-point correlation or bispectrum, one must address the complications which arise from nonlinearity, scale dependence, and stochasticity. For example, if the bias relation is established in coordinate space, then nonlinear biasing will produce scale dependence and stochasticity in Fourier space, and vice versa (e.g., Ref. [38]). This randomness will add to other sources of stochasticity which may arise, for example, from the fact that the formation of galaxies and halos depends on quantities other than the mass density field (e.g., the surrounding tidal field). Moreover, the bias may be established in the initial conditions (Lagrangian bias) or, alternatively, at the present time (Eulerian bias). In the former case, the bias between the tracers and the mass will be affected by the subsequent, nonlinear gravitational evolution. This will introduce additional nonlinearity, scale dependence, and stochasticity. Furthermore, if the velocities of the tracers differ from those of the mass elements, then this will complicate the application of the continuity equation to describe the redshift evolution of bias.

*dvince@physik.uzh.ch

Current analytic approaches to galaxy and dark matter halo clustering take into account some of these complications. In most models, the fundamental quantity is the overdensity of tracers $\delta_{\rm g,h}(R, \mathbf{x})$ within a sphere of radius R centered at position $\mathbf{x}$. It is commonly assumed that $\delta_{\rm g,h}$ is solely a function of the local mass overdensity $\delta_{\rm m}$ [39,40] (see, also, [41]), whose Taylor expansion coefficients are the galaxy or halo bias parameters b_N [42–44]. If this bias is established at a different time than the epoch at which the tracers are observed, then this local bias scheme is combined with some (Eulerian or Lagrangian) perturbative treatment of gravitational instability (see [45], for a review of perturbation theory) to predict the galaxy or halo power spectrum, bispectrum, etc. (e.g., Refs. [43,46–58]). This formalism can be extended to include stochasticity by formulating the biasing in terms of the conditional probability distribution $P(\delta_{\rm g,h}|\delta_{\rm m})$ of $\delta_{\rm g,h}$ at a given $\delta_{\rm m}$ (e.g., Ref. [38,49,59–61]). One of the main caveats with such local biasing schemes is that galaxies (or halos) are treated as though they define a continuous field smoothed on some scale R, whereas they are, in fact, discrete objects.

The peaks' approach to galaxy and dark matter halo clustering is interesting because it exhibits all of the complications mentioned above while also accounting for the discrete nature of the tracers (after all, peaks define a point process). In this model, the fundamental quantity is the set of positions which are local maxima of the density field (from which a peak overabundance $\delta_{\rm pk}(R, \mathbf{x})$ in spheres of radius R could in principle be derived). Since the evolved density field is highly nonlinear, the peak constraint is generally applied to the initial (Lagrangian) Gaussian density field, with the assumption that the most prominent peaks should be in one-to-one correspondence with luminous galaxies or massive halos in the low redshift Universe (see, e.g., Refs. [62–64], for numerical studies of this association). Peak abundances, profiles, and correlation functions in real and redshift space have been studied in the literature [36,65–75]. Some of these results have been used to interpret the abundance and clustering of rich clusters [42,76–80], constrain the power spectrum of mass fluctuations [81,82], and study evolution bias [83] and assembly bias [84].

On asymptotically large scales, peaks are linearly biased tracers of the mass density field, and this bias is scale independent [35,36,42,85]. However, these conclusions are based on a configuration space argument known as the peak-background split—which establishes a relation between the sensitivity of the peak bias factors and the peak abundances on the peak height—whereas a Fourier space analysis suggests that the linear bias factor of peaks is the sum of two terms, one of which is k dependent [38,74]. In configuration space, this leads to scale dependence of the bias and stochasticity. The k dependence of the linear peak bias arises from the peak constraint, i.e., the fact that one must specify not only the value of the mass density field but also its first two derivatives to define a peak. Therefore, this is a model in which the bias depends on quantities other than the local density. Moreover, as mentioned above, the peak biasing is applied to the initial Gaussian density field so that the late time peak bias is modified by nonlinear evolution and associated stochasticity. In this regard, peaks exhibit a nontrivial velocity bias [75], which further complicates the nonlinear evolution.

In the peak model, both the constant and the k-dependent piece of the linear bias factor depend on peak height. As shown in [75], the scale-independent contribution can be derived from the peak-background split argument. In the first half of this paper, we demonstrate that the Fourier space approach also predicts constant and k-dependent contributions to the second and higher-order peak bias factors. We then show that the scale-independent parts of all of these nonlinear peak bias factors can also be derived from the peak-background split argument, thus generalizing the result of [75]. We go on to show how the peak-background split approach can be used to determine the scale-dependent part of the peak bias factors, first at linear order, and then for all nonlinear orders as well. This is particularly interesting because it illustrates how the peak-background split argument should be implemented if the abundance of the biased tracers (in this case, peaks) depends on quantities other than the local mass overdensity (in this case, the first and second derivatives of the mass density field).

As recognized in [74], the k dependence of the *first order* peak bias strongly amplifies the contrast of the baryon acoustic oscillation (or BAO, see [86], and references therein) in the correlation of initial density maxima. However, this calculation was performed for peaks identified in the initial conditions, so there was no clear connection with the clustering of dark matter halos and galaxies. This is also true of all the results presented in the first half of this paper. To remedy this problem, we show in the second half how the effects of the (nonlinear, nonlocal) gravitational evolution of density peaks can be incorporated in the peak model. This allows us to ascertain the extent to which the initial scale dependence of bias across the BAO survives at late times. Our analysis incorporates two main complications that are usually ignored in local bias schemes. Namely, peak biasing depends on more than just the value of the local density, and peaks exhibit a velocity bias which (in addition to merging) complicates analyses based on the continuity equation. Finally, we show that taking into account these effects is of more than academic interest: Our peaks' model provides a very good description of the scale dependence of the bias of massive halos in numerical simulations—halos that are expected to host the luminous red galaxies which are often targeted in BAO experiments.

The paper is organized as follows. Section II briefly reviews known results and introduces some useful

definitions. Section III focuses on the correlation of initial density peaks of a Gaussian random field. It is shown that the scale-dependent and scale-independent parts of the peak bias parameters can be derived from a peak-background split argument. Section IV considers the gravitational evolution of the peak correlation function in the Zel'dovich approximation. It is shown that, in addition to gravity mode coupling, the Lagrangian peak biasing can generate a significant scale-dependent bias across the baryonic acoustic feature at the collapse epoch. Measurements of bias at BAO scales from the clustering of massive halos are also presented and compared with the model. Section V summarizes our results. Technical details of the calculation can be found in Appendixes A and B.

II. DEFINITIONS, NOTATIONS, AND KNOWN RESULTS

We begin by introducing some definitions and reviewing known results about the clustering of density peaks in Gaussian random fields. Next, we derive the peak correlation at second order. This result will serve as input to the calculation of the evolved correlation of density peaks.

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 δ and its first and second derivatives, $\partial_i\delta$ and $\partial_i\partial_j\delta$. In this regard, it is convenient to introduce the normalized variables $\nu = \delta/\sigma_0$, $\eta_i \equiv \partial_i\delta/\sigma_1$, and $\zeta_{ij} \equiv \partial_i\partial_j\delta/\sigma_2$, where the σ_n are the spectral moments of the matter power spectrum,

$$\sigma_n^2(R_S, z_0) \equiv \frac{1}{2\pi^2}\int_0^\infty dk k^{2(n+1)} P_\delta(k, z_0) W(kR_S)^2. \quad (1)$$

$P_\delta(k, z_0)$ denotes the dimensionless power spectrum of the linear density field at redshift z_0, and W is a spherically symmetric smoothing kernel of length R_S introduced to ensure convergence of all spectral moments. A Gaussian filter will be adopted throughout this paper. We will use the notation $P_{\delta_S}(k, z_0)$ to denote $P_\delta(k, z_0)W(kR_S)^2$. The ratio σ_0/σ_1 is proportional to the typical separation between zero crossings of the density field [36]. For subsequent use, we also define the spectral parameters

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

which reflect the range over which $k^{2(n-1)}P_{\delta_S}(k, z_0)$ is large. We will also work with the scaled velocities $\mathrm{v}_i \equiv \mathrm{v}_{pi}/(aHf)$ and with the curvature $u = -\partial^2\delta/\sigma_2$. Here, $\mathrm{v}_{pi}(\mathbf{x})$ is the ith component of the (proper) peculiar velocity, $H \equiv d\ln a/dt$, $f \equiv d\ln D/d\ln a$ is the logarithmic derivative of the linear theory growth rate $D(z_0)$ and $\partial^2 = \partial^i\partial_i$ is the Laplacian. Note that v_i has dimensions of length.

The analogous quantities to σ_n^2 at nonzero separation are defined as follows:

$$\xi_\ell^{(n)}(R_S, r, z_0) = \frac{1}{2\pi^2}\int_0^\infty dk k^{2(n+1)} P_{\delta_S}(k, z_0) j_\ell(kr), \quad (3)$$

where $j_\ell(x)$ are spherical Bessel functions. As ℓ gets larger, these harmonic transforms become increasingly sensitive to small scale power. The auto and cross correlations of the fields $v_i(\mathbf{x})$, $\eta_i(\mathbf{x})$, $\nu(\mathbf{x})$, $u(\mathbf{x})$, and $\zeta_{ij}(\mathbf{x})$ can generally be decomposed into components with definite transformation properties under rotations. Reference [74] gives explicit expressions for the isotropic and homogeneous linear density field.

B. Peak biasing and two-point correlation function at the first order

Although density peaks form a well-behaved point process, the large scale asympotics of the two-point correlation $\xi_{\rm pk}(r, z)$ and line-of-sight mean streaming $v_{12}(\mathbf{r}, z)\cdot\hat{\mathbf{r}}$ of peaks of height ν and curvature u identified on a scale R_S in the initial Gaussian density field linearly extrapolated at redshift z_0 can be thought of as arising from the continuous, deterministic bias relation [74,75]

$$\delta_{\rm pk}(\nu, u, R_S, \mathbf{x}) = b_\nu\delta_S(\mathbf{x}, z_0) - b_\zeta\partial^2\delta_S(\mathbf{x}, z_0), \quad (4)$$

$$\mathbf{v}_{\rm pk}(R_S, \mathbf{x}, z_0) = \mathbf{v}_S(\mathbf{x}, z_0) - \frac{\sigma_0^2}{\sigma_1^2}\partial\delta_S(\mathbf{x}, z_0), \quad (5)$$

which is nonlocal owing to the smoothing of the mass distribution. Here, $\delta_{\rm pk}$ and $\mathbf{v}_{\rm pk}$ are the average peak overdensity and velocity, δ_S and $\mathbf{v}_S$ are the mass density and 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

$$\begin{aligned} b_\nu(\nu, u, R_S, z_0) &\equiv \frac{1}{\sigma_0(R_S, z_0)}\left(\frac{\nu - \gamma_1 u}{1-\gamma_1^2}\right), \\ b_\zeta(\nu, u, R_S, z_0) &\equiv \frac{1}{\sigma_2(R_S, z_0)}\left(\frac{u - \gamma_1\nu}{1-\gamma_1^2}\right). \end{aligned} \quad (6)$$

The bias coefficient b_ν is dimensionless, whereas b_ζ has units of (length)^2. In fact, b_ν is precisely the amplification factor found by the authors of Ref. [36] who neglected derivatives of the density correlation function (i.e., their analysis assumes $b_\zeta \equiv 0$). Unlike $\mathbf{v}_{\rm pk}$, $\delta_{\rm pk}$ does not depend on z_0 (as expected) because the redshift dependence of b_ν, b_ζ cancels the factor $D(z_0)$ coming from $\delta_S(\mathbf{x}, z_0)$. Note also that, if $b_\nu > 0$ the effective peak density $\delta_{\rm pk}(\mathbf{x})$ can be less than -1 in deep voids. However, this is not a problem because $\delta_{\rm pk}(\mathbf{x})$ is not an observable quantity (this is *not* a count-in-cell density).

In what follows, we will focus on the clustering of initial density peaks of significance ν, for which the first order bias parameters are

$$\bar{b}_\nu(\nu, R_S, z_0) \equiv \frac{1}{\sigma_0(R_S, z_0)}\left(\frac{\nu - \gamma_1 \bar{u}}{1 - \gamma_1^2}\right) \tag{7}$$

$$\bar{b}_\zeta(\nu, R_S, z_0) \equiv \frac{1}{\sigma_2(R_S, z_0)}\left(\frac{\bar{u} - \gamma_1 \nu}{1 - \gamma_1^2}\right). \tag{8}$$

Here, the overline denotes the averaging over the peak curvature, so that $\bar{u} \equiv \bar{u}(\nu, R_S)$ is the mean curvature of peaks of height ν on filtering scale R_S. It is convenient to define the quantity $b_{\rm spk}$ as the Fourier space multiplication by

$$b_{\rm spk}(k, z_0) = b_\nu + b_\zeta k^2, \tag{9}$$

where we have omitted the explicit dependence on ν, u, and R_S for brevity. Although $b_{\rm spk}(k, z_0)$ has the same functional form as Eq. (57) of [38], this author approximated density peaks by density extrema. Therefore, our coefficients agree with his expressions only in the limit $\nu \gg 1$, in which $b_\nu \to \nu/\sigma_0$ and $b_\zeta \to 0$. As we will see shortly, the product of $b_{\rm spk}(k, z_0)$ factors can be used to define spatial bias parameters at all orders. For peaks of significance ν, the first order biasing is equivalent to the Fourier space multiplication by $\tilde{\mathfrak{b}}_{\rm I}(k, z_0) \equiv \bar{b}_{\rm spk}(k, z_0)$, i.e.,

$$\tilde{\mathfrak{b}}_{\rm I}(k, z_0) \equiv \tilde{\mathfrak{b}}_{10} + \tilde{\mathfrak{b}}_{01} k^2 \quad \text{where } \tilde{\mathfrak{b}}_{10} \equiv \bar{b}_\nu, \qquad \tilde{\mathfrak{b}}_{01} \equiv \bar{b}_\zeta. \tag{10}$$

We emphasize that this result is exact: there are no higher powers such as k^4, etc. In $\tilde{\mathfrak{b}}_{ij}$, i and j count the number of factors of b_ν and b_ζ, respectively (our notation should not be confused with that of [87]). In Sec. III B, we will demonstrate that the $\tilde{\mathfrak{b}}_{i0}$ are the bias parameters in the local bias model. Equation (10) defines the first order bias for peaks of height ν. Notice that, in real space, $\tilde{\mathfrak{b}}_{\rm I}(k, z_0) = \tilde{\mathfrak{b}}_{10} - \tilde{\mathfrak{b}}_{01}\partial^2$ is a differential operator acting on fields and correlation functions. Hence, the first order average peak overabundance can also be rewritten $\delta_{\rm pk}(\nu, \mathbf{x}, z_0) = (\tilde{\mathfrak{b}}_{\rm I}\delta_S)(\mathbf{x}, z_0)$.

Using the peak bias (9), it is straightforward to show that the real space cross and auto power spectrum are

$$P^{(1)}_{{\rm pk},\delta}(\nu, R_S, k, z_0) = \tilde{\mathfrak{b}}_{\rm I}(k, z_0) P_\delta(k, z_0) W(kR_S) \tag{11}$$

$$P^{(1)}_{\rm pk}(\nu, R_S, k) = \tilde{\mathfrak{b}}^2_{\rm I}(k, z_0) P_\delta(k, z_0) W^2(kR_S). \tag{12}$$

The corresponding relations for the correlation functions are

$$\xi^{(1)}_{\rm pk}(\nu, R_S, r, z_0) = (\tilde{\mathfrak{b}}_{\rm I}\xi^{(0)}_0) = \tilde{\mathfrak{b}}_{10}\xi^{(0)}_0(R_S, r, z_0) + \tilde{\mathfrak{b}}_{01}\xi^{(1)}_0(R_S, r, z_0) \tag{13}$$

$$\xi^{(1)}_{\rm pk}(\nu, R_S, r) = (\tilde{\mathfrak{b}}^2_{\rm I}\xi^{(0)}_0) = \tilde{\mathfrak{b}}^2_{10}\xi^{(0)}_0(R_S, r, z_0) + 2\tilde{\mathfrak{b}}_{10}\tilde{\mathfrak{b}}_{01}\xi^{(1)}_0(R_S, r, z_0) + \tilde{\mathfrak{b}}^2_{01}\xi^{(2)}_0(R_S, r, z_0). \tag{14}$$

Note that the cross correlations with the linear density field $\delta(\mathbf{x}, z_0)$ depend explicitly on z_0. As shown in [74,75], these expressions 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 (12) and (14) are only valid at leading order, the cross correlation functions (11) and (13) are exact to all orders.

We emphasize that the biasing (5) is a mean bias relation that does not contain any information about stochasticity. Because of the discrete nature of density peaks however, one can expect that the average peak overabundance $\delta_{\rm pk}(\mathbf{x})$ in a cell centered at $\mathbf{x}$ generally be a random function of the underlying matter density (and its derivatives) in some neighborhood of that point. In fact, while the bias is deterministic in Fourier space, it is generally stochastic and scale dependent in configuration space [75].

C. Velocities

In what follows, we will be interested in the gravitational evolution of the correlation of initial density peaks for which the velocity field also matters. As can be seen from, e.g., the average bias relation (5), peaks locally move with the dark matter (since the gradient of the density vanishes at the position of a peak). However, the three-dimensional velocity dispersion of peaks, $\sigma^2_{\rm vpk} = \langle \mathbf{v}^2_{\rm pk} \rangle$, is smaller than the mass velocity dispersion σ^2_{-1} [36,88],

$$\sigma^2_{\rm vpk} = \sigma^2_{-1}(1 - \gamma_0^2), \tag{15}$$

because large scale flows are more likely to be directed towards peaks than to be oriented randomly. As recognized in [75], the k dependence of the first order peak bias $b_{\rm spk}(k)$ leads to a k dependence of the peak velocity statistics even though the peaks move with the dark matter flows. Taking the divergence of the peak velocity Eq. (5) and Fourier transforming, we find

$$\theta_{\rm pk}(R_S, \mathbf{k}, z_0) = \left(1 - \frac{\sigma_0^2}{\sigma_1^2}k^2\right) W(kR_S)\theta(\mathbf{k}, z_0) \equiv b_{\rm vpk}(k)\theta_S(\mathbf{k}, z_0), \tag{16}$$

where $\theta \equiv \nabla \cdot \mathbf{v}$ is the velocity divergence. This defines the *linear* velocity bias factor $b_{\rm vpk}(k)$ for peaks of significance ν and curvature u. Note that it does not depend on ν,

u nor on redshift and, for the highest peaks, it remains scale dependent even though the spatial bias $\tilde{b}_I(k, z_0)$ has no k dependence. Nonetheless, for notational consistency, we define

$$\tilde{b}_{vpk}(k) \equiv \bar{b}_{vpk}(k) = b_{vpk}(k) \tag{17}$$

as being the velocity bias of peaks of height ν.

D. Smoothing scale and peak height

To illustrate the key predictions of the peak formalism, we will present results for the two-point correlation of density peaks in a ΛCDM cosmology with $h = 0.7$, $\Omega_m = 0.279$, $\Omega_b = 0.0462$, $n_s = 0.96$, and normalization $\sigma_8 = 0.81$ consistent with the latest constraints from the cosmic microwave background [89]. The sound horizon at recombination is $r_0 \approx 105h^{-1}$ Mpc.

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 $\nu = \delta_{sc}(z_0)/\sigma_0(R_S)$ identified in the smoothed density field linearly extrapolated at z_0 are related to dark matter halos of mass $M_S \propto R_S^3$ collapsing at redshift z_0, where $\delta_{sc}(z_0)$ is the critical density for collapse in the spherical model [90,91], and we use a Gaussian filter to relate smoothing scale to mass. A more realistic treatment should include nonspherical collapse since the maxima of a Gaussian density field are inherently triaxial [92–95]. In the background cosmology we assume, the linear critical density for spherical collapse at $z_0 = 0.3$ is $\delta_{sc} \approx 1.681$. The Gaussian smoothing scale at which $\nu = 1$ is $R_{S_\star} \approx 0.8h^{-1}$ Mpc, which corresponds to a characteristic mass scale $M_{S_\star} \approx 6.2 \times 10^{11} M_\odot/h$.

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, $M_{S_\star}$, for which the peak model is expected to work best. For the sake of illustration, we will present results for $\nu = 2$ (2σ) and $\nu = 3$ (3σ) density peaks. At redshift $z_0 = 0.3$, this corresponds to a filtering length $R_S = 2.9h^{-1}$ Mpc and $R_S = 5.3h^{-1}$ Mpc or, equivalently, a mass scale $M_S = 3.0 \times 10^{13} M_\odot/h$ and $1.6 \times 10^{14} M_\odot/h$. To help set scales in the discussion which follows, the associated values of $(\gamma_1, \sigma_0/\sigma_1, \tilde{b}_{10}, \tilde{b}_{01})$ are $(0.65, 3.7, 0.8, 21.1h^{-2}\ \mathrm{Mpc}^2)$ and $(0.68, 6.2, 3.5, 72.4h^{-2}\ \mathrm{Mpc}^2)$, respectively, (note that bias factors here are Lagrangian ones). The three-dimensional velocity dispersion of these peaks is $\sigma_{-1}^2(1 - \gamma_0^2)$: for our two smoothing scales, this corresponds to $(7.75h^{-1}\ \mathrm{Mpc})^2$ and $(7.03h^{-1}\ \mathrm{Mpc})^2$ (recall that our velocities are in units of $aHf \approx 58\,\mathrm{km\,s^{-1}}\,h\,\mathrm{Mpc}^{-1}$ at $z = 0.3$, so dispersions have dimensions of (length)^2).

III. CORRELATION OF INITIAL DENSITY PEAKS AT SECOND ORDER

A. The general formula

Correlations of density maxima can be evaluated using the Kac-Rice formula [96,97]. In this approach, $\eta_i(\mathbf{x})$ is Taylor expanded around the position $\mathbf{x}_{pk}$ of a local maximum. The number density of peaks of height ν' at position $\mathbf{x}$ in the smoothed density field δ_S reads as (we drop the subscript S in the right-hand side for notational convenience)

$$n_{pk}(\nu', R_S, \mathbf{x}) \equiv \frac{3^{3/2}}{R_1^3} |\det\zeta(\mathbf{x})| \delta^{(3)}[\boldsymbol{\eta}(\mathbf{x})]\theta[\lambda_3(\mathbf{x})]\delta[\nu(\mathbf{x}) - \nu'] \tag{18}$$

where

$$R_1 \equiv \sqrt{3}\frac{\sigma_1}{\sigma_2} \tag{19}$$

is the characteristic radius of a peak. Note that Eq. (18) is independent of the redshift z_0 for peaks at fixed ν. The three-dimensional Dirac delta $\delta^{(3)}(\boldsymbol{\eta})$ ensures that all extrema are included. The product of the theta function $\theta(\lambda_3)$, where λ_3 is the lowest eigenvalue of the shear tensor ζ_{ij}, and the Dirac delta $\delta(\nu - \nu')$ further restrict the set to density maxima with specific height ν'. The two-point correlation function for maxima of a given significance separated by a distance $r = |\mathbf{r}| = |\mathbf{x}_2 - \mathbf{x}_1|$ thus is

$$1 + \xi_{pk}(\nu, R_S, r) = \frac{\langle n_{pk}(\nu, R_S, \mathbf{x}_1) n_{pk}(\nu, R_S, \mathbf{x}_2)\rangle}{\bar{n}_{pk}^2(\nu, R_S)}, \tag{20}$$

where $\bar{n}_{pk}(\nu, R_S)$ is the differential average number density of peaks of height ν on filtering scale R_S [36],

$$\bar{n}_{pk}(\nu, R_S) = \frac{1}{(2\pi)^2 R_1^3} e^{-\nu^2/2} G_0^{(1)}(\gamma_1, \gamma_1\nu). \tag{21}$$

Note that it does not depend on z_0 (or, equivalently, on the amplitude of density fluctuations) at fixed ν. The function $G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)$ is defined in Eq. (A60). For the 2σ and 3σ density peaks considered here, the mean abundance is $\bar{n}_{pk} = 9.0 \times 10^{-5}$ and $4.8 \times 10^{-6} h^3\ \mathrm{Mpc}^{-3}$, respectively. While the calculation of Eq. (20) at first order in the mass correlation and its derivatives is rather straightforward [74] [this is Eq. (14)], at second order it is quite involved. The main steps are detailed in Appendix A. Fortunately, most of the terms nicely combine together, and the final result can be recast into the compact form

$$\xi_{\rm pk}(\nu, R_S, r) = (\tilde{\mathfrak{b}}_{\rm I}^2\xi_0^{(0)}) + \frac{1}{2}(\xi_0^{(0)}\tilde{\mathfrak{b}}_{\rm II}^2\xi_0^{(0)}) - \frac{3}{\sigma_1^2}(\xi_1^{(1/2)}\tilde{\mathfrak{b}}_{\rm II}\xi_1^{(1/2)}) - \frac{5}{\sigma_2^2}(\xi_2^{(1)}\tilde{\mathfrak{b}}_{\rm II}\xi_2^{(1)})\left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}\right)$$
$$+ \frac{5}{2\sigma_2^4}\left[(\xi_0^{(2)})^2 + \frac{10}{7}(\xi_2^{(2)})^2 + \frac{18}{7}(\xi_4^{(2)})^2\right]\left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}\right)^2 + \frac{3}{2\sigma_1^4}[(\xi_0^{(1)})^2 + 2(\xi_2^{(1)})^2]$$
$$+ \frac{3}{\sigma_1^2\sigma_2^2}[3(\xi_3^{(3/2)})^2 + 2(\xi_1^{(3/2)})^2]. \tag{22}$$

In the right-hand side of Eq. (22), all the correlations are function of R_S, r, and z_0. More precisely, the first line contains terms involving first and second order peak bias parameters $\tilde{\mathfrak{b}}_{\rm I}$ and $\tilde{\mathfrak{b}}_{\rm II}$, the second line has a ν dependence through the function $1 + (2/5)\partial_\alpha \ln G_0^\alpha(\gamma_1, \gamma_1\nu)|_{\alpha=1}$ (which is displayed in Fig. 9), and the last two terms depend on the separation r (and R_S) only. Note that this expression exhibits not only terms quadratic in bias parameters but, unlike standard local bias (Eulerian or Lagrangian), also terms linear in them. These terms involve derivatives $\xi_{j\neq 0}^{(n)}$ of the linear mass correlation $\xi_0^{(0)}$ that vanish at zero lag. They arise because the peak correlation depends also on the statistical properties of η_i and ζ_{ij}.

In analogy with $\tilde{\mathfrak{b}}_{\rm I}(k, z_0)$, the action of the second order peak bias $\tilde{\mathfrak{b}}_{\rm II}$ is defined as the Fourier space multiplication by

$$\tilde{\mathfrak{b}}_{\rm II}(q_1, q_2, z_0) \equiv \overline{b_{\rm spk}(q_1, z_0)b_{\rm spk}(q_2, z_0)} - (1-\gamma_1^2)^{-1}\left[\frac{1}{\sigma_0^2} + \frac{(q_1q_2)^2}{\sigma_2^2} - \frac{\gamma_1^2}{\sigma_1^2}(q_1^2 + q_2^2)\right]$$
$$\equiv \tilde{\mathfrak{b}}_{20} + \tilde{\mathfrak{b}}_{11}(q_1^2 + q_2^2) + \tilde{\mathfrak{b}}_{02}q_1^2q_2^2, \tag{23}$$

where q_1 and q_2 are wave modes and the coefficients $\tilde{\mathfrak{b}}_{20}$, $\tilde{\mathfrak{b}}_{11}$, and $\tilde{\mathfrak{b}}_{02}$ describing the peak bias at second order are

$$\tilde{\mathfrak{b}}_{20}(\nu, R_S, z_0) \equiv \bar{b}_{\nu\nu} - \frac{1}{\sigma_0^2(1-\gamma_1^2)}$$
$$= \frac{1}{\sigma_0^2}\left[\frac{\nu^2 - 2\gamma_1\nu\bar{u} + \gamma_1^2\bar{u}^2}{(1-\gamma_1^2)^2} - \frac{1}{(1-\gamma_1^2)}\right] \tag{24}$$

$$\tilde{\mathfrak{b}}_{11}(\nu, R_S, z_0) \equiv \bar{b}_{\nu\zeta} + \frac{\gamma_1^2}{\sigma_1^2(1-\gamma_1^2)}$$
$$= \frac{1}{\sigma_0\sigma_2}\left[\frac{(1+\gamma_1^2)\nu\bar{u} - \gamma_1[\nu^2 + \bar{u}^2]}{(1-\gamma_1^2)^2} + \frac{\gamma_1}{(1-\gamma_1^2)}\right] \tag{25}$$

$$\tilde{\mathfrak{b}}_{02}(\nu, R_S, z_0) \equiv \bar{b}_{\zeta\zeta} - \frac{1}{\sigma_2^2(1-\gamma_1^2)}$$
$$= \frac{1}{\sigma_2^2}\left[\frac{\bar{u}^2 - 2\gamma_1\nu\bar{u} + \gamma_1^2\nu^2}{(1-\gamma_1^2)^2} - \frac{1}{(1-\gamma_1^2)}\right]. \tag{26}$$

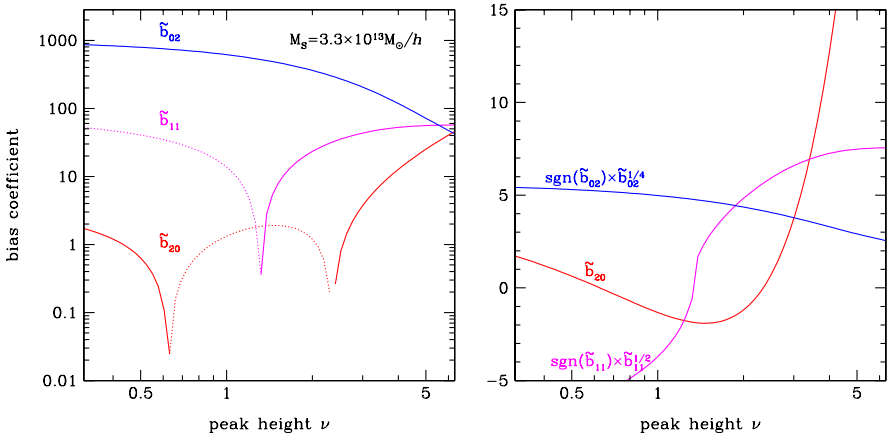

FIG. 1 (color online). Left panel: Lagrangian bias coefficients characterizing the second order peak bias $\tilde{\mathfrak{b}}_{\rm II}(q_1, q_2)$, Eqs. (24)–(26), as a function of peak height for a filtering radius $R_S = 2.9h^{-1}$ Mpc or, equivalently, a mass scale $M_S = 3 \times 10^{13}{\rm M}_\odot/h$. The shape parameter is $\gamma_1 \approx 0.65$. For the 2σ peaks considered in subsequent illustrations, $\tilde{\mathfrak{b}}_{20}$ is negative, $\tilde{\mathfrak{b}}_{20} \approx -1.2$. Right panel : The second and fourth root $\tilde{\mathfrak{b}}_{11}^{1/2}$ and $\tilde{\mathfrak{b}}_{02}^{1/4}$ define a characteristic scale below which the scale dependence of $\tilde{\mathfrak{b}}_{\rm II}$ is large. In the limit $\nu \to \infty$, $\tilde{\mathfrak{b}}_{02}$ becomes negative and converges towards $-\sigma_2^{-2}(1-\gamma_1^2)^{-1}$, whereas $\tilde{\mathfrak{b}}_{11}$ asymptotes to the constant $(\gamma_1/\sigma_1)^2(1-\gamma_1^2)^{-1}$. Note that, in contrast to $\tilde{\mathfrak{b}}_{11}^{1/2}$ and $\tilde{\mathfrak{b}}_{02}^{1/4}$ that have units of length, $\tilde{\mathfrak{b}}_{20}$ is dimensionless.

See Fig. 1 for the relative size of these contributions as a function of peak height. As shown in Appendix A, $\bar{b}_{\nu\nu}$, $\bar{b}_{\nu\zeta}$, and $\bar{b}_{\zeta\zeta}$ arise upon averaging the products $b_\nu b_\nu$, $b_\nu b_\zeta$, and $b_\zeta b_\zeta$ over the peak curvature. In this respect, $\bar{u}^n \equiv G_n^{(1)}(\gamma_1, \gamma_1\nu)/G_0^{(1)}(\gamma_1, \gamma_1\nu)$ is the nth moment of the peak curvature at a given significance ν. For the sake of completeness, $\tilde{\mathfrak{b}}_{\rm II}^2$ acts on the functions $\xi_{\ell_1}^{(n_1)}(r)$ and $\xi_{\ell_2}^{(n_2)}(r)$ according to

$$(\xi_{\ell_1}^{(n_1)}\tilde{\mathfrak{b}}_{\rm II}^2\xi_{\ell_2}^{(n_2)}) \equiv \frac{1}{4\pi^4}\int_0^\infty dq_1\int_0^\infty dq_2 q_1^{2(n_1+1)}q_2^{2(n_2+1)}\tilde{\mathfrak{b}}_{\rm II}^2 \times (q_1, q_2, z_0)P_{\delta_S}(q_1, z_0)P_{\delta_S}(q_2, z_0)j_{\ell_1} \times (q_1 r)j_{\ell_2}(q_2 r). \quad (27)$$

When $\ell_1 = \ell_2 = 0$, the real space counterpart of $\tilde{\mathfrak{b}}_{\rm II}(q_1, q_2, z_0)$ is readily obtained by making the replacement $q^2 \to -\partial^2$ [which reflects the fact that $\xi_0^{(n)}(R_S, r, z_0)$ is a solution to the Helmholtz equation $(\partial^2 + k^2)\xi(r) = 0$]. Equation (22) is the main result of this section. We note that [72] also computed second order corrections to the peak correlation $\xi_{\rm pk}(\nu, R_S, r)$ for which, however, they did not provide any explicit expression.

Before illustrating the impact of the second order terms on the correlation of initial density maxima, we remark that, although the calculation of the peak correlation at third order is very involved, the contribution proportional to $(\xi_0^{(0)})^3$ can be derived relatively easily. We find

$$\frac{1}{6}\left[\bar{b}_{\nu\nu\nu} - \frac{3\bar{b}_\nu}{\sigma_0^2(1-\gamma_1^2)}\right]^2(\xi_0^{(0)})^3 \equiv \frac{1}{6}\tilde{\mathfrak{b}}_{30}^2(\xi_0^{(0)})^3, \quad (28)$$

where the third order coefficient $\bar{b}_{\nu\nu\nu}$ is defined as

$$\bar{b}_{\nu\nu\nu}(\nu, R_S, z_0) = \frac{\nu^3 - 3\gamma_1\nu^2\bar{u} + 3\gamma_1^2\nu\bar{u}^2 - \gamma_1^3\bar{u}^3}{\sigma_0^3(1-\gamma_1^2)^3}. \quad (29)$$

Thus, up to third order, the peak correlation may be cast into the form

$$\xi_{\rm pk}(\nu, R_S, r) \approx \tilde{\mathfrak{b}}_{10}^2\xi_0^{(0)} + \tfrac{1}{2}\tilde{\mathfrak{b}}_{20}^2(\xi_0^{(0)})^2 + \tfrac{1}{6}\tilde{\mathfrak{b}}_{30}^2(\xi_0^{(0)})^3 + \text{additional terms}, \quad (30)$$

where the missing terms, while of the same order as the ones we display, have a more complicated structure [see Eq. (22) for second order contributions]. In the limit $\nu \gg 1$, the scale-independent pieces $\tilde{\mathfrak{b}}_{10}$, $\tilde{\mathfrak{b}}_{20}$, and $\tilde{\mathfrak{b}}_{30}$ to the bias asymptote to the values $\tilde{\mathfrak{b}}_{10} \to \nu/\sigma_0$, $\tilde{\mathfrak{b}}_{20} \to (\nu/\sigma_0)^2$, and $\tilde{\mathfrak{b}}_{30} \to (\nu/\sigma_0)^3$ obtained in the high level excursion set approximation [35]. We will see shortly that these bias factors are indeed equal to the peak-background split biases derived from the average peak abundance Eq. (21).

In Fig. 1, the second order Lagrangian biases $\tilde{\mathfrak{b}}_{20}$, $\tilde{\mathfrak{b}}_{11}$, and $\tilde{\mathfrak{b}}_{02}$ are shown as a function of the peak height for the mass scale $M_S = 3.3 \times 10^{13}{\rm M}_\odot/h$ (left panel). The second and fourth root $\tilde{\mathfrak{b}}_{11}^{1/2}$ and $\tilde{\mathfrak{b}}_{02}^{1/4}$, respectively, define a characteristic comoving scale below which the corresponding scale-dependent terms $\tilde{\mathfrak{b}}_{11}q^2$ and $\tilde{\mathfrak{b}}_{02}q^4$ are large. In the limit $\nu \to \infty$, the scale-independent piece $\tilde{\mathfrak{b}}_{20}$ increasingly dominates whereas $\tilde{\mathfrak{b}}_{11}$ and $\tilde{\mathfrak{b}}_{02}$ tend towards the constant value $(\gamma_1/\sigma_1)^2(1-\gamma_1^2)^{-1}$ and $-\sigma_2^{-2}(1-\gamma_1^2)^{-1}$. For a realistic threshold height $\nu < 4$ however, the scale-dependent contributions cannot be neglected since $|\tilde{\mathfrak{b}}_{11}|$ and/or $|\tilde{\mathfrak{b}}_{02}|$ are typically much larger than $|\tilde{\mathfrak{b}}_{20}|$. Although the exact value of the second order biases somewhat changes with the mass scale M_S, their overall behavior varies little over the range $M_S \sim 10^{12}$–$10^{14}{\rm M}_\odot/h$ as γ_1 weakly depends on R_S. Therefore, our conclusions hold regardless of the exact amount of smoothing. It should also be noted that if one wishes to associate these bias factors to halos of mass $M_S = 3.3 \times 10^{13}{\rm M}_\odot/h$, then the variation with ν is in fact a variation with redshift.

The correlation $\xi_{\rm pk}(\nu, R_S, r)$ is shown in Fig. 2 for the 2σ and 3σ initial density peaks collapsing at redshift $z_0 = 0.3$. The solid (green) curve represents the first order term $\tilde{\mathfrak{b}}_{\rm I}^2\xi_0^{(0)}$ [Eq. (14)] while the long-dashed-dotted curve is the full second order correlation [Eq. (22)]. We have also plotted the second order contributions quadratic in $\tilde{\mathfrak{b}}_{\rm II}$, linear in $\tilde{\mathfrak{b}}_{\rm II}$ and independent of $\tilde{\mathfrak{b}}_{\rm II}$ separately. They are shown as the short-dashed-dotted, short-dashed, and long-dashed curve, respectively. Notice that $(\tilde{\mathfrak{b}}_{20}, \tilde{\mathfrak{b}}_{11}, \tilde{\mathfrak{b}}_{02}) = (-1.2, 23, 363)$ and $(7.8, 285, 3927)$ for the low and high threshold, respectively. For the 3σ peaks, the term linear in $\tilde{\mathfrak{b}}_{\rm II}$ is negative over the range of distances considered and, thus, appears as a dotted line. In fact, the piece linear in $\tilde{\mathfrak{b}}_{\rm II}$ is the only negative contribution at small separations but it vanishes at zero lag. Since the true peak correlation rapidly converges to -1 for $r < R_1$ [as shown by [70], the small scale behavior of $\xi_{\rm pk}$ is dominated by an exponential term $\exp(-R_1^2/r^2)$], small scale exclusion should manifest itself in higher-order terms, but it is beyond the scope of this paper to calculate them. We can also observe that the correlation of 2σ peaks is negative on scales $r \sim 5$–$10h^{-1}$ Mpc. However, this is likely an artifact of truncating the expansion at second order in the correlations $\xi_\ell^{(n)}$.

Figure 3 focuses on the baryon acoustic oscillation. At separation $r > 80h^{-1}$ Mpc, second order corrections are negligibly small, so that $\xi_{\rm pk}(\nu, R_S, r)$ is given by Eq. (14) with an accuracy better than 1%. For comparison, we also plot the first order correlation $\tilde{\mathfrak{b}}_{10}^2\xi_\delta$ arising in a local biasing scheme with the same value of $\tilde{\mathfrak{b}}_{10}$. It is important to note that ξ_δ is the correlation of the *unsmoothed*, linear mass density field (in practice we use $R_S = 0.1h^{-1}$ Mpc). It is quite remarkable how the "sombrero-"shaped terms $2\tilde{\mathfrak{b}}_{10}\tilde{\mathfrak{b}}_{01}\xi_0^{(1)}$ and $\tilde{\mathfrak{b}}_{01}^2\xi_0^{(2)}$ restore the contrast of the baryonic feature otherwise smeared out by the large filtering (recall

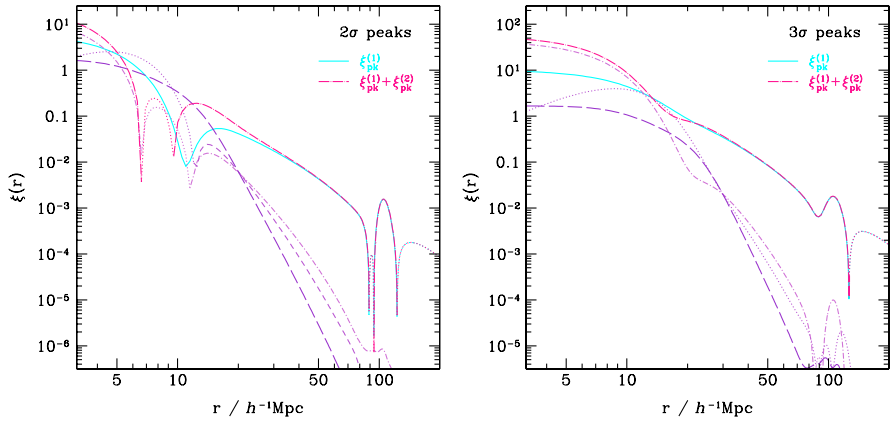

FIG. 2 (color online). The correlation of initial density peaks at the second order is shown as the long dashed-dotted (magenta) curve for 2σ (left panel) and 3σ (right panel) density peaks collapsing at redshift $z_0 = 0.3$ according to the spherical collapse prescription. For the Gaussian filter used in this paper, this corresponds to a mass scale $M_S = 3 \times 10^{13}$ and $2 \times 10^{14} \mathrm{M}_\odot/h$, respectively. The individual contributions appearing in Eq. (22) are shown separately. Namely, the solid (cyan) curve is the first order contribution $\tilde{\mathrm{b}}_{\mathrm{I}}^2 \xi_0^{(0)}$, whereas the second order term quadratic in $\tilde{\mathrm{b}}_{\mathrm{II}}$, linear in $\tilde{\mathrm{b}}_{\mathrm{II}}$ and independent of $\tilde{\mathrm{b}}_{\mathrm{II}}$ are shown as the short dashed-dotted, short-dashed, and long-dashed curves, respectively. A dotted line indicates negative values.

that $R_S = 2.9$ and $5.3h^{-1}$ Mpc for the 2σ and 3σ peaks, respectively). The BAO in the peak correlations is even sharpened relative to the BAO in ξ_δ. A thorough discussion of this effect can be found in [74]. In Sec. IV, we will see that although most of the initial enhancement of the BAO contrast is smeared out by the gravitational motions of the peaks, some of it survives at the epoch of collapse.

B. A peak-background split derivation of the peak bias factors

The peak-background split [33,36,37,98] is a heuristic argument that furnishes another way to derive the large scale bias of density peaks. This approach is quite different from ours because it is based on number counts in configuration space and, thus, does not make reference to the bias in Fourier space.

There are two ways in which the peak-background split is implemented. In the first [36,37,42], the Nth order bias parameter $b_N(\nu)$ is related to the Nth order derivative of the differential number density $\bar{n}(\nu)$ of virialized objects according to

$$b_N(\nu, z_0) \equiv \left(-\frac{1}{\sigma_0(z_0)}\right)^N \bar{n}(\nu)^{-1} \frac{\partial^N[\bar{n}(\nu)]}{\partial \nu^N}, \qquad (31)$$

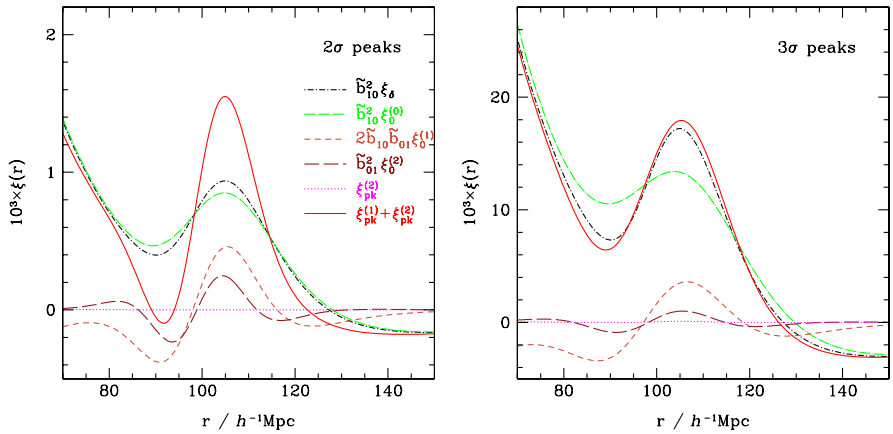

FIG. 3 (color online). A comparison between the initial *unsmoothed* density correlation $\xi(r, z_0)$ (black, dotted-dashed) and the initial peak correlation $\xi_{\mathrm{pk}}(\nu, R_S, r)$ (red, solid) around the BAO. To obtain the peak correlation, the density field was smoothed with a Gaussian filter on mass scale $M_S = 3 \times 10^{13}$ (left panel) and $2 \times 10^{14} \mathrm{M}_\odot/h$ (right panel). The dotted-long-dashed, short-dashed, and long-dashed curves represent the individual contributions $\tilde{\mathrm{b}}_{10}^2 \xi_0^{(0)}$, $2\tilde{\mathrm{b}}_{10}\tilde{\mathrm{b}}_{01}\xi_0^{(1)}$ and $\tilde{\mathrm{b}}_{01}^2 \xi_0^{(2)}$ to the first order peak correlation [Eq. (14)]. A nonzero $\tilde{\mathrm{b}}_{01}$ restores, and even amplifies the acoustic peak otherwise smeared out upon filtering the mass density field. The dotted curve indicates the second order correction to the peak correlation. Results are shown for the CDM transfer function considered in this paper.

with the important caveat that the mass function is universal (i.e., it depends on ν solely). For density peaks, on setting $\bar{n} = \bar{n}_{\rm pk}(\nu, R_S)$ and performing the derivatives with respect to ν at fixed smoothing radius R_S, one obtains [42]

$$b_N(\nu, R_S, z_0) \equiv \left(-\frac{1}{\sigma_0(R_S, z_0)}\right)^N \bar{n}_{\rm pk}(\nu, R_S)^{-1} \times \frac{\partial^N[\bar{n}_{\rm pk}(\nu, R_S)]}{\partial \nu^N}. \quad (32)$$

As noted by [75], the first order peak-background split bias is

$$b_{\rm I}(\nu, R_S, z_0) = \tilde{\mathfrak{b}}_{10}(\nu, R_S, z_0). \quad (33)$$

This means that the large scale, constant, and deterministic bias factor returned by the peak-background split argument is exactly the same as in our approach, when we are on large enough scales that the k dependence associated with the $\tilde{\mathfrak{b}}_{01}k^2$ term can be ignored. It turns out that Eq. (33) generalizes as follows: higher-order derivatives of the peak number density (21) with respect to ν (which are reported in [42]) result in the large scale, k-independent peak bias coefficients

$$b_{\rm II}(\nu, R_S, z_0) = \tilde{\mathfrak{b}}_{20}(\nu, R_S, z_0), \quad b_{\rm III}(\nu, R_S, z_0) = \tilde{\mathfrak{b}}_{30}(\nu, R_S, z_0), \text{ etc.} \quad (34)$$

However, derivatives of Eq. (21) cannot produce the k-dependent bias terms like $\tilde{\mathfrak{b}}_{01}$, $\tilde{\mathfrak{b}}_{11}$ etc., which arise owing to the constraints imposed by derivatives of the mass density field.

Therefore, we will now consider the second implementation of the peak-background split [33,98] in which the dependence of the mass function on the overdensity of the background is derived explicitly. The ratio of this conditional mass function to the universal one is then expanded in powers of the background density. The bias factors are the coefficients of this expansion. We will demonstrate below that this is the correct approach to recover the scale or k dependence of the peak bias parameters.

The key quantity is the average number density of peaks identified on scale R_S as a function of the overdensity δ_B defined on another smoothing scale R_B (we use the subscript B because we are mainly interested in the regime in which the scale R_B of the background satisfies $R_B \gg R_S$). This conditional peak number density is

$$\bar{n}_{\rm pk}(\nu, R_S|\delta_B, R_B) = \frac{G_0^{(0)}(\tilde{\gamma}_1, \tilde{\gamma}_1\tilde{\nu})}{(2\pi)^{3/2}R_1^3} \times \frac{\exp[-(\nu - \epsilon\nu_B)^2/2(1-\epsilon^2)]}{\sqrt{2\pi(1-\epsilon^2)}}, \quad (35)$$

where

$$\nu_B \equiv \frac{\delta_B}{\sigma_{0B}}, \qquad \langle \nu\nu_B\rangle \equiv \epsilon = \frac{\sigma_{0\times}^2}{\sigma_{0S}\sigma_{0B}},$$
$$\langle u\nu_B\rangle \equiv \gamma_1\epsilon r, \qquad r \equiv \frac{\langle k^2\rangle_\times}{\langle k^2\rangle_S} = \frac{\sigma_{1\times}^2/\sigma_{1S}^2}{\sigma_{0\times}^2/\sigma_{0S}^2}, \quad (36)$$

$$\langle u|\nu, \nu_B\rangle \equiv \tilde{\gamma}_1\tilde{\nu} = \gamma_1\nu\left(\frac{1-\epsilon^2 r}{1-\epsilon^2}\right) - \gamma_1\left(\frac{1-r}{1-\epsilon^2}\right)\epsilon\nu_B,$$
$$\mathrm{Var}(u|\nu, \nu_B) \equiv 1 - \tilde{\gamma}_1^2 \equiv 1 - \gamma_1^2\left[1 + \epsilon^2\frac{(1-r)^2}{1-\epsilon^2}\right], \quad (37)$$

with $\sigma_{nS} \equiv \sigma_n(R_S, z_0)$, $\sigma_{nB} \equiv \sigma_n(R_B, z_0)$, and we have defined

$$\sigma_{n\times}^2 \equiv \frac{1}{2\pi^2}\int_0^\infty dk k^{2(n+1)}P_\delta(k)W(kR_S)W(kR_B) \quad (38)$$

[see Eq. (E5) of [36]]. Here, the $\times$ denotes the splitting of smoothing scales, i.e., one filter is of size R_S, the other of size R_B. We have deliberately written r as an average of k^2 to emphasize that we naively expect it to give rise to the k^2 dependence of peak bias. This will eventually be proven correct. In addition, note that $\epsilon\nu_B = (\delta_B/\sigma_{0S})(\sigma_{0\times}^2/\sigma_{0B}^2)$. In what follows, it will be convenient to define $\Sigma_{\times B}^2 \equiv \sigma_{0\times}^2/\sigma_{0B}^2$. When $R_B \gg R_S$, this ratio is of order unity (there is a form factor that depends on the shape of the smoothing filter).

Notice that the integral of Eq. (35) over all δ_B gives the unconditional number density $\bar{n}_{\rm pk}(\nu, R_S)$ of Eq. (21). The peak-background split expands the ratio $\bar{n}_{\rm pk}(\nu, R_S|\delta_B, R_B)/\bar{n}_{\rm pk}(\nu, R_S)$ in powers of δ_B. This ratio is then interpreted as representing the average overabundance of peaks in regions which have mass overdensity δ_B although, strictly speaking, it is a statement about cells of overdensity δ_B that have a peak at their center. Therefore, it is *not* a statement about randomly placed cells, even though, as we discuss below, it is often treated as such.

If we set $r \to 0$ and $\epsilon \to 0$ then the coefficient of the term of order δ_B gives $\tilde{\mathfrak{b}}_{10}$, that of order gives δ_B^2 gives $\tilde{\mathfrak{b}}_{20}$, etc. Note that in this limit, $\tilde{\gamma}_1 \to \gamma_1$ and $\tilde{\gamma}_1\tilde{\nu} \to \gamma_1\nu(1 - \delta_B/\delta_S)$, so that expanding Eq. (35) in powers of δ_B will be the same as differentiating Eq. (21) with respect to δ_S. These derivatives result in the large scale, k-independent, peak bias factors we have been denoting as $\tilde{\mathfrak{b}}_{10}$, $\tilde{\mathfrak{b}}_{20}$, etc. As noted above, however, these derivatives cannot produce the k-dependent bias terms like $\tilde{\mathfrak{b}}_{01}$, $\tilde{\mathfrak{b}}_{11}$, etc.

Setting $\epsilon \to 0$ but keeping the r dependence means that $\tilde{\gamma}_1 \to \gamma_1$ but $\tilde{\gamma}_1\tilde{\nu} \to \gamma_1\nu[1 - (1-r)(\delta_B/\delta_S)\Sigma_{\times B}^2]$. As a result, derivatives of $G_0^{(0)}$ with respect to δ_B will introduce terms which depend on r; these are terms which could *not* have been obtained by differentiating the unconditional mass function. For example, to first order in δ_B,

$$G_0^{(0)}(\tilde{\gamma}_1, \tilde{\gamma}_1\tilde{\nu}) \approx G_0^{(0)}(\gamma_1, \gamma_1\nu)\Big[1 - (1-r)\frac{(\bar{u} - \gamma_1\nu)\gamma_1\nu}{1-\gamma_1^2} \times \frac{\delta_B}{\delta_S}\Sigma_{\times B}^2\Big] \tag{39}$$

so that Eq. (35) becomes

$$\bar{n}_{\rm pk}(\nu, R_S|\delta_B, R_B) \approx \bar{n}_{\rm pk}(\nu, R_S)\Big[1 + \nu\frac{\delta_B}{\sigma_{0S}}\Sigma_{\times B}^2 - (1-r)\frac{(\bar{u} - \gamma_1\nu)\gamma_1\nu}{1-\gamma_1^2}\frac{\delta_B}{\delta_S}\Sigma_{\times B}^2\Big]. \tag{40}$$

Hence, in this limit,

$$\begin{aligned}\langle\delta_{\rm pk}|\delta_B\rangle &\equiv \frac{\bar{n}_{\rm pk}(\nu, R_S|\delta_B, R_B)}{\bar{n}_{\rm pk}(\nu, R_S)} - 1 \\ &\approx \frac{\delta_B}{\sigma_{0S}}\Sigma_{\times B}^2\left(\nu - \frac{(\gamma_1\bar{u} - \gamma_1^2\nu)}{1-\gamma_1^2}\right) \\ &\quad + r\gamma_1\nu\frac{\delta_B}{\delta_S}\Sigma_{\times B}^2\frac{(\bar{u} - \gamma_1\nu)}{1-\gamma_1^2} \\ &= \frac{\sigma_{0\times}^2}{\sigma_{0B}^2}\frac{\delta_B}{\sigma_{0S}}\left(\frac{\nu - \gamma_1\bar{u}}{1-\gamma_1^2}\right) + \frac{\sigma_{1\times}^2}{\sigma_{0B}^2}\frac{\delta_B}{\sigma_{2S}}\left(\frac{\bar{u} - \gamma_1\nu}{1-\gamma_1^2}\right) \\ &= \frac{\sigma_{0\times}^2}{\sigma_{0B}^2}\tilde{\mathfrak{b}}_{10}\delta_B + \frac{\sigma_{1\times}^2}{\sigma_{0B}^2}\tilde{\mathfrak{b}}_{01}\delta_B,\end{aligned} \tag{41}$$

where $\tilde{\mathfrak{b}}_{10}$ and $\tilde{\mathfrak{b}}_{01}$ were defined in Eq. (10). Therefore, the cross correlation between the average overdensity of peaks defined on scale R_S and the mass overdensity on scale R_B is

$$\langle\delta_{\rm pk}\delta_B\rangle(R_B) = \sigma_{0\times}^2\tilde{\mathfrak{b}}_{10} + \sigma_{1\times}^2\tilde{\mathfrak{b}}_{01} \quad \text{on large scales } R_B \gg R_S. \tag{42}$$

Note that this final expression is the Fourier transform of $(\tilde{\mathfrak{b}}_{10} + \tilde{\mathfrak{b}}_{01}k^2)P_\delta(k, z_0)W(kR_S)W(kR_B)$. Thus, we have shown explicitly that our implementation of the peak-background split argument has produced the same linear, scale-dependent bias as the Fourier space argument.

At this point, one may be worried there is an inconsistency in the above expansion since $\epsilon \to 0$ necessarily implies $r \to 0$. For example, for Gaussian smoothing of a power law spectrum $P(k) \propto k^n$, we find $r \simeq 2(R_S/R_B)^2$ and $\epsilon \simeq (2R_S/R_B)^{3+n/2}$ at sufficiently large R_B [36]. For a realistic spectral index $-3 < n < 3$, ϵ and, obviously, r always vanish in the limit $R_B \to \infty$. In fact, there is no problem here since one can, at least formally, treat ϵ and r as two independent, small parameters in addition to δ_B. Hence, setting $r \to 0$ at a fixed value of δ_B and ϵ for instance corresponds to retaining the change in the small scale density δ_S while neglecting the change in the curvature u induced by the background wave δ_B.

The higher-order bias factors can be derived in an analogous way. Each term of order δ_B^N will include terms of order r^m with $m \leq N$. Terms proportional to r give $\tilde{\mathfrak{b}}_{01}$, $\tilde{\mathfrak{b}}_{11}$, etc.; terms proportional to r^2 give $\tilde{\mathfrak{b}}_{02}$, $\tilde{\mathfrak{b}}_{12}$, etc. Thus, our analysis provides a simple way of determining all the additional k-dependent higher-order bias terms which arise in the peaks' model. The k-dependent polynomials associated to these bias factors are determined from symmetry considerations. For instance, $\tilde{\mathfrak{b}}_{12}$ multiplies the polynomial $q_1^2q_2^2 + q_1^2q_3^2 + q_2^2q_3^2$. As a general rule, the bias factor $\tilde{\mathfrak{b}}_{ij}$ is associated to the monomial symmetric function $m_{(1_1,\cdots,1_j)}(q_1^2, \cdots, q_{i+j}^2)$.

Our peak-background split derivation is very interesting for the following reasons. First, all the analytic models of halo/galaxy clustering *assume* that the (scale-independent) bias coefficients multiplying powers of the mass correlation function $\xi_0^{(0)}$ [Eq. (30)] are the peak-background split biases [Eq. (31)]. In the peak model, this equivalence can be derived from first principles. Second, the peak-background split holds even though the mean number density $\bar{n}_{\rm pk}(\nu, R_S)$ is not universal (due to its R_S dependence). For a spherical collapse prescription $\nu = \delta_{\rm sc}(z_0)/\sigma_{0S}$, this implies that it is the abundance of virialized objects of mass M_S as a function of collapse redshift z_0 which is related to the bias parameters, and not the mass function (the abundance as a function of mass M_S) at fixed redshift z_0. Third, when correctly implemented, the second of our prescriptions yields the correct scale-dependent bias factors, despite the nonuniversality which comes from the fact that the background density is correlated with the curvature. And finally, the peak-background split approximation has been shown to provide a good description of large scale peak bias in simulations [42,99] although, recently, deviations at the 5%–10% level have been reported [44,100]. Since our expressions reproduce this limit, we have confidence that our approach will furnish a good approximation at the smaller scales where the bias parameters become scale dependent (which we showed is reproduced by the peak-background split).

Summarizing, we have shown how to determine the cross correlation between the peak point process and the smoothed, linear mass overdensity δ_B order by order. In fact, because we have the full expression for the correlation between peaks and the surrounding density field, this cross correlation can be computed exactly. The calculation simplifies by noting that the integral to be done is $f(u)g(u|\nu, \nu_B)g(\nu|\nu_B)\nu_B g(\nu_B)$, where g is a Gaussian variate. This can be rewritten as $f(u)g(\nu)g(u|\nu)\nu_B g(\nu_B|\nu, u)$. Assuming that ν_B ranges over all $[-\infty, \infty]$ simplifies the integrals to compute, and one finds

$$\sigma_{{\rm pk},\delta}^2(R_B) = \sigma_{0\times}^2\tilde{\mathfrak{b}}_{10} + \sigma_{1\times}^2\tilde{\mathfrak{b}}_{01} \qquad \text{at all scales.} \tag{43}$$

This expression, which is exact to all orders, coincides with the average mass profile around an initial density peak

[36,75]. This average density profile has been shown to provide an accurate description of the cross correlation between the Stokes Q and U polarization parameters and the temperature peaks of (two-dimensional) Gaussian cosmic microwave background maps [89].

C. Relation to the local bias model

The local bias model is commonly used to model the clustering of dark matter halos and galaxies. In its Lagrangian formulation, this model assumes that the tracer overdensity $\delta_{\rm h}$ of halos of mass M_S is a local deterministic function of the linear mass density field smoothed on scale R_B. This function may be written as the Taylor series (e.g., [39])

$$\delta_{\rm h}(\nu, M_S, R_B, \mathbf{x}) = b_0 + \sum_{N=1}^{\infty} b_N(\nu)\frac{\delta_B^N(\mathbf{x})}{N!}, \tag{44}$$

where the dependence on M_S arises through the peak-background split bias parameters $b_N(\nu)$ solely. The value of b_0 is set by requiring $\langle\delta_{\rm h}(\nu, M_S, R_B, \mathbf{x})\rangle = 0$. The choice of R_B is arbitrary as long as the constraint $R_B \gg R_S$ is satisfied. The large scale clustering is then independent of the smoothing on scales $kR_B \ll 1$ (see [52,55]; see also [51] for a discussion of smoothing in the context of local Eulerian biasing). Therefore, the cross correlation between the tracers overabundance and the mass overdensity in spheres of radius R_B is

$$\begin{aligned}\sigma^2_{{\rm h},\delta}(R_B) &\equiv \langle\delta_B(\mathbf{x})\delta_{\rm h}(\nu, M_S, R_B, \mathbf{x})\rangle \\ &= \sum_{N=1}^{\infty} b_N(\nu)\frac{\langle\delta_B(\mathbf{x})\delta_B^N(\mathbf{x})\rangle}{N!} \\ &\approx \left(b_{\rm I} + \frac{1}{2}b_{\rm III}\sigma^2_{0B}\right)\sigma^2_{0B}.\end{aligned} \tag{45}$$

A comparison with Eq. (43) reveals that, even on the largest scales, the cross correlation between the peaks and the mass distributions, $\sigma^2_{{\rm pk},\delta_B} \approx b_{\rm I}\sigma^2_{0\times}$, differs from the local bias expression, which involves all the bias factors with odd N values. This can be traced to the fact that the peaks' expression is for cells centered on peaks, and only the mass field is smoothed on scale R_B, whereas the local bias expansion is for randomly placed cells, where both the tracer field and the mass have been smoothed on scale R_B.

For similar reasons, the autocorrelation function, at second order in the local bias model, is

$$\begin{aligned}\xi_{\rm h}(\nu, M_S, R_B, r) \approx\ & b_{\rm I}^2\xi_0^{(0)}(R_B, r, z_0) \\ & + \tfrac{1}{2}b_{\rm II}^2[\xi_0^{(0)}(R_B, r, z_0)]^2,\end{aligned} \tag{46}$$

where we have ignored extra terms involving powers of σ^2_{0B}, which arise owing to the continuous nature of the bias relation (44).

This demonstrates that, for large separations $r \gg R_B$, the peak correlation function $\xi_{\rm pk}(\nu, R_S, r)$ is consistent with that of the local bias model although, for peaks, the bias factors are k dependent. It is pretty clear from Eqs. (46) and (30) that the local bias scheme is a special case of the more general peak model which, on large scales, approaches a local, deterministic, scale-independent relation only in the high peak limit $\nu \gg 1$. Notwithstanding this, the previous analysis shows that

$$\begin{aligned}&\xi_{\rm pk}(\nu \gg 1, R_S, r) \\ &\quad\approx b_{\rm I}^2\xi_0^{(0)}(R_S, r, z_0) + \frac{1}{2}\left(b_{\rm II}^2 + \frac{2b_{\rm II}\gamma_1^2}{\sigma_1^2(1-\gamma_1^2)}\partial^2\right) \\ &\qquad\times[\xi_0^{(0)}(R_S, r, z_0)]^2 + \text{other terms},\end{aligned} \tag{47}$$

on smaller scales where the k dependence of peak bias matters. As can be seen, the second order peak bias (the term in curly brackets) remains scale dependent even for the highest peaks. Therefore, if the peaks' model is correct, then we shall expect deviations from the local bias model on mildly nonlinear scales $k \sim 0.01 - 0.1h\,{\rm Mpc}^{-1}$, even for the most prominent density maxima.

IV. GRAVITATIONAL EVOLUTION OF THE PEAK CORRELATION FUNCTION

Thus far, we have explored the scale dependence of bias in the two-point correlation of local maxima of the primordial density field. However, nonlinear collapse and pairwise motions induced by gravitational instabilities will distort the initial correlation. Since the precise calculation of the dynamical evolution of $\xi_{\rm pk}(\nu, R_S, r)$ is rather involved, we will assume that the initial density peaks are test particles which flow locally with the dark matter according to the Zel'dovich approximation (i.e., peaks move along straight lines). We will calculate the correlation of their positions as a function of redshift and show that we recover a velocity damping factor and a mode-coupling power similar to that found in (Eulerian) renormalized perturbation theory (RPT, see [101]).

A. The peak correlation function in the Zel'dovich approximation

In the Lagrangian approach, the Eulerian comoving position and proper velocity of a density peak can generally be expressed as a mapping

$$\mathbf{x}_{\rm pk}(z) = \mathbf{q}_{\rm pk} + \mathbf{S}(\mathbf{q}_{\rm pk}, z), \qquad \mathbf{v}_{\rm pk}(z) = a(z)\dot{\mathbf{S}}(\mathbf{q}_{\rm pk}, z), \tag{48}$$

where $\mathbf{q}_{\rm pk}$ is the initial peak position, $\mathbf{S}(\mathbf{q}, z)$ is the displacement field, a is the scale factor, and the peak velocity is in standard units (i.e., not scaled by aHf). A dot denotes a derivative with respect to cosmic time. We will assume that the local maxima are test particles that do not interact

with each other. Therefore, at the first order, the peak position is described by the Zeldovich approximation [102], in which the displacement factorizes into a time and a spatial component,

$$\mathbf{S} = -D(z)\nabla\Phi(\mathbf{q}), \qquad \dot{\mathbf{S}} = -\beta(z)\nabla\Phi(\mathbf{q}). \tag{49}$$

Here, $D(z)$ is the growth factor of linear mass density perturbations and $\Phi(\mathbf{q})$ is the perturbation potential linearly extrapolated to present time. Explicitly, $\Phi(\mathbf{q}) = \phi(\mathbf{q}, z)/4\pi G\bar{\rho}_m(z)a^2D(z)$ where $\phi(\mathbf{q}, z)$ is the Newtonian gravitational potential and $\bar{\rho}_m(z)$ is the average matter density. Finally, $\beta(z) = HDf$, where $H(z)$ is the Hubble constant, is proportional to the logarithmic derivative $f = d\ln D/d\ln a$. Note that $f(z)$ scales as $\Omega_m(z)^{0.6}$ for a wide range of cold dark matter (CDM) cosmologies [103].

Let us now consider an ensemble of realizations of initial peak positions. The correlation function $\xi_{\rm pk}(\nu, R_S, r, z)$ (we will henceforth omit the dependence on ν and R_S for brevity) is related to the zeroth moment of the joint probability $P_2(\mathbf{v}_1, \mathbf{v}_2; \mathbf{r}, z|\mathrm{pk})$ to have a pair of peaks separated by a distance $\mathbf{r}$ and with normalized velocities $\mathbf{v}_1$ and $\mathbf{v}_2$. Following [104], we write

$$\bar{n}^2_{\rm pk}[1 + \xi_{\rm pk}(r, z)] = \int d^3\mathbf{v}_1 d^3\mathbf{v}_2 P_2(\mathbf{v}_1, \mathbf{v}_2; \mathbf{r}, z|\mathrm{pk}). \tag{50}$$

As we will see shortly, even though the probability $P_2(\mathbf{v}_1, \mathbf{v}_2; \mathbf{r}, z|\mathrm{pk})$ depends upon the distance $r = \mathbf{r}\cdot\hat{\mathbf{r}}$ *and* the unit direction vector $\hat{\mathbf{r}}$, the peak correlation depends only on r. When the peak motions are governed by the Zel'dovich approximation Eq. (49), we can easily relate $P_2(\mathbf{v}_1, \mathbf{v}_2; r, z|\mathrm{pk})$ to the joint probability distribution at the initial redshift $z_i \gg 1$,

$$\begin{aligned} P_2(\boldsymbol{v}_1, \boldsymbol{v}_2; r, z|\mathrm{pk}) &= P_2(\boldsymbol{v}_1, \boldsymbol{v}_2; \mathbf{r} - \sigma_v\Delta\boldsymbol{v}_{12}, z_i|\mathrm{pk}) \\ &= \int d^3\mathbf{r}'\delta^{(3)}(\mathbf{r}' - \mathbf{r} \\ &\quad + \sigma_v\Delta\boldsymbol{v}_{12})P_2(\boldsymbol{v}_1, \boldsymbol{v}_2; \mathbf{r}', z_i|\mathrm{pk}), \end{aligned} \tag{51}$$

where $\sigma_v \equiv \sigma_{-1}(z)$ is the three-dimensional rms velocity variance of the matter, and $\Delta\boldsymbol{v}_{12}$ is the velocity difference $\boldsymbol{v}_2 - \boldsymbol{v}_1$. Note that in this equation and those that follow, velocities have been scaled by $aHf\sigma_{-1}$, i.e., $\boldsymbol{v} \equiv \mathbf{v}/\sigma_{-1}$. Equation (51) is a consequence of Liouville's theorem, which states that the phase space density of peaks is conserved (so, as shown in [104], one can easily obtain a differential equation describing the evolution of the n-particle distribution functions). It is especially useful because we know how to calculate two-point distributions subjects to the peak constraint in the Gaussian initial conditions,

$$\begin{aligned} P_2(\boldsymbol{v}_1, \boldsymbol{v}_2; \mathbf{r}', z_i|\mathrm{pk}) = \int d^6\zeta_1 d^6\zeta_2 n_{\rm pk}(\mathbf{x}'_1)n_{\rm pk}(\mathbf{x}'_2) \\ \times P_2(\mathbf{w}_1, \mathbf{w}_2; \mathbf{r}', z_i). \end{aligned} \tag{52}$$

Here, P_2 is a joint probability for the 13-dimensional vector $\mathbf{w} = (v_i, \eta_i, \nu, \zeta_A)$ of variables at position $\mathbf{x}'_1$ and $\mathbf{x}'_2$, and $n_{\rm pk}(\mathbf{x})$ is the Klimontovitch density Eq. (18). Expressing the Dirac delta as the Fourier transform of a uniform distribution, we find

$$\begin{aligned} \bar{n}^2_{\rm pk}[1 + \xi_{\rm pk}(r, z)] &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int d^3\mathbf{r}' e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \int d^6\zeta_1 d^6\zeta_2 n_{\rm pk}(\mathbf{x}'_1)n_{\rm pk}(\mathbf{x}'_2)\left\{\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}_2 P_2(\mathbf{w}_1, \mathbf{w}_2; \mathbf{r}', z_i)e^{i\sigma_v\mathbf{k}\cdot\Delta\boldsymbol{v}_{12}}\right\} \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int d^3\mathbf{r}' e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \int d^6\zeta_1 d^6\zeta_2 n_{\rm pk}(\mathbf{x}'_1)n_{\rm pk}(\mathbf{x}'_2)P_2(\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}', z_i) \\ &\quad \times\left\{\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}_2 P_2(\boldsymbol{v}_1, \boldsymbol{v}_2|\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}', z_i)e^{i\sigma_v\mathbf{k}\cdot\Delta\boldsymbol{v}_{12}}\right\}, \end{aligned} \tag{53}$$

where the vector $\mathbf{y}$ corresponds to (η_i, ν, ζ_A). The second equality follows from Bayes' theorem. To integrate over the velocities, we use the identity

$$\begin{aligned} \int d^N\mathbf{y} y_{i_1}\cdots y_{i_n}P(\mathbf{y})e^{i\mathbf{J}\cdot\mathbf{y}} &= (-i)^n\frac{\partial}{\partial J_{i_1}}\cdots\frac{\partial}{\partial J_{i_n}} \\ &\quad \times \exp\left(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J} + i\mathbf{J}^\dagger\Xi\right), \end{aligned} \tag{54}$$

which follows from the relation

$$P(\mathbf{y}) = \frac{1}{(2\pi)^N}\int d^N\mathbf{J}\exp\left[-i\mathbf{J}^\dagger(\mathbf{y} - \Xi) - \frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J}\right]. \tag{55}$$

Here, $P(\mathbf{y})$ is an N-dimensional Gaussian multivariate of covariance matrix $\Sigma = \langle\mathbf{y}\mathbf{y}^\dagger\rangle$ and centered at $\mathbf{y} = \Xi$. Equation (54) is a very useful relation since it allows us to circumvent the inversion of the covariance matrix. On inserting the above identity into Eq. (53), the peak correlation function may now be formulated as

$$\begin{aligned} \bar{n}^2_{\rm pk}[1 + \xi_{\rm pk}(r, z)] &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int d^3\mathbf{r}' e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \\ &\quad \times \int d^6\zeta_1 d^6\zeta_2 n_{\rm pk}(\mathbf{x}'_1)n_{\rm pk}(\mathbf{x}'_2) \\ &\quad \times P_2(\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}', z_i) \\ &\quad \times \exp\left(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J} + i\mathbf{J}^\dagger\Xi\right), \end{aligned} \tag{56}$$

where $\mathbf{J} = \sigma_v(\mathbf{k}, -\mathbf{k})$. The task of computing the redshift evolution of the peak correlation function boils down to the

evaluation of the six-dimensional covariance matrix Σ and mean vector Ξ.

B. A simple illustration: The mass correlation function

To understand the physical meaning of nonlinear corrections as well as emphasize the relation with, e.g., RPT, it is instructive to consider first the *unbiased* case, i.e., the evolution of the matter correlation function. Therefore, there is no peak constraint, so the two-particle probability density $P_2(\boldsymbol{v}_1, \boldsymbol{v}_2; \mathbf{r}, z)d^3\boldsymbol{v}_1 d^3\boldsymbol{v}_2$ is simply the probability to find a pair of dark matter particles separated by a distance $\mathbf{r}$ and with velocities $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, respectively. As a consequence, Σ is the covariance of matter velocity components, i.e., $\Sigma_{ij} = \langle v_i v_j \rangle$, and $\Xi \equiv 0$ because there is no net mean streaming between two randomly selected locations. After some simplifications, the mass correlation function evolved with the Zel'dovich ansatz takes the simple form

$$1 + \xi_{\rm m}(r, z) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} e^{-(1/3)k^2\sigma_v^2} \int d^3\mathbf{r}' e^{-i\mathbf{k}\cdot\mathbf{r}'} \times \exp\left\{\frac{1}{3}\frac{\sigma_v^2}{\sigma_{-1}^2}[\xi_0^{(-1)} + \xi_2^{(-1)}]k^2 - \frac{\sigma_v^2}{\sigma_{-1}^2}\xi_2^{(-1)}(\mathbf{k}\cdot\hat{\mathbf{r}})^2\right\}. \quad (57)$$

This agrees with Eqs. (16) and (17) of [104] provided that his $\phi(r)$ corresponds to our $\xi_0^{(-2)}(r)$, so that $\Delta\phi(0) \equiv -\sigma_{-1}^2$. It should be noted that σ_{-1} and $\xi_\ell^{(-1)}$ are evaluated at redshift $z_i \gg 1$. Hence, the ratio σ_v/σ_{-1} is equal to $D(z)/D(z_i)$. After some manipulations, the last exponent can be reexpressed as (in the notation of [105])

$$I(\mathbf{k}, \mathbf{r}) \equiv \left(\frac{D(z)}{D(z_i)}\right)^2 \left\{\frac{1}{3}[\xi_0^{(-1)} + \xi_2^{(-1)}]k^2 - \xi_2^{(-1)}(\mathbf{k}\cdot\hat{\mathbf{r}})^2\right\} = \left(\frac{D(z)}{D(z_i)}\right)^2 \int \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{(\mathbf{k}\cdot\hat{\mathbf{q}})^2}{q^2} P_\delta(q) e^{i\mathbf{q}\cdot\mathbf{r}}, \quad (58)$$

where we have used that

$$\frac{1}{4\pi}\int d\Omega_{\hat{\mathbf{q}}} \hat{q}_i \hat{q}_j e^{i\mathbf{q}\cdot\mathbf{r}} = \frac{1}{3}[j_0(qr) + j_2(qr)]\delta_{ij} - j_2(qr)\hat{r}_i\hat{r}_j. \quad (59)$$

On taking the Fourier transform of Eq. (57), we recover the well-known expression for the nonlinear mass power spectrum in the Zel'dovich approximation,

$$P_{\rm m}(k, z) = e^{-(1/3)k^2\sigma_{-1}^2(z)} \int d^3\mathbf{r}' e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{n=1}^{\infty} \frac{[I(\mathbf{k}, \mathbf{r})]^2}{n!} = e^{-(1/3)k^2\sigma_{-1}^2(z)} \sum_{n=1}^{\infty} (2\pi)^3 n! \left(\frac{D(z)}{D(z_i)}\right)^{2(n+1)} \times \prod_{j=1}^{n}\left\{\int \frac{d^3\mathbf{q}_j}{(2\pi)^3} P_\delta(q_j, z_i)\right\} \times [F_n(\mathbf{q}_1, \cdots, \mathbf{q}_n)]^2 \delta^{(3)}(\mathbf{k} - \mathbf{q}_{1\cdots n}), \quad (60)$$

where $\mathbf{q}_{1\cdots n} = \mathbf{q}_1 + \cdots + \mathbf{q}_n$ and the kernels F_n [106,107] are symmetric, homogeneous functions of the wave vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ that describe the nonlinear evolution of the density field in the Zel'dovich approximation [108],

$$F_n(\mathbf{q}_1, \cdots \mathbf{q}_n) = \frac{1}{n!}\frac{(\mathbf{k}\cdot\hat{\mathbf{q}}_1)}{q_1} \cdots \frac{(\mathbf{k}\cdot\hat{\mathbf{q}}_n)}{q_n}. \quad (61)$$

They are nearly independent of $\Omega_{\rm m}$ and Ω_Λ. Note that, in the exact dynamics, $F_2(\mathbf{q}_1, \mathbf{q}_2) = 5/7 + 1/2(q_1/q_2 + q_2/q_1)\hat{\mathbf{q}}_1\cdot\hat{\mathbf{q}}_2 + 2/7(\hat{\mathbf{q}}_1\cdot\hat{\mathbf{q}}_2)^2$. In the Zel'dovich approximation however, the coefficients are all equal to $1/2$, reflecting the fact that momentum is only conserved at first order.

In Eq. (60), the sum represents all the mode-coupling corrections, whereas the decaying exponential prefactor corresponds to the propagator [105], which for Gaussian initial conditions describes the (imperfect) correlation between the nonlinear and linear density field. Hence, the Lagrangian formulation of [104] furnishes an easy way to obtain the resummed RPT propagator (see also [105,109] for a similar Lagrangian description). At the second order, the mass correlation is

$$\xi_{\rm m}(r, z) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} e^{-(1/3)k^2\sigma_{-1}^2(z)} \left\{P_\delta(k, z) + \frac{2}{(2\pi)^3}\int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 [F_2(\mathbf{q}_1, \mathbf{q}_2)]^2 \times P_\delta(q_1, z) P_\delta(q_2, z)\delta^{(3)}(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2)\right\}. \quad (62)$$

The second order kernel scales as $F_2(\mathbf{q}_1, \mathbf{q}_2) \propto k^2$ in the (squeezed) limit where the total momentum $\mathbf{k}$ goes to zero as a consequence of mass-momentum conservation [110,111].

C. Including the peak constraint

The computation of Eq. (56) is rather involved when the peak constraint is taken into account. Here, we will explain the basic result and show that it generalizes previous formulas based on the local bias scheme. Details of the calculation can be found in Appendix B where it is shown, among others, that it is quite convenient to work with the Fourier transforms of Σ and Ξ. We eventually arrive at

$$\xi_{\rm pk}(r, z) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\left\{e^{-(1/3)k^2\sigma_{\rm vpk}^2(z)}\left[\left(\frac{D(z)}{D(z_i)}\right)\tilde{\mathfrak{b}}_{\rm vpk}(k) + \tilde{\mathfrak{b}}_{\rm I}(k, z_i)\right]^2 P_{\delta_s}(k, z_i) + P_{\rm MC}(k, z)\right\} e^{i\mathbf{k}\cdot\mathbf{r}} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} G^2(k, z)[\tilde{\mathfrak{b}}_{\rm I}^{\rm E}(k, z)]^2 P_{\delta_s}(k, z_0) e^{i\mathbf{k}\cdot\mathbf{r}} + \xi_{\rm MC}(k, z). \quad (63)$$

The last equality follows upon making the replacement $z_i \to z_0$ (everywhere but in the exponential prefactor). Here, z_0 is some fiducial redshift which we take to be the

collapse redshift (because results are usually normalized to low redshift quantities). The function

$$G^2(k, z) = D_+^2(z) e^{-(1/3)k^2\sigma^2_{\rm vpk}(z)}, \tag{64}$$

where $D_+(z) \equiv D(z)/D(z_0)$ is the linear growth rate normalized to its value at the epoch of collapse, is a damping term induced by velocity diffusion around the mean displacement [104]. Namely, $G(k, z)$ is the peaks' propagator in the Zel'dovich approximation, and is analogous to that introduced in [101,105] for the matter evolution (the latter depends on σ_{-1}). In the exponent, the factor of $1/3$ reflects the fact that the dynamics governing $\xi_{\rm pk}(r, z)$ is effectively one dimensional (because, on average, only the streaming along the separation vector of a peak pair matters). The first order Eulerian and Lagrangian peak biases $\tilde{\rm b}_{\rm I}^{\rm E}(k, z)$ and $\tilde{\rm b}_{\rm I}(k, z_0)$, both defined with respect to $P_{\delta_S}(k, z_0)$, are then related according to

$$\tilde{\rm b}_{\rm I}^{\rm E}(k, z) \equiv \tilde{\rm b}_{\rm vpk}(k) + D_+^{-1}(z)\tilde{\rm b}_{\rm I}(k, z_0). \tag{65}$$

In the limit $z \to \infty$, we recover $\tilde{\rm b}_{\rm I}^2(k, z_0)P_{\delta_S}(k, z_0)$ at the first order. The presence of $\tilde{\rm b}_{\rm vpk}(k)$ reflects the fact that peaks stream towards (or move apart from) each other in high (low) density environments, but this effect is k dependent owing to the statistical velocity bias. Still, the k dependence of $\tilde{\rm b}_{\rm vpk}(k)$ is such that the linear Eulerian peak bias $\tilde{\rm b}_{\rm I}^{\rm E}$ is scale independent in the limit $k \ll 1$, in agreement with the "local bias theorem" [55,112,113]. On writing the Eulerian linear peak bias as $\tilde{\rm b}_{\rm I}^{\rm E} \equiv \tilde{\rm b}_{10}^{\rm E} + \tilde{\rm b}_{01}^{\rm E}k^2$, Eq. (65) becomes

$$\begin{aligned}\tilde{\rm b}_{10}^{\rm E}(z) &\equiv 1 + D_+^{-1}(z)\tilde{\rm b}_{10}(z_0),\\ \tilde{\rm b}_{01}^{\rm E}(z) &\equiv D_+^{-1}(z)\tilde{\rm b}_{01}(z_0) - \frac{\sigma_0^2}{\sigma_1^2}.\end{aligned} \tag{66}$$

The first relation is the usual formula for the Eulerian, linear scale-independent bias [33]. It shows that, unsurprisingly, the bias of the peak distribution eventually relaxes to unity [114–116]. The second relation implies that $\tilde{\rm b}_{01}^{\rm E}$ approaches the negative, R_S-dependent constant $-\sigma_0^2/\sigma_1^2$ as the gravitational instability grows. A scale dependence in the linear peak bias $\tilde{\rm b}_{\rm I}^{\rm E}$ thus persists in the long term if the linear velocities are statistically biased (otherwise $\tilde{\rm b}_{01}^{\rm E} \to 0$ as $z \to \infty$).

Thinking of the first order peak statistics as arising from the *continuous* bias relation Eq. (5) furnishes a straightforward derivation of Eq. (66). Namely, the linear continuity equation for the mass density field reads $\dot{\delta}(t) = -\nabla \cdot \mathbf{v}(t) = \theta(t)$ whereas, for density peaks, $\dot{\delta}_{\rm pk}(t) = -\nabla \cdot \mathbf{v}_{\rm pk}(t) = \theta_{\rm pk}(t) = \tilde{\rm b}_{\rm vpk}W\theta$. Recall that $\tilde{\rm b}_{\rm vpk}W$ is an operator which, in Fourier space, is a multiplication by $(1 - \sigma_{0S}^2/\sigma_{1S}^2k^2)W(kR_S)$. The solution is

$$\begin{pmatrix}\delta(t_0)\\ \delta_{\rm pk}(t_0)\end{pmatrix} = \begin{pmatrix}D_+^{-1}(t) & 0\\ (D_+^{-1}(t) - 1)\tilde{\rm b}_{\rm vpk}W & 1\end{pmatrix}\begin{pmatrix}\delta(t)\\ \delta_{\rm pk}(t)\end{pmatrix}. \tag{67}$$

As a result, the leading-order contribution to the cross correlation between the average overdensity of peaks defined on scale R_S and the mass density field smoothed on an arbitrary scale R_B satisfies

$$\begin{aligned}\langle\delta_{\rm pk}(t_0)\delta_B(t_0)\rangle &= \langle D_+^{-1}(t)\delta_B(t)[(D_+^{-1}(t) - 1)\tilde{\rm b}_{\rm vpk}W\delta(t) + \delta_{\rm pk}(t)]\rangle\\ &= D_+^{-1}(t)[(D_+^{-1}(t) - 1)\langle\delta_B(t)\tilde{\rm b}_{\rm vpk}\delta_S(t)\rangle + \langle\delta_B(t)\delta_{\rm pk}(t)\rangle]\\ &= D_+^{-1}(t)\left[(D_+^{-1}(t) - 1)\left(\sigma_{0\times}^2(t) - \frac{\sigma_{0S}^2}{\sigma_{1S}^2}\sigma_{1\times}^2(t)\right) + \tilde{\rm b}_{10}(t)\sigma_{0\times}^2(t) + \tilde{\rm b}_{01}(t)\sigma_{1\times}^2(t)\right].\end{aligned} \tag{68}$$

On writing the left-hand side as $\tilde{\rm b}_{10}(t_0)\sigma_{0\times}^2(t_0) + \tilde{\rm b}_{01}(t_0)\sigma_{1\times}^2(t_0)$ and isolating the terms in $\sigma_{0\times}^2$ and $\sigma_{1\times}^2$, we arrive at

$$\begin{aligned}\tilde{\rm b}_{10}(t_0)\sigma_{0\times}^2(t_0) = D_+^{-1}(t)[&(D_+^{-1}(t) - 1)\sigma_{0\times}^2(t)\\ &+ \tilde{\rm b}_{10}(t)\sigma_{0\times}\sigma_{0\times}^2(t)]\end{aligned} \tag{69}$$

$$\begin{aligned}\tilde{\rm b}_{01}(t_0)\sigma_{1\times}^2(t_0) = D_+^{-1}(t)\Bigg[&-(D_+^{-1}(t) - 1)\frac{\sigma_{0S}^2}{\sigma_{1S}^2}\sigma_{1\times}^2(t)\\ &+ \tilde{\rm b}_{01}(t)\sigma_{1\times}^2(t)\Bigg].\end{aligned} \tag{70}$$

Using the fact that the spectral moments of the linear mass density field scale as $\sigma_{n\times}(t_0) = D_+^{-1}(t)\sigma_{n\times}(t)$, we can rewrite the above as

$$\begin{aligned}\tilde{\rm b}_{10}(t) &= 1 + D_+^{-1}(t)(\tilde{\rm b}_{10}(t_0) - 1),\\ \tilde{\rm b}_{01}(t) &= -\frac{\sigma_{0S}^2}{\sigma_{1S}^2} + D_+^{-1}(t)\left(\tilde{\rm b}_{01}(t_0) + \frac{\sigma_{0S}^2}{\sigma_{1S}^2}\right),\end{aligned} \tag{71}$$

which is precisely Eq. (66). On linear scales, although the peak bias is deterministic in Fourier space, it is generally stochastic and scale dependent in configuration space [75]. In principle, it would be possible to solve also for the cross correlation coefficient between the peaks and the smoothed mass density field δ_B by considering the time evolution of the peak rms variance $\langle\delta_{\rm pk}^2(t)\rangle$ for instance (see, e.g., [115]).

With the peak constraint, the one-loop contribution to the mode-coupling power $P_{\rm MC}(k, z)$ in the Zel'dovich approximation eventually becomes

$$
\begin{aligned}
P_{\rm MC}(k,z) = {} & \frac{2}{(2\pi)^3} D_+^4(z) e^{-(1/3)k^2\sigma_{\rm vpk}^2} \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 [\tilde{\mathrm{b}}_{\rm II}^{\rm E}(\mathbf{q}_1,\mathbf{q}_2,z)]^2 P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2) \\
& + \frac{6}{(2\pi)^3} \frac{1}{\sigma_{1S}^2} D_+^2(z) e^{-(1/3)k^2\sigma_{\rm vpk}^2} \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 q_1 q_2 (\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2) \tilde{\mathrm{b}}_{\rm II}^{\rm E}(\mathbf{q}_1,\mathbf{q}_2,z) P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2) \\
& - \frac{5}{2(2\pi)^3} \frac{1}{\sigma_{2S}^2} D_+^2(z) e^{-(1/3)k^2\sigma_{\rm vpk}^2} \left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}\right) \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 q_1^2 q_2^2 [3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2 - 1] \\
& \times [3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2 - 1] \tilde{\mathrm{b}}_{\rm II}^{\rm E}(\mathbf{q}_1,\mathbf{q}_2,z) P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2) \\
& + \frac{25}{64(2\pi)^3} \frac{1}{\sigma_{2S}^4} e^{-(1/3)k^2\sigma_{\rm vpk}^2} \left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}\right)^2 \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 q_1^4 q_2^4 \{11 - 30(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2 + 27(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^4 \\
& - 6(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2 \times [5 - 42(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2 + 45(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^4] + 9(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^4 [3 - 30(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2 + 35(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^4]\} \\
& \times P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2) \\
& + \frac{27}{8(2\pi)^3} \frac{1}{\sigma_{1S}^4} e^{-(1/3)k^2\sigma_{\rm vpk}^2} \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 q_1^2 q_2^2 [3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2 (\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2 - 2(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2 + 1] \\
& \times P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2) \\
& - \frac{15}{4(2\pi)^3} \frac{1}{\sigma_{1S}^2 \sigma_{2S}^2} e^{-(1/3)k^2\sigma_{\rm vpk}^2} \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 q_1^3 q_2^3 (\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2) [15(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2 (\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2 - 18(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2 \\
& + 7] P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2), \qquad (72)
\end{aligned}
$$

where all spectral moments σ_n and power spectra $P_{\delta_s}(q)$ are evaluated at z_0. The second order Eulerian bias $\tilde{\mathrm{b}}_{\rm II}^{\rm E}(\mathbf{q}_1,\mathbf{q}_2,z)$ is a symmetric function of q_1 and q_2,

$$
\begin{aligned}
\tilde{\mathrm{b}}_{\rm II}^{\rm E}(\mathbf{q}_1,\mathbf{q}_2,z) \equiv {} & \mathcal{F}_2(\mathbf{q}_1,\mathbf{q}_2) + \tfrac{1}{2} D_+^{-1}(z) [\mathcal{F}_1(\mathbf{q}_1)\tilde{\mathrm{b}}_{\rm I}(q_2,z_0) \\
& + \mathcal{F}_1(\mathbf{q}_2)\tilde{\mathrm{b}}_{\rm I}(q_1,z_0)] \\
& + \tfrac{1}{2} D_+^{-2}(z) \tilde{\mathrm{b}}_{\rm II}(q_1,q_2,z_0). \qquad (73)
\end{aligned}
$$

Here, $\tilde{\mathrm{b}}_{\rm II}(q_1, q_2, z_0)$ represents the second order Lagrangian peak bias Eq. (23) and, in analogy with standard perturbation theory (PT), we have introduced the kernel $\mathcal{F}_n$ which characterize the nth order evolution of the peak correlation function in the Zel'dovich approximation,

$$
\begin{aligned}
\mathcal{F}_n(\mathbf{q}_1,\cdots\mathbf{q}_n) \equiv {} & \frac{1}{n!} \\
& \times \frac{(\mathbf{k}\cdot\hat{\mathbf{q}}_1)}{q_1} \cdots \frac{(\mathbf{k}\cdot\hat{\mathbf{q}}_n)}{q_n} \tilde{\mathrm{b}}_{\rm vpk}(q_1) \cdots \tilde{\mathrm{b}}_{\rm vpk}(q_n). \qquad (74)
\end{aligned}
$$

$\mathcal{F}_n$ is identical to the standard PT kernels except for the velocity bias $\tilde{\mathrm{b}}_{\rm vpk}(q)$. As can be seen, the second order mode-coupling power is simply obtained from Eq. (22) upon a replacement $\tilde{\mathrm{b}}_{\rm II} \to \tilde{\mathrm{b}}_{\rm II}^{\rm E}$ in Eq. (22), a Fourier transformation and a multiplication by the diffusion damping prefactor $\exp(-(1/3)k^2\sigma_{\rm vpk}^2)$. Therefore, the mode coupling induced by gravity is Eq. (72) minus the second order terms in Eq. (22). We also note the plus sign in the second term of (72), which follows from the fact that the Fourier transform of $\xi_1^{(1/2)}(r)$ is $iq(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})$. Finally, we simply Fourier transform $P_{\rm MC}(k, z)$ to obtain the mode-coupling contribution in configuration space.

In Eq. (56), each power of $\mathbf{J}$ brings a factor of $\mathcal{F}_1(q)$. At the second order, there is only one contribution proportional to $\mathbf{J}^4 \sim \mathcal{F}_2^2$ but, contrary to the Eulerian PT expression, it does not involve the first order bias. In fact, taking the local bias limit in which $\tilde{\mathrm{b}}_{\rm I} = b_{\rm I}$, $\tilde{\mathrm{b}}_{\rm II} = b_{\rm II}$ and ignoring the exponential damping prefactor and a possible statistical velocity bias (i.e., $G \equiv 1$ and $\tilde{\mathrm{b}}_{\rm vpk} \equiv 1$), we *do not recover* the familiar PT expression (e.g., [46])

$$
\begin{aligned}
P_{\rm MC}(k,z) = {} & \frac{2}{(2\pi)^3} \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 \left[b_{\rm I}^{\rm E} F_2(\mathbf{q}_1,\mathbf{q}_2) \right. \\
& \left. + \frac{b_{\rm II}^{\rm E}}{2} \right]^2 P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2), \qquad (75)
\end{aligned}
$$

where $b_N^{\rm E}$ are Eulerian peak-background split bias factors, but rather (omitting subleading powers in the growth factors for simplicity)

$$
\begin{aligned}
P_{\rm MC}(k,z) \sim {} & \frac{2}{(2\pi)^3} \int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 \left[F_2(\mathbf{q}_1,\mathbf{q}_2) \right. \\
& \left. + \frac{b_{\rm I}}{2}(F_1(\mathbf{q}_1) + F_1(\mathbf{q}_2)) + \frac{b_{\rm II}}{2} \right]^2 \\
& \times P_{\delta_s}(q_1) P_{\delta_s}(q_2) \delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2). \qquad (76)
\end{aligned}
$$

The difference is not surprising. In the first expression, local bias is applied to the *evolved* density field and, thus, bias coefficients enter as multiplicative factors in each term

of the perturbation series, including the mode-coupling contribution. In our approach, however, peaks are identified in the initial conditions and then evolved gravitationally such that, in the long term, the Lagrangian bias factors drop to zero and the unbiased case is reproduced. In fact, our Eq. (76) agrees with the local Lagrangian bias expression of [117],

$$P_{\rm MC}(k,z) \sim \frac{2}{(2\pi)^3}\int d^3\mathbf{q}_1 \int d^3\mathbf{q}_2 \Big[(1+b_0)F_2(\mathbf{q}_1,\mathbf{q}_2) + \frac{b_{\rm I}}{2}(F_1(\mathbf{q}_1)+F_1(\mathbf{q}_2)) + \frac{b_{\rm II}}{2}\Big]^2 \times P_{\delta_s}(q_1)P_{\delta_s}(q_2)\delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2), \quad (77)$$

except for a prefactor $(1+b_0)$ multiplying the second order kernel $F_2(\mathbf{q}_1,\mathbf{q}_2)$ (in the peak model, this prefactor is unity since the zeroth order bias is $b_0=0$). Therefore, in the local bias limit considered here, the one-loop contribution to the mode coupling can be thought of as arising from the (nonlocal) relation [47]

$$1+\delta_{\rm pk}(\mathbf{x},z) = [1+\delta_{\rm pk}(\mathbf{q})][1+\delta(\mathbf{x},z)]. \quad (78)$$

Here, $\delta_{\rm pk}(\mathbf{x},z)$ and $\delta(\mathbf{x},z)$ are the Eulerian peak and matter overdensity, $\delta_{\rm pk}(\mathbf{q}) = b_{\rm I}\delta(\mathbf{q}) + (1/2)b_{\rm II}\delta(\mathbf{q})^2$ is the Lagrangian peak overabundance, and $\mathbf{x}=\mathbf{q}+\mathbf{S}(\mathbf{q},z)$ is the mapping from Lagrangian to Eulerian coordinates. We expect also our Eq. (76) to agree with the local Lagrangian bias expressions of [55] if we include higher-order perturbative corrections to the peak displacement. We defer the calculation of the mode coupling at second order Lagrangian perturbation theory (2LPT) to a future work.

D. Mode-coupling power: Shot noise and shift in the baryonic acoustic scale

To help visualize the results of the previous section, Fig. 4 shows the evolved power spectrum $P_{\rm pk}(k,z)$ of 2σ density peaks as computed from Eq. (63) at redshift $z=z_0=0.3$ and $z=1$. The peak power spectrum is split into two distinct parts. The first piece is the linear mass power spectrum P_δ smoothed with the filter $W(R_S,k)^2$, amplified by a scale-dependent prefactor $(\tilde{b}_{\rm vpk}(k) + D_+^{-1}(z)\tilde{b}_{\rm I}(k,z_0))^2$ and damped with a diffusion kernel. The second piece is a sum of the mode-coupling power present in the initial conditions (induced by the peak biasing) and of that generated during the nonlinear evolution (induced by gravity).

At small wave number, the mode-coupling power arising from the peak biasing dominates and, in the limit $k \ll 1$, contribute a pure white noise term whose amplitude is independent of redshift. Namely, the low-k white noise tail is generated by the peak biasing and not by gravity. While it can be shown that the redistribution of the matter caused by the nonlinear interactions cannot build a white noise tail $P_\delta(k) \propto k^0$ in the mass power spectrum (see, e.g., [103]), it is interesting that this holds also for tracers of the mass distribution (such as density peaks) which do not conserve momentum. In Eq. (72), the redshift independence of $P_{\rm MC}$ $(k=0)$ is ensured by the fact that all the $D(z)$-dependent terms come at least with a factor of $\mathcal{F}_{n=1,2}$ which vanishes in the limit $k\to 0$. The amplitude of $P_{\rm MC}$ $(k=0)$, however, strongly depends upon the peak height.

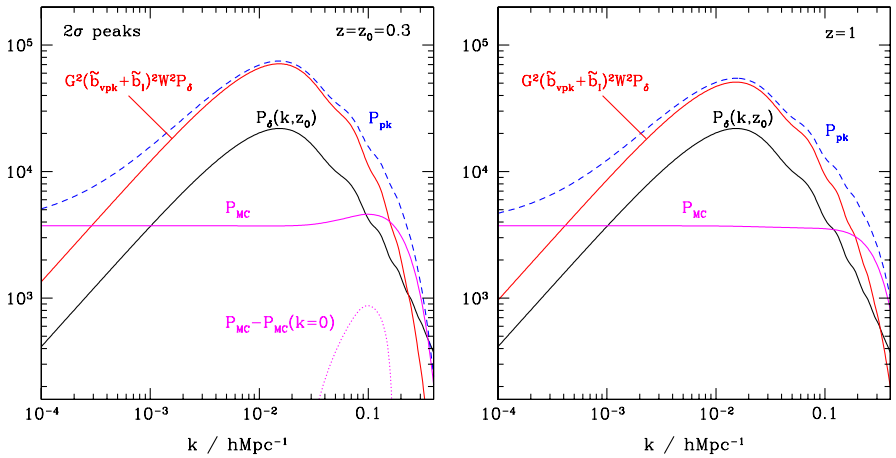

FIG. 4 (color online). The evolved power spectrum $P_{\rm pk}(k,z)$ (dashed curve) for the 2σ peaks as predicted by Eq. (63) at the redshift of collapse $z=z_0=0.3$ (left panel) and at $z=1$ (right panel). The first order term and the mode-coupling contribution are shown as solid curves together with the unsmoothed linear mass power spectrum $P_\delta(k,z_0)$. At low wave number, the scale-dependent contribution to the mode-coupling power, $P_{\rm MC}(k)-P_{\rm MC}(k=0)$, scales as k^2 since there is no mass-momentum conservation. The Poisson expectation $1/\bar{n}_{\rm pk}$ (not shown on this figure) is $\approx 10^4$.

A quick calculation shows that, as $k \to 0$,

$$P_{\rm MC}(k,z) \to \frac{1}{(2\pi)^2}\int_0^\infty dq q^2 [\tilde{\rm b}_{\rm II}(q,q,z_0)]^2 P_{\delta_s}(q)^2 - \frac{1}{2\pi^2}\int_0^\infty dq q^4 \Big[\frac{1}{\sigma_1^2}+\frac{q^2}{\sigma_2^2}\Big(1+\frac{2}{5}\partial_\alpha \times \ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu)_{\alpha=1}|\Big)\Big]\tilde{\rm b}_{\rm II}(q,q,z_0)P_{\delta_s}(q)^2 + \frac{1}{(2\pi)^2}\int_0^\infty dq q^6\Big[\frac{63}{10\sigma_1^4}+\frac{46q^2}{7\sigma_1^2\sigma_2^2}+\frac{3525q^4}{448\sigma_2^4} \times \Big(1+\frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu)|_{\alpha=1}\Big)^2\Big]P_{\delta_s}(q)^2. \quad (79)$$

All three integrals converge owing to the filtering of the mass density field. However, the first and third terms are always positive whereas the second term can be negative. For the 2σ peaks, these are 3.2, -1.4, and 1.9×10^3, respectively. As the threshold height is raised, the first term increasingly dominates (we find 10.6, -2.1, and 1.0×10^4 for the 3σ peaks). Therefore, at second order in the expansion, the shot noise, which we define as the sum of Eq. (79) and the Poisson expectation $1/\bar{n}_{\rm pk}$, is super-Poisson for $\nu>2$. This prediction seems to be at odds with recent lines of evidence suggesting that the shot noise of massive halos is sub-Poisson [52,118,119]. However, we caution that this super-Poisson behavior may be an artifact of truncating the computation of the peak power spectrum at second order. We expect that, as higher powers of $\xi_\ell^{(n)}$ are included in the description of the peak correlation, small scale exclusion will increase and eventually make the large scale shot noise correction sub-Poisson (see, [52], for a rough estimate of the effect).

Before continuing, we note that any local nonlinear biasing introduces constant power at small wave number in addition to the conventional $1/\bar{n}$ shot noise [46,49,51,52,113,120]. At second order, this white noise contribution is always positive as it is given by the first integral of Eq. (79) with $\tilde{\rm b}_{\rm II}(q,q,z_0)$ replaced by $b_{\rm II}$ and δ smoothed on some arbitrary scale R. Therefore, in contrast to the prediction of the peak model, the magnitude of this term is not well defined in local biasing schemes owing to the freedom at filtering the mass density field.

Ignoring the k^0 tail, the next-to-leading term should scale as k^2 since there is no local conservation of momentum (which would otherwise enforce a k^4 behavior). One can easily check that this is indeed the case by noticing that, in the low-k limit,

$$F_1(\mathbf{q}_1)=\frac{k}{q_1}\mu,$$
$$F_1(\mathbf{q}_2)=-\frac{k}{q_1}\mu+\frac{k^2}{q_1^2}(1-2\mu^2)+\mathcal{O}(k^3), \quad (80)$$
$$F_2(\mathbf{q}_1,\mathbf{q}_2)=\frac{k^2}{2q_1^2}\mu\Big[-\mu+\frac{k}{q_1}(1-2\mu^2)\Big]+\mathcal{O}(k^4).$$

As $k\to0$, this implies $F_1(\mathbf{q}_1)+F_1(\mathbf{q}_2)\approx(k/q_1)^2\times(1-2\mu^2)$, so the first two terms of $\tilde{\rm b}_{\rm II}^{\rm E}$ scale as k^2 at a low wave number.

In configuration space, the mode-coupling contribution $\xi_{\rm MC}(r,z)$ changes sign across the acoustic scale $r_0\approx105h^{-1}$ Mpc because the dominant term involves the derivative $\xi_1^{(1/2)}(r)$ of the mass correlation which is positive (negative) to the left (right) of the acoustic peak [121]. This leads to a shift Δr_0 of the inferred acoustic scale towards small scales [121–123]. Typically, a 1% shift in the acoustic scale translates to a fairly substantial $\sim$4% bias in the estimated dark energy equation of state (e.g., [124]).

As shown in Fig. 5, the magnitude of $\xi_{\rm MC}(r,z)$ strongly depends on the first and second order bias parameters.

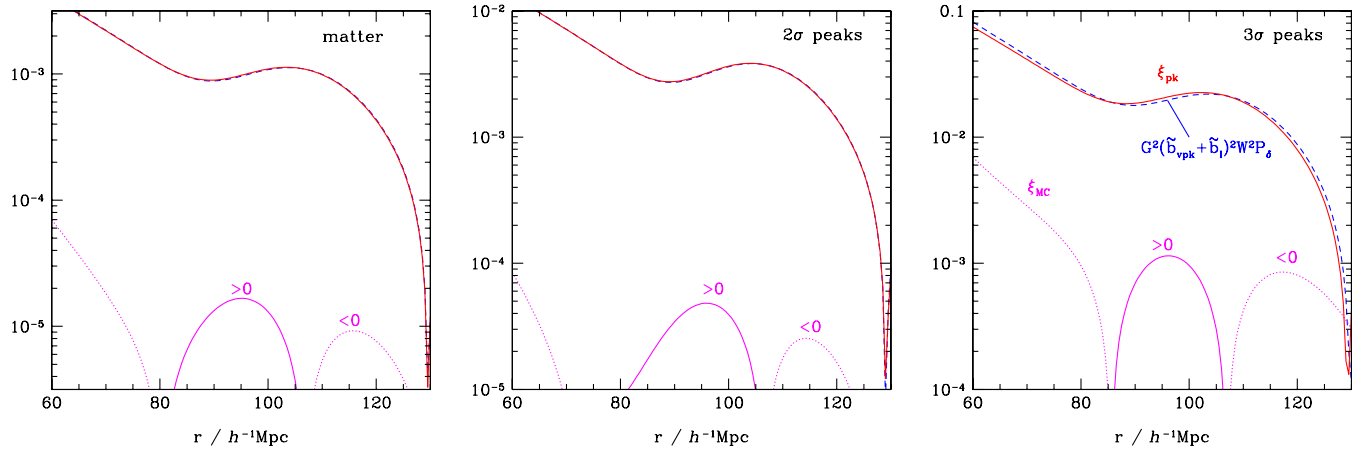

FIG. 5 (color online). Matter and peak correlation functions at the redshift of collapse $z=z_0$ as predicted by Eqs. (62) and (63), respectively. In configuration space, the mode-coupling contribution $\xi_{\rm MC}$ is always positive left to the BAO and negative right to the BAO, so it introduces a shift in the acoustic scale. The magnitude of $\xi_{\rm MC}$ depends strongly on the bias parameters.

For the 2σ peaks, the strength of the mode-coupling contribution relative to the linear term $\xi^{(1)}_{\rm pk}$ is slightly smaller than for the matter ($\sim 1\%$), whereas for the 3σ peaks, it is much larger ($\sim 5\%$) due to the large, positive first and second order bias factors. To estimate the "physical" shift of the acoustic peak generated by mode coupling, we must correct beforehand for the "apparent" shift induced by the convolution with the propagator G^2 (we use the terminology of [121,122]). The latter is usually taken into account in the data analysis and cosmological parameter forecast [124,125]. Here, we will simply ignore the diffusion damping both in the linear and mode-coupling piece and, therefore, calculate the shift of the acoustic peak in the perfectly deconvolved correlation function

$$\xi^{\rm un}_{\rm pk}(r, z_0) = [\tilde{\rm b}^{\rm E}_{\rm I}]^2 \otimes \xi^{(0)}_0(r, z_0) + \xi^{\rm un}_{\rm MC}(r, z_0). \qquad (81)$$

It is important to note that, in the peak model, the physical shift is generated by second order mode-coupling *and* by scale dependence in the linear biases $\tilde{\rm b}_{\rm I}$ and $\tilde{\rm b}_{\rm vpk}$. The relative contribution of the second source of shift, which is not present in the local bias approximation, turns out to be equally important at high and low peak height (this is somewhat counterintuitive since the scale dependence is more pronounced at low ν). For the 2σ and 3σ peaks, we find that the location of the acoustic peak in the linear mass correlation is "physically" shifted by $\Delta r_0 = -0.25$ and $-1.65h^{-1}$ Mpc, respectively. The relative contribution of the first order scale dependence is in both cases $\sim 20\%$ only, mainly because $\tilde{\rm b}_{\rm I}$ does not shift the Fourier phases. We believe these values should change somewhat at second order in the peak displacement (2LPT) since the cross term between the monopole and dipole (of $[\tilde{\rm b}^{\rm E}_{\rm II}]^2$), which generates most of the shift, will not remain the same.

E. Scale dependence across the acoustic peak: Theoretical predictions and comparison with simulations

We will now present results for the evolved peak correlation function and compare them to the autocorrelation of simulated dark matter halos.

To begin, Fig. 6 shows the redshift evolution of the correlation $\xi_{\rm pk}(\nu, R_S, r, z)$ of 2σ and 3σ peaks from the initial conditions at $z = \infty$ (bottom dotted curve) until collapse at $z = z_0 = 0.3$ (top solid curve). The intermediate redshift values are $z = 5$, 2, 1, and 0.5. It should be noted that only the correlation at the collapse epoch can be measured in real data (assuming that the tracers are observed at the epoch their host dark matter halos collapse). For comparison, the bottom and top dashed curves represent the initial and final correlation for the local bias approximation in which $(\tilde{\rm b}_{\rm I}, \tilde{\rm b}_{\rm vpk}) \equiv (b_{\rm I}, 1)$ and the mode-coupling power is given by Eq. (76). As the redshift decreases, gravitational instability generates coherent motions which amplify the large scale amplitude of the peak correlation and random motions which increasingly smear out the initial BAO feature. Although velocity diffusion due to large scale flows is less important for the peaks than for the locally biased tracers (owing to the fact that $\tilde{\rm b}_{\rm vpk} < 1$), the final correlations are noticeably more similar than they were initially. Still, mild differences subsist at $z = z_0$ between the peak and local bias predictions, especially around the baryonic acoustic feature.

In order to quantify these deviations, we take the square root of the ratio between the peak correlation $\xi_{\rm pk}$ [Eq. (63)] and the mass correlation $\xi_{\rm m}$ [Eq. (62)]. Both are consistently evolved at second order with the Zel'dovich approximation. $\sqrt{\xi_{\rm pk}/\xi_{\rm m}}$ measures the scale dependence of peak bias as a function of separation. For illustration purposes,

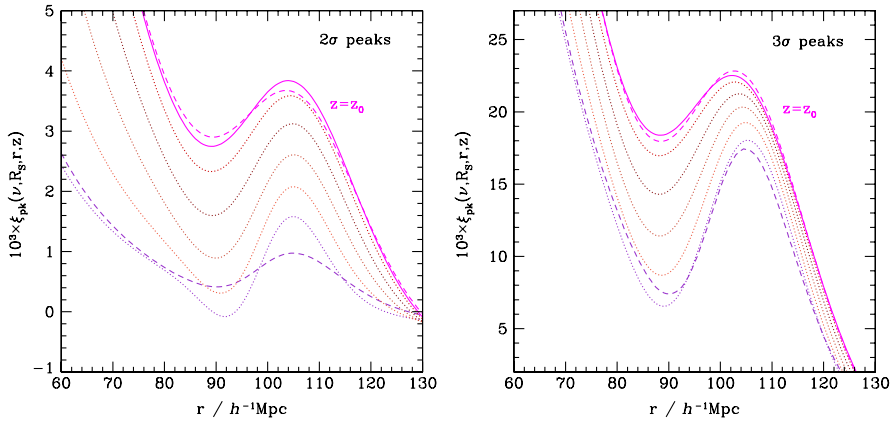

FIG. 6 (color online). Redshift evolution of the correlation of 2σ and 3σ peaks collapsing at $z_0 = 0.3$ as predicted by Eq. (63). The curves from bottom to top represent $\xi_{\rm pk}(\nu, R_S, r, z)$ at redshift $z = \infty$, 5, 2, 1, 0.5 (dotted curves) and $z = z_0$ (solid curve). Only the correlation at the collapse epoch ($z = z_0$) can be measured in real data. For comparison, the dashed curves show the correlation at $z = \infty$ and $z = z_0$ in a local bias approximation (see text).

we normalize this ratio to $\tilde{b}^{E}_{10}(z)$ so that, on scales much larger than the acoustic scale (not shown in this figure), the normalized ratio rapidly converges to unity. Results are shown in Fig. 7 for the 2σ and 3σ peaks. The dashed curve represents the local bias prediction at the collapse redshift $z = z_0$. Across the baryonic acoustic feature, the peak bias exhibits a residual scale dependence of $\sim$5% amplitude. For the 2σ peaks, this scale dependence arises principally from the initial amplification of the BAO contrast (see Fig. 3) whereas, for the 3σ peaks, it is mostly induced by the gravity mode coupling. In stark contrast to the peak model, the local bias approximation predicts negligible scale dependence for 2σ tracers (there is a sharp upturn at $r \simeq 130h^{-1}$ Mpc due to the fact that the zero crossings of $\xi_{\rm pk}$ and $\xi_{\rm m}$ are different). At 3σ however, the discrepancy between both models is relatively smaller because the contribution of the mode coupling generated during gravitational evolution, which is weakly sensitive to the bias factors $\tilde{b}_{01}$, $\tilde{b}_{11}$, and $\tilde{b}_{02}$ in the limit $\nu \gg 1$, dominates the scale dependence of bias. The mode coupling also contributes to suppress the peak bias by 2%–3% at separations $r \sim 60$–$80h^{-1}$ Mpc (see also [11]).

Clearly, a very large simulated volume is required to search for similar scale dependences in the bias of the most massive objects created by gravitational collapse. Hence, we will present measurements of the baryonic acoustic feature in the autocorrelation of halos extracted from a single realization of 2048^3 particles in a cubical box of side $7.68h^{-1}$ Gpc. The simulated volume thus is more than 16 times the Hubble volume. Halos were subsequently found using a friends-of-friends algorithm with a linking length of 0.2 times the mean interparticle distance, leading to a final catalog of more than 15×10^6 halos of mass larger than $\approx 7 \times 10^{13}{\rm M}_\odot/h$ (with 20 particles or more). This N-body run and the associated halo catalog is a part of the MICE simulations project (see [12,126,127], for an exhaustive description of the runs and [128] for publicly available data). A more detailed clustering analysis of this simulation will be presented elsewhere [129].

The filled symbols in Fig. 8 show the measured bias (defined as the squared root ratio of their autocorrelation function to that of the dark matter field) for halos with significance larger than $\nu_t = 2$ and 3. Here, we exceptionally adopted a top hat filter to define the peak significance $\delta_{\rm sc}(z)/\sigma_0$. Therefore, the ν_t threshold corresponds to a mass cut $M_t = 2 \times 10^{14}$ and $7 \times 10^{14}{\rm M}_\odot/h$, respectively, (these halos have at least 35 and 200 particles, respectively). The correlation was computed by direct pair counting using the estimator of [130]. Measurements were done first dividing the full box size of $452h^{-3}$ Gpc^3 into 27 nonoverlapping regions of equal volume. The bias shown is the average over these regions and the error bars those associated with the corresponding variance of the autocorrelation measurements (we actually depict the error on the mean, i.e., $\sigma_b/\sqrt{27}$).

From Fig. 8, it is clear that while the $>3\sigma$ halo sample is too sparse to furnish useful information, the correlation of $>2\sigma$ halos shows unambiguous evidence for a scale dependence of bias of $\sim$5% relative magnitude around the baryonic acoustic feature. To the best of our knowledge, this is the first time that such a scale dependence is reported from N-body simulations (see, however, [131]).

To compare predictions from the peak model with the measurements from the N-body simulation, we remark that the set of dark matter halos with significance larger than ν_t includes all halos of mass larger than some threshold value M_t. As an attempt to reproduce the number counts of halos above a mass cut M_t, one may want to smooth the mass

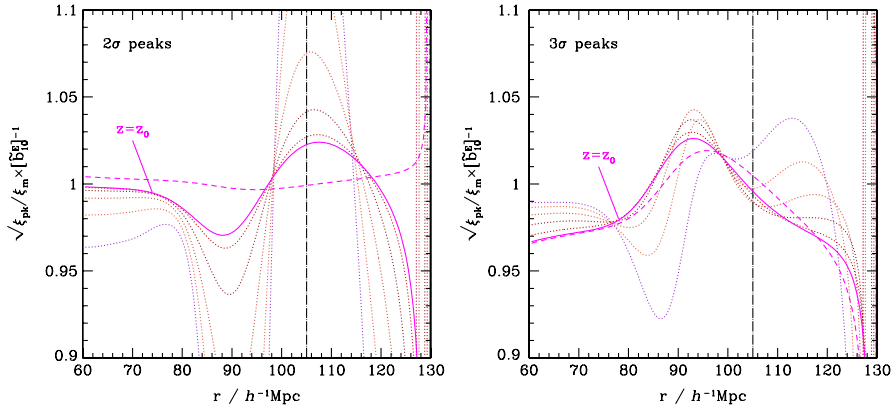

FIG. 7 (color online). Ratio of the peak correlation $\xi_{\rm pk}(\nu, R_S, r, z)$ to the mass correlation $\xi_{\rm m}(r, z)$ at redshift $z = \infty$, 5, 2, 1, 0.5, and z_0 (at separation $60h^{-1}$ Mpc, curves from bottom to top in the left panel). For convenience, the ratio is normalized to $\tilde{b}^{E}_{10}(z)$ so that the results approach unity as $r \to \infty$. The dashed curve shows the quantity at redshift of collapse $z_0 = 0.3$ in the local bias approximation. At $z = z_0$, the bias of density peaks exhibits a $\sim$5% scale dependence across the baryonic acoustic feature. For the 2σ peaks however, most of this scale dependence is caused by the initial amplification of the BAO contrast relative to that of the mass (see Fig. 3) whereas, for 3σ peaks, this scale dependence is mainly generated by the gravity mode coupling. The vertical line denotes the position of the linear BAO feature.

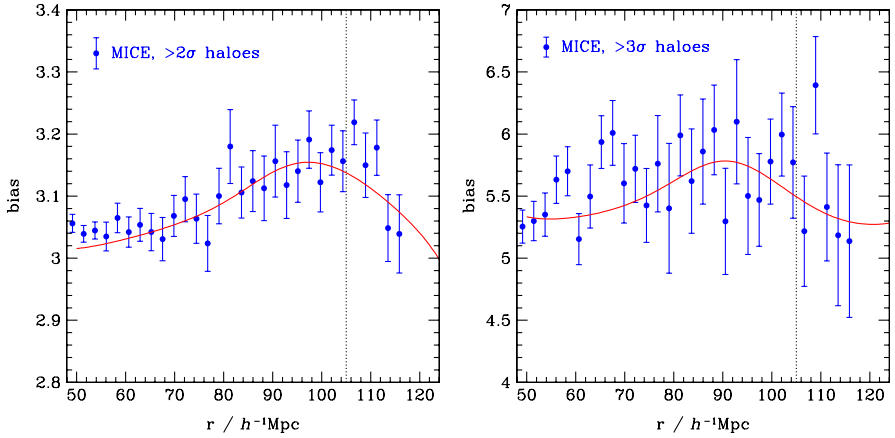

FIG. 8 (color online). Scale dependence of bias around the baryonic acoustic feature of the halo autocorrelation. Filled symbols with error bars represent the measurement for $>2\sigma$ and $>3\sigma$ halos extracted from the MICE N-body simulation (see text) at $z = 0$. The corresponding mass cuts are $M > 1.2 \times 10^{14}$ and $M > 7.0 \times 10^{14} \mathrm{M}_{\odot}/h$, respectively. The solid curve corresponds to a prediction based on the peak model with linear peak-background split bias b_{I} similar to that of the simulated halo catalogs. For the low mass cut, the threshold height is $\nu_t = 2.15$ whereas, for the high mass cut, it is $\nu_t = 3$ [see Eq. (82)]. A vertical line denotes the position of the linear BAO feature.

density field with a range of filter radii R_S and identify peaks of height $\delta_{\mathrm{sc}}(z_0)/\sigma_0(R_S)$ with virialized halos of mass $M = M_S$. We adopt a sharp threshold even though the actual selection function $w(M|\nu, R_S)$, which gives the probability that a peak of height ν in the linear density field smoothed on scale R_S forms a virialized halo of mass M, may be fairly smooth (see, e.g., [36] for more realistic selection functions). Consequently, we approximate the correlation function of dark matter halos as

$$\xi_{\mathrm{h}}(r, z) = \frac{1}{N_{\mathrm{h}}(\nu_t)} \int_{\nu_t}^{\infty} d\nu \bar{n}_{\mathrm{pk}}(\nu, R_S(\nu)) \xi_{\mathrm{pk}}(r, z)$$
$$\text{where } N_{\mathrm{h}}(\nu_t) = \int_{\nu_t}^{\infty} d\nu \bar{n}_{\mathrm{pk}}(\nu, R_S(\nu)). \tag{82}$$

Here, $R_S(\nu)$ is the filtering radius at which $\delta_{\mathrm{sc}}(z)/\sigma_0(R_S, z) \equiv \nu$. The fraction of mass $f(\nu)d\nu$ (or mass function) in peaks of height ν thus is

$$f(\nu)d\nu \equiv (M_S/\bar{\rho})\bar{n}_{\mathrm{pk}}(\nu, R_S(\nu))d\nu, \tag{83}$$

and $M_S/\bar{\rho} = (2\pi)^{3/2} R_S^3$ is the mass enclosed in the Gaussian filter. In our simulation, the measured cumulative number density for the $>2\sigma$ and $>3\sigma$ halos is $N_{\mathrm{h}} \approx 1.45 \times 10^{-5}$ and $\approx 5.10 \times 10^{-7} h^3$ Mpc^{-3}, respectively. For these threshold heights, Eq. (82) predicts slightly larger densities, i.e., $N_{\mathrm{h}} \approx 1.77 \times 10^{-5}$ and $\approx 8.40 \times 10^{-7}$. To further improve the agreement with the simulated mass function, one shall ensure that a peak of height δ_{sc} on a smoothing scale R_S is not embedded in a region of height δ_{sc} on any larger smoothing scale. Carefully accounting for clouds in clouds in the peak formalism is a nontrivial problem even though, in a first approximation, we may simply enforce that the height of the peak be less than δ_{sc} on scale $R_S + dR_S$ [132,133]. Figure 1 of [80] demonstrates that this substantially improves the agreement with the measured halo abundances at $\nu \gtrsim 3$ but underestimates the counts at $\nu \lesssim 2$. Furthermore, adding this extra constraint also modifies the peak bias factors and, possibly, the $\tilde{\mathrm{b}}_{\mathrm{II}}$-independent terms in Eq. (22). For these reasons, we hereafter stick to the naive albeit reasonably good approximation Eq. (82), and defer a more detailed modeling of the halo mass function to a future work.

The predicted halo correlation Eq. (82) is shown in Fig. 8 as the solid curve. For the low mass cut (left panel), the theoretical prediction assumes in fact $\nu_t = 2.15$, which furnishes a better match to the linear, scale-independent halo bias measured in the simulation. For the high mass cut, $\nu_t = 3$ as in the halo sample. Overall, since we consider the clustering of high peaks at $z = 0$, the scale dependence induced by gravity mode coupling is important. Consequently, all the theoretical curves exhibit a broad peak structure centered at $r \sim 90$–$100h^{-1}$ Mpc. By contrast, Lagrangian peak biasing solely would induces a sombrero-like scale dependence with a maximum close to the acoustic scale (see Fig. 7). Focusing on the low-ν halo sample, it is difficult to assess the quality of the fit without knowing the data covariances. Nevertheless, the measurements strongly suggest that our prediction based on the peak model is a very good approximation. An improved description of peak motions, collapse, and mass function will somewhat modify the theoretical curve but, because the predicted abundances are close to the measured ones, we believe the change should not be dramatic. It will be interesting to search for a similar effect in the clustering of low mass and/or high redshift halos, for which the contribution of the peak biasing should be larger.

V. CONCLUSION

In standard approaches to clustering, biased tracers are treated as a continuous field. Their overabundance is commonly assumed to be a local function of the smoothed mass density field. Gravitational motions can be straightforwardly included through a spherical collapse prescription or perturbation theory to predict galaxy and dark matter halo correlation functions.

In the peak model where the tracers form a well-behaved point process, the peak constraint renders the calculation almost prohibitive. For this reason, all analytic studies of peak correlations thus far [36,38,72,74,75] have obtained results which are strictly valid at leading order and in the (Gaussian) initial conditions solely.

To make predictions which can be directly compared with observational data or the outcome of numerical experiments, we have calculated the correlation of density peaks in the Gaussian linear mass field and computed its redshift evolution, consistently within the Zel'dovich approximation. We have then used our model to explore the effect of peak biasing on the baryonic acoustic feature (BAO).

In the first part of this work (Sec. III), we obtained a compact expression for the correlation of initial density maxima up to second order. We demonstrated that the k-independent pieces of the peak bias factors ($\tilde{b}_{10}$, $\tilde{b}_{20}$, etc.) are equal to the peak-background split biases derived from the peak number density $\bar{n}_{\rm pk}(\nu, R_S)$, even though the later is not universal. Furthermore, we also showed that the peak-background split approach can be applied to derive the k-dependent bias factors: $\tilde{b}_{01}$, $\tilde{b}_{11}$, etc. This is especially interesting as it illustrates how the peak-background split should be implemented in any other model where the abundance of tracers depends on variables other than the local mass overdensity. So far however, we have not found a straightforward derivation of the terms linear and independent of $\tilde{b}_{\rm II}$ in the initial peak correlation.

In the second part (Sec. IV), we computed the gravitational evolution of the peak correlation under the assumption that the peaks locally flow with the matter. For simplicity, we considered the Zel'dovich approximation in which the initial density peaks execute a motion along straight line trajectories. Using the approach laid out in [104], we showed that the evolved peak correlation function can be expressed, like in renormalized perturbation theory [101,109] as the sum of a linear contribution dominant at large scales, and a nonlinear contribution which describes the generation of power at a given scale due to the coupling between two modes (at one-loop level in our calculation) at different scales. The linear term can be easily derived from a continuity equation argument for the average peak overabundance if one includes the peak velocity bias $\tilde{b}_{\rm vpk}$. If one ignores the peak constraint, then our results reduce to those predicted by previous local Lagrangian bias calculations. Thus, the peak model is a useful generalization of this model. In particular, local Lagrangian biasing provides a good approximation to our results only in the limit of large peak height and separation, and only if velocity bias is ignored.

We applied our model to predict the shape and amplitude of the baryon acoustic oscillation in the correlation of peaks at the epoch of collapse. Most of the initially strong scale-dependent bias across the BAO reported in [74] is washed out by velocity diffusion, which manifests itself as an exponential damping kernel. Still, for 2σ peaks collapsing at $z = 0.3$, our model predicts a residual $\sim 5\%$ scale-dependent bias which should be quite asymmetric about the acoustic scale. This prediction stands in stark contrast to the negligible scale dependence predicted by the local bias approximation for the same peak height. For 3σ peaks, the mode-coupling power dominates so that the predictions of both models are quite similar: the bias exhibits a broad bump at distances $r \sim 90$–$100h^{-1}$ Mpc smaller than the acoustic scale.

We then measured the clustering of massive halos extracted from a very large N-body simulation. For halos with mass $M > 2 \times 10^{14} {\rm M}_\odot/h$, which corresponds roughly to a peak significance $\nu > 2$, the large scale bias shows strong evidence for a $\sim 5\%$ scale dependence around the baryonic acoustic feature. We have tried to reproduce this measurement using our model combined with a simple prescription to account for the observed halo counts. Our prediction, which is a weighted contribution of the scale dependence induced by Lagrangian peak bias and gravity mode coupling, is in good agreement with the data. However, the model slightly overestimates the cumulative halo abundance.

Clearly, there are a number of important missing ingredients, such as the second order contribution to the peak displacement and a better treatment of nonlinear collapse, which must be accounted to achieve more accurate predictions. In this regards, all the machinery developed for the local bias scheme can be applied to the peak model, with the caveat that the bias factors, including the linear one, are k dependent. Numerical simulations will be essential to calibrate some of the parameters, such as the diffusion scale appearing in the propagator (see, e.g., [134]) or determine what is the actual shape of the filtering kernel. Future work should also include redshift-space distortions, non-Gaussian initial conditions, and massive neutrino species, and investigate the effect of small scale exclusion and stochasticity on the peak power spectrum and correlation function.

ACKNOWLEDGMENTS

V. D. wishes to thank the Center for Cosmology and Particle Physics at New York University, the Center for Particle Cosmology at University of Pennsylvania, and the Institut de Ciències de l'Espai at Universitat Autònoma de Barcelona for their hospitality during the redaction of this

paper. V. D., R. S., and R. K. S. would like to thank the Centro de Ciencias de Benasque ''Pedro Pascual'' and the organizers of the Benasque 2010 workshop on modern cosmology for a very enjoyable and productive meeting. We are grateful to Takahiko Matsubara for correspondence and to Cristiano Porciani, Eniko Regös, and Robert Smith for interesting discussions. We acknowledge the use of numerical simulations from the MICE collaboration (http://www.ice.cat/mice). We acknowledges partial support from FK UZH 57184001 and the Swiss National Foundation under Contract No. 200021-116696/1 (V. D.); from Spanish Ministerio de Ciencia e Innovacion (MICINN), projects AYA2009-13936 Consolider-Ingenio CSD2007-00060, Juan de la Cierva program, and Project No. 2009SGR1398 from Generalitat de Catalunya (M. C.); from NASA NNX10AI71GS02 and NSF AST-0607747 (R. S.) and from NSF AST-0908241 (R. K. S.).

APPENDIX A: CORRELATION OF INITIAL DENSITY PEAKS AT THE SECOND ORDER

In this Appendix, we present the derivation of the correlation of initial density maxima Eq. (20) at the second order in the mass correlation $\xi_0^{(0)}$ and its derivatives.

1. SVT (scalar-vector-tensor) decomposition of the covariance matrix

Computing $\xi_{\rm pk}(r)$ requires knowledge of the joint probability distribution $P_2(\mathbf{y}(\mathbf{x}_1), \mathbf{y}(\mathbf{x}_2); \mathbf{r})$ where $\mathbf{y}^\top = (\eta_i, \nu, \zeta_A)$ is a ten-dimensional vector whose components ζ_A, $A = 1, \cdots, 6$ symbolize the independent entries $ij = 11$, 22, 33, 12, 13, 23 of ζ_{ij}.

We shall proceed in a way analogous to cosmological perturbation theory (e.g. [135–142] and references therein), and decompose the variables into irreducible components according to their transformation properties under spatial rotations. Although this decomposition is not adequate for imposing the peak constraint, it greatly simplifies the structure of the covariance matrix $C(r)$ of the joint probability density P_2 (see, e.g., [72]). In this work, we choose the *covariant* helicity basis $(\mathbf{e}_+, \hat{\mathbf{r}}, \mathbf{e}_-)$ as reference frame, where

$$\mathbf{e}_+ \equiv \frac{i\hat{\mathbf{e}}_\phi - \hat{\mathbf{e}}_\theta}{\sqrt{2}}, \qquad \hat{\mathbf{r}} \equiv \mathbf{r}/r, \qquad \mathbf{e}_- \equiv \frac{i\hat{\mathbf{e}}_\phi + \hat{\mathbf{e}}_\theta}{\sqrt{2}} \tag{A1}$$

and $\hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_\phi$ are orthonormal vectors in spherical coordinates (θ, ϕ). The orthogonality relations between these vectors are $\mathbf{e}_\pm \cdot \mathbf{e}_\pm = \hat{\mathbf{r}} \cdot \hat{\mathbf{r}} = 1$ and $\mathbf{e}_+ \cdot \mathbf{e}_- = \mathbf{e}_\pm \cdot \hat{\mathbf{r}} = 0$, where the inner product between two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined as $\mathbf{u} \cdot \mathbf{v} \equiv u_i \bar{v}_i \equiv u_i v^i$. An overline will denote complex conjugation throughout this section. The property $\overline{\mathbf{u} \cdot \mathbf{v}} = \mathbf{v} \cdot \mathbf{u}$ follows from our definition of the inner product. For the first derivative of the density field, the SVT decomposition is

$$\boldsymbol{\eta} \equiv \boldsymbol{\eta}^{(S)} + \boldsymbol{\eta}^{(V)} \equiv \eta^{(0)}\hat{\mathbf{r}} + \eta^{(+1)}\hat{\mathbf{e}}_+ + \eta^{(-1)}\mathbf{e}_-. \tag{A2}$$

Here, $\eta^{(0)} \equiv \boldsymbol{\eta} \cdot \hat{\mathbf{r}}$ is the *contravariant* component of an irrotational vector (spin-0), $\hat{\mathbf{r}} \wedge \boldsymbol{\eta} = 0(S)$, whereas $\eta^{(\pm 1)} \equiv \boldsymbol{\eta} \cdot \mathbf{e}_\pm$ are the two independent *contravariant* components of the transverse vector $\boldsymbol{\eta}^{(V)}$ (spin-1), $\boldsymbol{\eta}^{(V)} \cdot \hat{\mathbf{r}} = 0$. The correlation properties of $\eta^{(0)}$ and $\eta^{(\pm 1)}$ can be obtained by projecting out the scalar and vector parts of the correlation of the Cartesian components η_i which, because of statistical isotropy and symmetry in i, j, is of the form

$$\langle \eta_i(\mathbf{x}_1)\eta_j(\mathbf{x}_2)\rangle = H_\parallel(r)\hat{r}_i\hat{r}_j + H_\perp(r)P_{ij}, \tag{A3}$$

where

$$P_{ij} \equiv (\mathbf{I}_3 - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}})_{ij} = e_{+i}\bar{e}_{+j} + e_{-i}\bar{e}_{-j} \tag{A4}$$

is the projection operator onto the plane perpendicular to $\hat{\mathbf{r}}$, and $\mathbf{I}_n$ is the $n \times n$ identity matrix. The vectors $\bar{\mathbf{e}}_\pm$ will be called *contravariant* basis vectors. They satisfy $\mathbf{e}_\pm \cdot \bar{\mathbf{e}}_\pm = e_{\pm i}e_{\pm i} \equiv 0$, which follows from the relations $\bar{e}^i_\pm \equiv e_{\pm i}$ and $\bar{e}_{\pm i} \equiv e^i_\pm$ between contravariant and covariant basis. The *covariant* components of $\boldsymbol{\eta}^{(V)}$ thus are $\bar{\eta}^{\pm 1} \equiv \boldsymbol{\eta} \cdot \bar{\mathbf{e}}_\pm$, whereas $\bar{\eta}^{(0)} = \eta^{(0)}$. The functions $H_\parallel = (\xi_0^{(1)} - 2\xi_2^{(1)})/3\sigma_1^2$ and $H_\perp = (\xi_0^{(1)} + \xi_2^{(1)})/3\sigma_1^2$ are radial and transverse correlations (see, e.g., [143]). Using $\eta^{(\pm 1)} = \eta_i e^i_\pm$ and $\bar{\eta}^{(\pm 1)} = \eta_i \bar{e}^i_\pm$, we find that the correlations of the spin-0 and spin-1 components read

$$\langle \eta_1^{(0)}\eta_2^{(0)}\rangle = H_\parallel(r), \qquad \langle \eta_1^{(\pm 1)}\bar{\eta}_2^{(\pm 1)}\rangle = H_\perp(r), \qquad \langle \eta_1^{(\pm 1)}\bar{\eta}_2^{(\mp 1)}\rangle = 0. \tag{A5}$$

Here and henceforth, the subscripts ''1'' and ''2'' will denote variables evaluated at position $\mathbf{x}_1$ and $\mathbf{x}_2$ for shorthand convenience. Let us now consider the symmetric tensor ζ_{ij}. Writing ζ_{ij} as the sum of a traceless symmetric matrix $\tilde{\zeta}_{ij}$ and its orthogonal complement, $\zeta_{ij} \equiv \tilde{\zeta}_{ij} - (u/3)\mathbf{I}_3$ where $u \equiv -\mathrm{tr}\zeta = -\zeta^i_i$, a suitable parametrization of the SVT decomposition for ζ_{ij} is (e.g., [144])

$$\zeta_{ij} \equiv -\tfrac{1}{3}u\delta_{ij} + S_{ij}\zeta^{(S)} + \sqrt{\tfrac{1}{3}}(\zeta_i^{(V)}\hat{r}_j + \zeta_j^{(V)}\hat{r}_i) + \sqrt{\tfrac{2}{3}}\zeta_{ij}^{(T)}. \tag{A6}$$

The functions $u \equiv -\mathrm{tr}\zeta = -\zeta^i_i$ and $\zeta^{(S)} \equiv \zeta^{(0)}$ are the longitudinal and transverse spin-0 modes, $\zeta_i^{(V)}$ are the components of a spin-1 vector, $\boldsymbol{\zeta}^{(V)} \cdot \hat{\mathbf{r}} = 0$, and $\zeta_{ij}^{(T)}$ is a symmetric, traceless, transverse (spin-2) tensor, $\delta^{ij}\zeta_{ij}^{(T)} = \zeta_{ij}^{(T)}\hat{r}^j = 0$. Explicit expressions for these functions are

$$\zeta^{(S)} \equiv \tfrac{3}{2}S^{lm}\zeta_{lm} = \tfrac{1}{2}(3\hat{r}^l\hat{r}^m - \delta^{lm})\zeta_{lm} \tag{A7}$$

$$\zeta_i^{(V)} \equiv \sqrt{3}V_i^{lm}\zeta_{lm} = \sqrt{3}(\delta^l_i - \hat{r}_i\hat{r}^l)\hat{r}^m\zeta_{lm} \tag{A8}$$

$$\zeta_{ij}^{(T)} \equiv \sqrt{\tfrac{3}{2}}T_{ij}^{lm}\zeta_{lm} = \sqrt{\tfrac{3}{2}}\left(P_i^lP_j^m - \tfrac{1}{2}P_{ij}P^{lm}\right)\zeta_{lm}. \tag{A9}$$

We have introduced factors of $\sqrt{1/3}$ and $\sqrt{2/3}$ in the decomposition (A6) such that the zero-point moments of the spin-0, spin-1, and spin-2 tensor variables all equal 1/5 [see Eqs. (A15)–(A17) below]. Note that $P_i^j \equiv e_{+i}e_+^j + e_{-i}e_-^j$ and $P^{ij} \equiv e_+^i\bar{e}_+^j + e^i_-\bar{e}^j_-$ to ensure nilpotency of the operator P. The two-point correlation function of any isotropic, symmetric tensor field ζ_{ij} is of the form (e.g., [145])

$$\begin{aligned}\langle\zeta_{ij}(\mathbf{x}_1)\zeta_{lm}(\mathbf{x}_2)\rangle = {} & Z_1(r)\hat{r}_i\hat{r}_j\hat{r}_l\hat{r}_m + Z_2(r)(\hat{r}_i\hat{r}_l\delta_{jm} \\ & + \hat{r}_i\hat{r}_m\delta_{jl} + \hat{r}_j\hat{r}_l\delta_{im} + \hat{r}_j\hat{r}_m\delta_{il}) \\ & + Z_3(r)(\hat{r}_i\hat{r}_j\delta_{lm} + \hat{r}_l\hat{r}_m\delta_{ij}) \\ & + Z_4(r)\delta_{ij}\delta_{lm} + Z_5(r)(\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl}).\end{aligned} \tag{A10}$$

In the case of the tensor $\zeta_{ij} = \partial_i\partial_j\delta$ considered here, we have $Z_2 = Z_3$ and $Z_4 = Z_5$. To project out the transverse spin-0, spin-1, and spin-2 parts of the 4-rank correlation tensor Eq. (A10), we act with the scalar, vector, and tensor projection operators S_{ab}, V_a^{bc}, T_{ab}^{cd} on $\langle\zeta_{ij}(\mathbf{x}_1)\zeta_{lm}(\mathbf{x}_2)\rangle$ and obtain

$$\langle\zeta^{(S)}(\mathbf{x}_1)\zeta^{(S)}(\mathbf{x}_2)\rangle = Z_1(r) + 4Z_3(r) + 3Z_5(r) \tag{A11}$$

$$\langle\zeta_i^{(V)}(\mathbf{x}_1)\zeta_j^{(V)}(\mathbf{x}_2)\rangle = 3P_{ij}[Z_3(r) + Z_5(r)] \tag{A12}$$

$$\langle\zeta_{ij}^{(T)}(\mathbf{x}_1)\zeta_{lm}^{(T)}(\mathbf{x}_2)\rangle = 3T_{ijlm}Z_5(r), \tag{A13}$$

where

$$\begin{aligned}Z_1 + 4Z_3 + 3Z_5 &= \frac{1}{\sigma_2^2}\left(\frac{1}{5}\xi_0^{(2)} - \frac{2}{7}\xi_2^{(2)} + \frac{18}{35}\xi_4^{(2)}\right), \\ Z_3 + Z_5 &= \frac{1}{\sigma_2^2}\left(\frac{1}{15}\xi_0^{(2)} - \frac{1}{21}\xi_2^{(2)} - \frac{4}{35}\xi_4^{(2)}\right), \\ Z_5 &= \frac{1}{\sigma_2^2}\left(\frac{1}{15}\xi_0^{(2)} + \frac{2}{21}\xi_2^{(2)} + \frac{1}{35}\xi_4^{(2)}\right).\end{aligned} \tag{A14}$$

The correlation for the spin-0 variable $\zeta^{(0)}$ thus is

$$\langle\zeta_1^{(0)}\zeta_2^{(0)}\rangle = Z_1 + 4Z_3 + 3Z_5. \tag{A15}$$

Taking the 2 independent components of $\boldsymbol{\zeta}^{(V)}$ and their complex conjugate as being $\zeta^{(\pm1)} \equiv \boldsymbol{\zeta}^{(V)}\cdot\mathbf{e}_\pm = \sqrt{3}e_\pm^i\hat{r}^j\zeta_{ij}$ and $\bar{\zeta}^{(\pm1)} \equiv \boldsymbol{\zeta}^{(V)}\cdot\bar{\mathbf{e}}_\pm = \sqrt{3}\bar{e}_\pm^i\hat{r}^j\zeta_{ij}$ yields

$$\langle\zeta_1^{(\pm1)}\bar{\zeta}_2^{(\pm1)}\rangle = 3(Z_3 + Z_5), \qquad \langle\zeta_1^{(\pm1)}\bar{\zeta}_2^{(\mp1)}\rangle = 0. \tag{A16}$$

Similarly, we choose $\zeta^{(\pm2)} \equiv \zeta_{ij}^{(T)}e_\pm^ie_\pm^j = \sqrt{3/2}e_\pm^ie_\pm^j\zeta_{ij}$ and $\bar{\zeta}^{(\pm2)} \equiv \zeta_{ij}^{(T)}\bar{e}_\pm^i\bar{e}_\pm^j = \sqrt{3/2}\bar{e}_\pm^i\bar{e}_\pm^j\zeta_{ij}$ for the two independent polarizations and their complex conjugate, respectively. The correlation properties of these variables are

$$\langle\zeta_1^{(\pm2)}\bar{\zeta}_2^{(\pm2)}\rangle = 3Z_5, \qquad \langle\zeta_1^{(\pm2)}\bar{\zeta}_2^{(\mp2)}\rangle = 0. \tag{A17}$$

As expected, these are the only spin-2 degrees of freedom since $\zeta_{ij}^{(T)}e_+^i\bar{e}_-^j \equiv 0$. Therefore, the second rank tensor ζ_{ij} is fully characterized by the set of variables $\{u, \zeta^{(0)}, \zeta^{(\pm1)}, \zeta^{(\pm2)}\}$. Furthermore, all the correlations are real despite the fact that some of these variables are complex. We note that, in the particular case $\hat{\mathbf{r}} \equiv \hat{\mathbf{z}}$, our variables are directly related to the variables y_{lm}^n defined in [72], who work in the spherical basis

$$\left(\frac{i\hat{\mathbf{y}} - \hat{\mathbf{x}}}{\sqrt{2}}, \hat{\mathbf{z}}, \frac{i\hat{\mathbf{y}} + \hat{\mathbf{x}}}{\sqrt{2}}\right). \tag{A18}$$

The transformation from the helicity to the spherical basis vectors is performed by the rotation operator $D^1(0, \theta, \phi)$ (whose matrix elements are the Wigner D functions $D_{mm'}^{\ell=1}(0, \theta, \phi)$). For the traceless tensor $\tilde{\zeta}_{ij}$, the spin-0, spin-1, and spin-2 components in the spherical basis simplify to

$$\begin{aligned}\zeta^{(0)} &= \tfrac{1}{2}(2\zeta_{33} - \zeta_{11} - \zeta_{22}), \\ \zeta^{(\pm1)} &= \mp\sqrt{\tfrac{3}{2}}(\zeta_{13} \pm i\zeta_{23}), \\ \zeta^{(\pm2)} &= \sqrt{\tfrac{3}{8}}(\zeta_{11} - \zeta_{22} \pm 2i\zeta_{12}).\end{aligned} \tag{A19}$$

The relationship between the two sets of variables thus is $\zeta^{(m)}(\mathbf{x}) = y_{2m}^0(\mathbf{x})/\sqrt{5}$ [see Eq. (22) of [72]]. For the gradient $\boldsymbol{\eta}(\mathbf{x})$ of the density field, the correspondence is $\eta^{(m)}(\mathbf{x}) = y_{1m}^0(\mathbf{x})/\sqrt{3}$.

The 20-dimensional covariance matrix $\mathrm{C}(r) \equiv \langle\mathbf{y}\mathbf{y}^\dagger\rangle$, where $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ and $\mathbf{y}^\dagger$ is its conjugate transpose, describes the correlations of the fields at positions $\mathbf{x}_1$ and $\mathbf{x}_2$. For simplicity, we will assume in what follows that these are smoothed on the same mass scale. $\mathrm{C}(r)$ may be partitioned into four 10×10 block matrices, the zero-point contribution M in the top left and bottom right corners, and the cross correlation matrix $\mathrm{B}(r)$ and its transpose in the bottom left and top right corners, respectively. In terms of the variables $\mathbf{y}_i = (\eta_i^{(0)}, \nu_i, u_i, \zeta_i^{(0)}, \eta_i^{(+1)}, \zeta_i^{(+1)}, \eta_i^{(-1)}, \zeta_i^{(-1)}, \zeta_i^{(+2)}, \zeta_i^{(-2)})$, M and B have the block-diagonal decomposition $\mathrm{M} = \mathrm{diag}(\mathrm{M}^{(0)}, \mathrm{M}^{(1)}, \mathrm{M}^{(1)}, \mathrm{M}^{(2)}, \mathrm{M}^{(2)})$ and $\mathrm{B} = \mathrm{diag}(\mathrm{B}^{(0)}, \mathrm{B}^{(1)}, \mathrm{B}^{(1)}, \mathrm{B}^{(2)}, \mathrm{B}^{(2)})$. Explicitly,

$$\begin{aligned}\mathrm{M}^{(0)} &= \begin{pmatrix} 1/3 & 0 & 0 & 0 \\ 0 & 1 & \gamma_1 & 0 \\ 0 & \gamma_1 & 1 & 0 \\ 0 & 0 & 0 & 1/5 \end{pmatrix}, \\ \mathrm{M}^{(1)} &= \begin{pmatrix} 1/3 & 0 \\ 0 & 1/5 \end{pmatrix}, \\ \mathrm{M}^{(2)} &= \frac{1}{5},\end{aligned} \tag{A20}$$

and

$$\mathrm{B}^{(0)} = \begin{pmatrix} \frac{1}{3\sigma_1^2}(\xi_0^{(1)} - 2\xi_2^{(1)}) & -\frac{1}{\sigma_0\sigma_1}\xi_1^{(1/2)} & -\frac{1}{\sigma_1\sigma_2}\xi_1^{(3/2)} & -\frac{1}{5\sigma_1\sigma_2}(3\xi_3^{(3/2)} - 2\xi_1^{(3/2)}) \\ \frac{1}{\sigma_0\sigma_1}\xi_1^{(1/2)} & \frac{1}{\sigma_0^2}\xi_0^{(0)} & \frac{1}{\sigma_0\sigma_2}\xi_0^{(1)} & \frac{1}{\sigma_0\sigma_2}\xi_2^{(1)} \\ \frac{1}{\sigma_1\sigma_2}\xi_1^{(3/2)} & \frac{1}{\sigma_0\sigma_2}\xi_0^{(1)} & \frac{1}{\sigma_2^2}\xi_0^{(2)} & \frac{1}{\sigma_2^2}\xi_2^{(2)} \\ \frac{1}{5\sigma_1\sigma_2}(3\xi_3^{(3/2)} - 2\xi_1^{(3/2)}) & \frac{1}{\sigma_0\sigma_2}\xi_2^{(1)} & \frac{1}{\sigma_2^2}\xi_2^{(2)} & \frac{1}{5\sigma_2^2}(\xi_0^{(2)} - \frac{10}{7}\xi_2^{(2)} + \frac{18}{7}\xi_4^{(2)}) \end{pmatrix}, \tag{A21}$$

$$\mathrm{B}^{(1)} = \begin{pmatrix} \frac{1}{3\sigma_1^2}(\xi_0^{(1)} + \xi_2^{(1)}) & -\frac{\sqrt{3}}{5\sigma_1\sigma_2}(\xi_1^{(3/2)} + \xi_3^{(3/2)}) \\ \frac{\sqrt{3}}{5\sigma_1\sigma_2}(\xi_1^{(3/2)} + \xi_3^{(3/2)}) & \frac{1}{5\sigma_2^2}(\xi_0^{(2)} - \frac{5}{7}\xi_2^{(2)} - \frac{12}{7}\xi_4^{(2)}) \end{pmatrix} \tag{A22}$$

$$\mathrm{B}^{(2)} = \frac{1}{\sigma_2^2}\left(\frac{1}{5}\xi_0^{(2)} + \frac{2}{7}\xi_2^{(2)} + \frac{3}{35}\xi_4^{(2)}\right). \tag{A23}$$

The second and third entries of the first row of $\mathrm{B}^{(0)}$ are the correlations of $\eta_2^{(0)}$ with ν_1 and u_1, respectively. As expected, these correlations are negative if $\nu_1 > 0$ or $u_1 > 0$ since, in this case, the line-of-sight derivative is preferentially directed towards $\mathbf{x}_1$. The determinant of $C(r)$ reads as

$$\mathrm{detC}(r) = \mathrm{detM}^{(0)}\,\mathrm{det}[\mathrm{M}^{(0)} - \mathrm{B}^{(0)\top}(\mathrm{M}^{(0)})^{-1}\mathrm{B}^{(0)}] \times \prod_{s=1,2}(\mathrm{detM}^{(s)})^2\,\mathrm{det}[\mathrm{M}^{(s)} - \mathrm{B}^{(s)\top}(\mathrm{M}^{(s)})^{-1}\mathrm{B}^{(s)}]^2. \tag{A24}$$

It is worth noticing that, although $\mathrm{C}(r)$ does not depend on the direction $\hat{\mathbf{r}}$ of the separation vector $\mathbf{r}$, it is *not* equal to the angular average covariance matrix $\hat{\mathrm{C}}(r) \equiv (1/4\pi) \times \int d\Omega_{\hat{\mathbf{r}}}\mathrm{C}(\mathbf{r})$. The latter follows upon setting $\xi_j^{(n)} \equiv 0$ whenever $j \neq 0$ in Eqs. (A20) and (A21). Furthermore, whereas the two-point probability distribution $P_2(\mathbf{y}_1, \mathbf{y}_2; \mathbf{r})$ associated to the correlation matrix $\mathrm{C}(r)$ cannot be easily expressed in closed form, the joint probability density $\hat{P}_2(\mathbf{y}_1, \mathbf{y}_2; r)$ of covariance $\hat{\mathrm{C}}(r)$ may be exactly written as the product

$$\hat{P}_2(\mathbf{y}_1, \mathbf{y}_2; r) = \hat{P}_2(\nu_1, u_1, \nu_2, u_2; r)\hat{P}_2(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2; r)\hat{P}_2(\tilde{\zeta}_1, \tilde{\zeta}_2; r). \tag{A25}$$

This factorization property reflects the fact that, upon angle averaging, the $\ell = 0$ (ν_i, u_i), $\ell = 1$ (η_i) and $\ell = 2$ ($\tilde{\zeta}_{ij}$) representations of SO(3) decouple from each other. Consequently, it should be possible to cast the two-point probability densities in terms of rotational invariants such as the scalar product and the matrix trace. After some manipulations, we find the joint probability density for the $\ell = 1$ and $\ell = 2$ variables is

$$\hat{P}_2(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2; r) = \left(\frac{3}{2\pi}\right)^3(1 - \Upsilon_1^2)^{-3/2} \times \exp\left[-\frac{3\boldsymbol{\eta}_1^2 + 3\boldsymbol{\eta}_2^2 - 6\Upsilon_1\boldsymbol{\eta}_1\cdot\boldsymbol{\eta}_2}{2(1 - \Upsilon_1^2)}\right] \tag{A26}$$

$$\hat{P}_2(\tilde{\zeta}_1, \tilde{\zeta}_2; r) = \frac{1}{20}\frac{15^6}{(2\pi)^5}(1 - \Upsilon_2^2)^{-5/2}\exp\left[-\left(\frac{15}{4}\right) \times \frac{\mathrm{tr}(\tilde{\zeta}_1^2) + \mathrm{tr}(\tilde{\zeta}_2^2) - 2\Upsilon_2\,\mathrm{tr}(\tilde{\zeta}_1\tilde{\zeta}_2)}{(1 - \Upsilon_2^2)}\right], \tag{A27}$$

where we have defined $\Upsilon_n(r) = \xi_0^{(n)}(r)/\sigma_n^2$ for the sake of conciseness. The following relations

$$\eta_1^{(0)}\eta_2^{(0)} + \eta_1^{(+1)}\bar{\eta}_2^{(+1)} + \eta_1^{(-1)}\bar{\eta}_2^{(-1)} = \boldsymbol{\eta}_1\cdot\boldsymbol{\eta}_2 \tag{A28}$$

$$\zeta_1^{(0)}\zeta_2^{(0)} + \sum_{s=1,2}(\zeta_1^{(+s)}\bar{\zeta}_2^{(+s)} + \zeta_1^{(-s)}\bar{\zeta}_2^{(-s)}) = \frac{3}{2}\,\mathrm{tr}(\tilde{\zeta}_1\tilde{\zeta}_2) \tag{A29}$$

can be useful to derive Eqs. (A26) and (A27). When the peak constraint is enforced, $\hat{P}_2(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2; r)$ reduces to a simple multiplicative factor. Finally, the joint probability density $\hat{P}_2(\nu_1, u_1, \nu_2, u_2; r)$ for the $\ell = 0$ degrees of freedom evaluates to

$$\hat{P}_2(\nu_1, u_1, \nu_2, u_2; r) = \frac{e^{-\Phi(\nu_1,u_1,\nu_2,u_2;r)}}{(2\pi)^2\sqrt{(1 - \gamma_1^2)\Delta_{\mathrm{P}}}}. \tag{A30}$$

Here, $\Phi(\nu_1, u_1, \nu_2, u_2; r)$ is the quadratic form associated to the inverse covariance matrix $\mathrm{C}_{\nu u}^{-1}(r) \equiv (\mathrm{P}, \mathrm{R}^\top; \mathrm{R}, \mathrm{P})$. The 2×2 matrix P in the top left and bottom right corners reads as

$$\mathrm{P} = \frac{1}{\Delta_{\mathrm{P}}}\begin{pmatrix} \mathrm{P}_{11} & \mathrm{P}_{12} \\ \mathrm{P}_{12} & \mathrm{P}_{22} \end{pmatrix} = \frac{1}{\Delta_{\mathrm{P}}}\begin{pmatrix} 1 - \frac{\Upsilon_2^2 - 2\gamma_1^2\Upsilon_1\Sigma_2 + \gamma_1^2\Upsilon_1^2}{1-\gamma_1^2} & -\gamma_1 + \frac{\gamma_1\Upsilon_1(\Upsilon_0 - \gamma_1^2\Upsilon_1) - \gamma_1\Upsilon_2(\Upsilon_0 - \Upsilon_1)}{1-\gamma_1^2} \\ -\gamma_1 + \frac{\gamma_1\Upsilon_1(\Upsilon_0 - \gamma_1^2\Upsilon_1) - \gamma_1\Upsilon_2(\Upsilon_0 - \Upsilon_1)}{1-\gamma_1^2} & 1 - \frac{\Upsilon_0^2 - 2\gamma_1^2\Upsilon_0\Upsilon_1 + \gamma_1^2\Upsilon_1^2}{1-\gamma_1^2} \end{pmatrix}, \tag{A31}$$

whereas the matrix R in the top right and bottom left corners is

$$\mathrm{R} = -\frac{1}{(1-\gamma_1^2)\Delta_{\mathrm{P}}}\begin{pmatrix} (\Upsilon_0 - \gamma_1^2\Upsilon_1)\mathrm{P}_{11} - \gamma_1(\Upsilon_2 - \Upsilon_1)\mathrm{P}_{12} & (\Upsilon_0 - \gamma_1^2\Upsilon_1)\mathrm{P}_{12} - \gamma_1(\Upsilon_2 - \Upsilon_1)\mathrm{P}_{22} \\ (\Upsilon_0 - \gamma_1^2\Upsilon_1)\mathrm{P}_{12} - \gamma_1(\Upsilon_2 - \Upsilon_1)\mathrm{P}_{22} & (\Upsilon_2 - \gamma_1^2\Upsilon_1)\mathrm{P}_{22} - \gamma_1(\Upsilon_0 - \Upsilon_1)\mathrm{P}_{12} \end{pmatrix}. \tag{A32}$$

Notice that the determinant $\Delta_{\mathrm{P}} = \mathrm{P}_{11}\mathrm{P}_{22} - \mathrm{P}_{12}^2$ asymptotes to $1 - \gamma_1^2$ in the limit $r \to \infty$.

2. Series expansion of the joint probability density

To calculate the correlation function of initial density peaks at the second order, we first separate the covariance matrix into $\mathrm{C}(r) \equiv \hat{\mathrm{C}}(r) + \delta\mathrm{C}(r)$. The angular average $\hat{\mathrm{C}}(r) \equiv (\mathrm{M}, \hat{\mathrm{B}}^{\top}; \hat{\mathrm{B}}, \mathrm{M})$ contains the zero-point moments and the cross correlation entries with $\xi_0^{(n)}$ solely, whereas $\delta\mathrm{C}(r) \equiv (0_{10}, \delta\mathrm{B}^{\top}; \delta\mathrm{B}, 0_{10})$, where 0_{10} is the 10×10 zero matrix, encodes the cross correlations $\xi_{j\neq 0}^{(n)}$. We have for instance $\hat{\mathrm{B}}^{(2)} = \xi_0^{(2)}/(5\sigma_2^2)$ and $\delta\mathrm{B}^{(2)} = (10\xi_2^{(2)} + 3\xi_4^{(2)})/(35\sigma_2^2)$. Using the identity $\det(\mathrm{I} + \mathrm{X}) = 1 + \mathrm{tr}\mathrm{X} + (1/2)[(\mathrm{tr}\mathrm{X})^2 - \mathrm{tr}(\mathrm{X}^2)] + \cdots$, we expand the joint density $P_2(\mathbf{y}_1, \mathbf{y}_2; r)$ in the small perturbation $\delta\mathrm{C}(r)$ and arrive at

$$\begin{aligned} P_2(\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}) \approx \hat{P}_2(\mathbf{y}_1, \mathbf{y}_2; r)[1 &- \tfrac{1}{2}\mathrm{tr}(\hat{\mathrm{C}}^{-1}\delta\mathrm{C}) \\ &+ \tfrac{1}{4}\mathrm{tr}(\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\delta\mathrm{C})][1 \\ &+ \tfrac{1}{2}\mathbf{y}^{\dagger}\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\mathbf{y} + \tfrac{1}{8}(\mathbf{y}^{\dagger}\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\mathbf{y})^2 \\ &- \tfrac{1}{2}\mathbf{y}^{\dagger}\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\mathbf{y}], \end{aligned} \tag{A33}$$

at second order in $\delta\mathrm{C}$. Here, the product $\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}$ is of order $\mathcal{O}(\xi)$ in the correlation functions $\xi_{\ell}^{(n)}$, whereas $\mathrm{tr}(\hat{\mathrm{C}}^{-1}\delta\mathrm{C})$, $\mathrm{tr}(\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\delta\mathrm{C})$, and $\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}\delta\mathrm{C}\hat{\mathrm{C}}^{-1}$ are of order $\mathcal{O}(\xi^2)$. Expressing the matrices $\hat{\mathrm{C}}$ and $\delta\mathrm{C}$ in terms of the auto and cross covariances yields

$$\begin{aligned} P_2(\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}) \approx \hat{P}_2(\mathbf{y}_1, \mathbf{y}_2; r)\{1 &+ \mathbf{y}_2^{\dagger}\mathrm{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1 \\ &+ \tfrac{1}{2}(\mathbf{y}_2^{\dagger}\mathbf{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1)^2 - \tfrac{1}{2}(\mathbf{y}_1^{\dagger}\mathrm{Q}\mathbf{y}_1 \\ &+ \mathbf{y}_2^{\dagger}\mathrm{Q}\mathbf{y}_2) + \tfrac{1}{2}\mathrm{tr}[\mathrm{M}^{-1}\hat{\mathrm{B}}\mathrm{M}^{-1}(\delta\mathrm{B} + \delta\mathrm{B}^{\top})] \\ &+ \tfrac{1}{2}\mathrm{tr}(\mathrm{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}\delta\mathrm{B}^{\top})\} \end{aligned} \tag{A34}$$

at order $\mathcal{O}(\xi^2)$, where

$$\mathrm{Q} \equiv 2(\mathrm{M}^{-1}\hat{\mathrm{B}}\mathrm{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}) + \mathrm{M}^{-1}\delta\mathrm{B}^{\top}\mathrm{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}. \tag{A35}$$

Note that $\delta\mathrm{B}$ is *not* symmetric, so one must distinguish between $\delta\mathrm{B}$ and its transpose.

Let us consider the term $\mathbf{y}_2^{\dagger}\mathrm{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1$ linear in the correlation functions. Owing to the block-diagonal nature of C, it is a sum of contributions from the spin-0, spin-1, and spin-2 degrees of freedom. While the matrix $(\mathrm{M}^{(s)})^{-1}\delta\mathrm{B}^{(s)}(\mathrm{M}^{(s)})^{-1}$ for $s = 1, 2$ generally has nonvanishing elements, it is easy to check that $(\mathrm{M}^{(0)})^{-1}\delta\mathrm{B}^{(0)}(\mathrm{M}^{(0)})^{-1}$ has zero entries for the elements $ij = 22$, 23, 32, and 33. Furthermore, imposing the constraint $\boldsymbol{\eta}_1 = \boldsymbol{\eta}_2 \equiv 0$ implies that $\mathbf{y}_2^{(s)\dagger}(\mathrm{M}^{(s)})^{-1}\delta\mathrm{B}^{(s)} \times (\mathrm{M}^{(s)})^{-1}\mathbf{y}_1^{(s)}$, $s = 0$, 1, 2 contains only terms linear in $\zeta_1^{(s)}$, $\bar{\zeta}_2^{(s)}$ and products of the form $\zeta_1^{(s)}\bar{\zeta}_2^{(s)}$. At this point, we shall remember that the principal axes of the tensors $\zeta_1 \equiv \zeta(\mathbf{x}_1)$ and $\zeta_2 \equiv \zeta(\mathbf{x}_2)$ are not necessarily aligned with those of the coordinate frame. Without loss of generality, we can write $\zeta_1 = -\mathrm{R}\Lambda\mathrm{R}^{\top}$, where R is an orthogonal matrix that contains the angular variables (e.g., Euler angles) and Λ is the diagonal matrix consisting of the three ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ of $-\zeta_1$. The value of $u_i = -\mathrm{tr}\zeta_1$ is invariant under rotations of the principal axes, while $\tilde{\zeta}_1$ transforms in the same manner as the $\ell = 2$ eigenfunctions of the (orbital) angular momentum operator, i.e., the spherical harmonics $Y_{\ell=2}^{m}(\hat{\mathbf{r}})$. Namely, on inspecting a table of spherical harmonics in Cartesian coordinates, we can write

$$\begin{aligned} r^2 Y_2^0(\hat{\mathbf{r}}) &= \frac{1}{4}\sqrt{\frac{5}{\pi}}(3z^2 - r^2), \\ r^2 Y_2^{\pm 1}(\hat{\mathbf{r}}) &= \mp\frac{1}{2}\sqrt{\frac{15}{2\pi}}z(x \pm iy), \\ r^2 Y_2^{\pm 2}(\hat{\mathbf{r}}) &= \frac{1}{4}\sqrt{\frac{15}{2\pi}}(x \pm iy)^2. \end{aligned} \tag{A36}$$

A comparison with Eq. (A19) shows that $\zeta^{(m)} \sim \sqrt{5/4\pi}Y_2^m(\hat{\mathbf{r}})$. Therefore, the variables $\zeta^{(m)}$ must transform in accordance with

$$\zeta^{(m')}(\mathbf{x}) = \sum_m D^2_{mm'}(\varphi, \vartheta, \psi)\zeta^{(m)}(\mathbf{x}) \tag{A37}$$

under rotations of the principal axis frame. Here, $D^2_{mm'}(\varphi, \vartheta, \psi)$ are quadrupole Wigner D functions with the Euler angles $(\varphi, \vartheta, \psi)$ as arguments, whereas $\zeta^{(m)}$ and $\zeta^{(m')}$ are the components of $\tilde{\zeta}$ in the original and final eigenvector frames, respectively. Therefore, averaging over distinct orientations of the principal axes gives $\langle\zeta^{(m)}\rangle = 0$. Noticing that, at the zeroth order, the joint density $\hat{P}_2(\mathbf{y}_1, \mathbf{y}_2; r)$ factorizes into the product $\hat{P}_1(\mathbf{y}_1)\hat{P}_1(\mathbf{y}_2)$ of

one-point probability densities $\hat{P}_1(\mathbf{y}_i)$ which do not depend upon R, we find the correlations $\xi_j^{(n)}$, $j \neq 0$ *do not contribute* to the peak correlation at the first order. This is in agreement with the findings of [72,74].

The second order terms in the right-hand side of Eq. (A34) will also yield products in the variables $\zeta_1^{(m)}$ and $\bar{\zeta}_2^{(m')}$ whose angle average can be reduced using the orthogonality conditions

$$\int_{\mathrm{SO}(3)} d\mathrm{R} D^{\ell_1}_{m_1 m_1'}(\varphi, \vartheta, \psi) D^{\ell_2 \star}_{m_2 m_2'}(\varphi, \vartheta, \psi) = \frac{1}{2\ell_1 + 1} \delta_{\ell_1 \ell_2} \delta_{m_1 m_2} \delta_{m_1' m_2'} \tag{A38}$$

$$\int_{\mathrm{SO}(3)} d\mathrm{R} D^{\ell_1}_{m_1 m_1'}(\varphi, \vartheta, \psi) D^{\ell_2}_{m_2 m_2'}(\varphi, \vartheta, \psi) = (-1)^{m_2 - m_2'} \frac{1}{2\ell_2 + 1} \delta_{\ell_1 \ell_2} \delta_{-m_1 m_2} \delta_{-m_1' m_2'}. \tag{A39}$$

Clearly, cross products of the form $\zeta_1^{(m)} \bar{\zeta}_2^{(m')}$ vanish upon averaging over the principal axis frames because they involve distinct rotation operators. However, when the angle average is taken at a single position $\mathbf{x}_1 = \mathbf{x}_2$, we obtain

$$\begin{aligned} \langle \zeta^{(m_1')} \bar{\zeta}^{(m_2')} \rangle &= \sum_{m_1 m_2} \zeta^{(m_1)} \bar{\zeta}^{(m_2)} \langle D^2_{m_1 m_1'} D^{2\star}_{m_2 m_2'} \rangle \\ &= \frac{1}{5} \sum_{m_1 m_2} \zeta^{(m_1)} \bar{\zeta}^{(m_2)} \delta_{m_1 m_2} \delta_{m_1' m_2'} \\ &= \frac{3}{10} \delta_{m_1' m_2'} \mathrm{tr}(\tilde{\zeta}^2), \end{aligned} \tag{A40}$$

where $\langle \cdots \rangle$ denotes the average over orientations and we have also omitted the arguments of the Wigner D functions for brevity. Similarly, it is easy to show that

$$\langle \zeta^{(m_1')} \zeta^{(m_2')} \rangle = \langle \bar{\zeta}^{(m_1')} \bar{\zeta}^{(m_2')} \rangle = \tfrac{3}{10} (-1)^{m_2'} \delta_{-m_1' m_2'} \mathrm{tr}(\tilde{\zeta}^2). \tag{A41}$$

These relations can be used to integrate out the orientation of the two eigenframes in the series expansion of the joint probability density $P_2(\mathbf{y}_1, \mathbf{y}_2; \mathbf{r})$. For example, let us consider the contribution $(\mathbf{y}_2^\dagger \mathrm{M}^{-1} \hat{\mathrm{B}} \mathrm{M}^{-1} \mathbf{y}_1)(\mathbf{y}_2^\dagger \mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \mathbf{y}_1)$. After some algebra and with the aid of Eq. (22) of [74], we can write

$$\begin{aligned} (\mathbf{y}_2^\dagger \mathrm{M}^{-1} \hat{\mathrm{B}} \mathrm{M}^{-1} \mathbf{y}_1)(\mathbf{y}_2^\dagger \mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \mathbf{y}_1) = &\left[5 \zeta_1^{(0)} \zeta_2^{(0)} \Upsilon_2 + 5 \sum_{s=1,2} (\zeta_1^{(+s)} \bar{\zeta}_2^{(+s)} + \zeta_1^{(-s)} \bar{\zeta}_2^{(-s)} \Upsilon_2) + \text{terms in } \nu_i, u_i \right] \\ &\times \left\{ \left(\frac{5}{1 - \gamma_1^2} \right) \left[\left(\frac{1}{\sigma_0 \sigma_2} \xi_2^{(1)} - \frac{\gamma_1}{\sigma_2^2} \xi_2^{(2)} \right) (\nu_2 \zeta_1^{(0)} + \nu_1 \zeta_2^{(0)}) \right. \right. \\ &\left. + \left(\frac{1}{\sigma_2^2} \xi_2^{(2)} - \frac{\gamma_1}{\sigma_0 \sigma_2} \xi_2^{(1)} \right) (u_2 \zeta_1^{(0)} + u_1 \zeta_2^{(0)}) \right] + \frac{5}{7\sigma_2^2} \left[(-10 \xi_2^{(2)} + 18 \xi_4^{(2)}) \zeta_1^{(0)} \zeta_2^{(0)} \right. \\ &\left. \left. - (5 \xi_2^{(2)} + 12 \xi_4^{(2)}) \sum_{s=\pm 1} \zeta_1^{(s)} \bar{\zeta}_2^{(s)} + (10 \xi_2^{(2)} + 3 \xi_4^{(2)}) \sum_{s=\pm 2} \zeta_1^{(s)} \bar{\zeta}_2^{(s)} \right] \right\}, \end{aligned} \tag{A42}$$

on enforcing the constraint $\boldsymbol{\eta}_1 = \boldsymbol{\eta}_2 \equiv 0$. To average over the orientation of the tensors ζ_1 and ζ_2, we note that products of the form $\zeta_1^{(m_1)} \bar{\zeta}_2^{(m_1)} \zeta_1^{(m_2)} \bar{\zeta}_2^{(m_2)}$ simplify to

$$\langle \zeta_1^{(m_1)} \bar{\zeta}_2^{(m_1)} \zeta_1^{(m_2)} \bar{\zeta}_2^{(m_2)} \rangle = (\tfrac{3}{10})^2 \delta_{-m_1 m_2} \mathrm{tr}(\tilde{\zeta}_1^2) \mathrm{tr}(\tilde{\zeta}_2^2). \tag{A43}$$

After some further manipulation, this leads to the cancellation of the term (A42).

However, the matrix traces in Eq. (A34) do not vanish on integrating over the angular variables,

$$\begin{aligned} &\mathrm{tr}[\mathrm{M}^{-1} \hat{\mathrm{B}} \mathrm{M}^{-1} (\delta\mathrm{B} + \delta\mathrm{B}^\top)] + \mathrm{tr}(\mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \delta\mathrm{B}^\top) \\ &= \frac{6}{\sigma_1^2 \sigma_2^2} [3(\xi_3^{(3/2)})^2 + 2(\xi_1^{(3/2)})^2] + \frac{6}{\sigma_1^4} (\xi_2^{(1)})^2 \\ &\quad + \frac{10}{7\sigma_2^4} [5(\xi_2^{(2)})^2 + 9(\xi_4^{(2)})^2] + \left(\frac{6}{1 - \gamma_1^2} \right) \left[\frac{1}{\sigma_0^2 \sigma_1^2} (\xi_1^{(1/2)})^2 \right. \\ &\quad \left. - 2 \frac{\gamma_1^2}{\sigma_1^4} \xi_1^{(1/2)} \xi_1^{(3/2)} + \frac{1}{\sigma_1^2 \sigma_2^2} (\xi_1^{(3/2)})^2 \right] + \left(\frac{10}{1 - \gamma_1^2} \right) \\ &\quad \times \left[\frac{1}{\sigma_0^2 \sigma_2^2} (\xi_2^{(1)})^2 - 2 \frac{\gamma_1^2}{\sigma_1^2 \sigma_2^2} \xi_2^{(1)} \xi_2^{(2)} + \frac{1}{\sigma_2^4} (\xi_2^{(2)})^2 \right], \end{aligned} \tag{A44}$$

nor does the second order contribution

$$\begin{aligned} &\tfrac{1}{2} \hat{P}_2(\mathbf{y}_1, \mathbf{y}_2; r) [(\mathbf{y}_2^\dagger \mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \mathbf{y}_1)^2 - (\mathbf{y}_1^\dagger \mathrm{Q} \mathbf{y}_1 + \mathbf{y}_2^\dagger \mathrm{Q} \mathbf{y}_2)] \\ &\approx \tfrac{1}{2} \hat{P}_1(\mathbf{y}_1) \hat{P}_1(\mathbf{y}_2) [(\mathbf{y}_2^\dagger \mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \mathbf{y}_1)^2 \\ &\quad - (\mathbf{y}_1^\dagger \mathrm{Q} \mathbf{y}_1 + \mathbf{y}_2^\dagger \mathrm{Q} \mathbf{y}_2)]. \end{aligned} \tag{A45}$$

To evaluate the latter, we set the first derivatives to zero and recast the term $\mathbf{y}_2^\dagger \mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \mathbf{y}_1$, whose explicit expression is enclosed inside the curly bracket in the right-hand side of Eq. (A42), into the following compact form

$$\begin{aligned} \mathbf{y}_2^\dagger \mathrm{M}^{-1} \delta\mathrm{B} \mathrm{M}^{-1} \mathbf{y}_1 = &f_2(r) \zeta_1^{(0)} + f_1(r) \zeta_2^{(0)} \\ &+ g_0(r) \zeta_1^{(0)} \zeta_2^{(0)} + g_1(r) \sum_{s=\pm 1} \zeta_1^{(s)} \bar{\zeta}_2^{(s)} \\ &+ g_2(r) \sum_{s=\pm 2} \zeta_1^{(s)} \bar{\zeta}_2^{(s)}, \end{aligned} \tag{A46}$$

where f_1, f_2, g_0, g_1, and g_2 are functions of r and, possibly, also ν_i and u_i [the exact expressions can be read off from Eq. (A42)]. Upon squaring Eq. (A46), taking the average over the principal axis frames and substituting the relations (A40) and (A41), we arrive at

$$\begin{aligned}\langle(\mathbf{y}_2^\dagger \mathrm{M}^{-1}\delta\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1)^2\rangle &= \frac{3}{10}[f_1^2\,\mathrm{tr}(\tilde{\zeta}_2^2) + f_2^2\,\mathrm{tr}(\tilde{\zeta}_1^2)] + \left(\frac{3}{10}\right)^2 \\ &\times (g_0^2 + 2g_1^2 + 2g_2^2)\,\mathrm{tr}(\tilde{\zeta}_1^2)\,\mathrm{tr}(\tilde{\zeta}_2^2) \\ &= \frac{15}{2\sigma_2^2}[\mathrm{tr}(\tilde{\zeta}_1^2)(b_{\nu 2}\xi_2^{(1)} + b_{\zeta 2}\xi_2^{(2)})^2 \\ &+ \mathrm{tr}(\tilde{\zeta}_2^2)(b_{\nu 1}\xi_2^{(1)} + b_{\zeta 1}\xi_2^{(2)})^2] \\ &+ h(r)\,\mathrm{tr}(\tilde{\zeta}_1^2)\,\mathrm{tr}(\tilde{\zeta}_2^2), \end{aligned} \tag{A47}$$

$$h(r) = \left(\frac{45}{14\sigma_2^4}\right)[5(\xi_2^{(2)})^2 + 9(\xi_4^{(2)})^2]. \tag{A48}$$

The variables $b_{\nu i}$ and $b_{\zeta i}$ are defined as

$$b_{\nu i} = \frac{1}{\sigma_0}\left(\frac{\nu_i - \gamma_1 u_i}{1-\gamma_1^2}\right), \qquad b_{\zeta i} = \frac{1}{\sigma_2}\left(\frac{u_i - \gamma_1 \nu_i}{1-\gamma_1^2}\right). \tag{A49}$$

They characterize the large scale bias of density peaks of significance ν_i and curvature u_i. As we will see shortly, product of these two variables generate bias parameters beyond first order. Likewise, the angular average of the scalar-valued function $(\mathbf{y}_1^\dagger \mathrm{Q}\mathbf{y}_1 + \mathbf{y}_2^\dagger \mathrm{Q}\mathbf{y}_2)$ can eventually be expressed as

$$\begin{aligned}\langle \mathbf{y}_1^\dagger \mathrm{Q}\mathbf{y}_1 + \mathbf{y}_2^\dagger \mathrm{Q}\mathbf{y}_2\rangle &= \frac{5}{\sigma_2^2}(b_{\nu 1}\xi_2^{(1)} + b_{\zeta 1}\xi_2^{(2)})^2 \\ &+ \frac{3}{\sigma_1^2}(b_{\nu 1}\xi_1^{(1/2)} + b_{\zeta 1}\xi_1^{(3/2)})^2 \\ &+ q(r)\,\mathrm{tr}(\tilde{\zeta}_1^2) + 1 \leftrightarrow 2, \end{aligned} \tag{A50}$$

where

$$\begin{aligned} q(r) = \frac{15\gamma_1^2}{2(1-\gamma_1^2)}&\left[\frac{1}{\sigma_0^2}(\xi_2^{(1)})^2 - 2\frac{\gamma_1^2}{\sigma_1^2}\xi_2^{(1)}\xi_2^{(2)} + \frac{1}{\sigma_2^2}(\xi_2^{(2)})^2\right] \\ &+ \frac{15}{7\sigma_2^4}[5(\xi_2^{(2)})^2 + 9(\xi_4^{(2)})^2] + \frac{9}{2\sigma_1^2\sigma_2^2}[3(\xi_3^{(3/2)})^2 \\ &+ 2(\xi_1^{(3/2)})^2] \end{aligned} \tag{A51}$$

is a function of the separation r solely.

3. Second order approximation to the peak correlation function

At this point, we follow [36] and transform the eigenvalues of $\zeta(\mathbf{x}_1)$ and $\zeta(\mathbf{x}_2)$ to the new set of variables $\{u_i, \upsilon_i, w_i, i = 1, 2\}$. Here, υ_i and w_i are shape parameters that characterize the asymmetry of the density profile in the neighborhood of density maxima. After some algebra, the two-point correlation of density peaks at second order in $\xi_{j\neq 0}^{(n)}$ can be written as

$$\begin{aligned} 1 + \xi_{\mathrm{pk}}(r) = &\frac{1}{\bar{n}_{\mathrm{pk}}^2}\frac{5^5 3^4}{(2\pi)^6}\frac{(1-\Upsilon_1^2)^{-3/2}(1-\Upsilon_2^2)^{-5/2}}{R_\star^6\sqrt{(1-\gamma_1^2)\Delta_{\mathrm{P}}}}\int\prod_{i=1,2}\{du_i d\upsilon_i dw_i F(u_i,\upsilon_i,w_i)e^{-(5/2)((3\upsilon_i^2+w_i^2)/(1-\Upsilon_2^2))}\}e^{-\Phi}\mathcal{I}_\beta(\tilde{\Lambda}_1,\tilde{\Lambda}_2) \\ &+ \frac{1}{2\bar{n}_{\mathrm{pk}}^2}\frac{5^5 3^4}{(2\pi)^6}R_\star^{-6}(1-\gamma_1^2)^{-1}\int\prod_{i=1,2}\{du_i d\upsilon_i dw_i F(u_i,\upsilon_i,w_i)e^{-\Phi_{0i}}\}\Big[\frac{2}{9}(3\upsilon_1^2+w_1^2)(3\upsilon_2^2+w_2^2)h(r) \\ &+ \frac{5}{\sigma_2^2}(3\upsilon_1^2+w_1^2-1)(b_{\nu 2}\xi_2^{(1)} + b_{\zeta 2}\xi_2^{(2)})^2 - \frac{3}{\sigma_1^2}(b_{\nu 2}\xi_1^{(1/2)} + b_{\zeta 2}\xi_1^{(3/2)})^2 - \frac{2}{3}(3\upsilon_1^2+w_1^2)q(r) + 1\leftrightarrow 2\Big]. \end{aligned} \tag{A52}$$

In what follows, we will restrict ourselves to the cross correlation $\xi_{\mathrm{pk}}(r) \equiv \xi_{\mathrm{pk}}(\nu_1, \nu_2, R_S, r)$ of peaks of height ν_1 and ν_2. Therefore, we must integrate over the peak curvatures u_i. The peak constraint implies that the integration at fixed $u_i \geq 0$ is restricted to the interior of the triangle bounded by $(\upsilon_i, w_i) = (0, 0)$, $(u_i/4, -u_i/4)$, and $(u_i/2, u_i/2)$. Moreover, Φ_{0i} is the quadratic form that appears in the one-point probability density $\hat{P}_1(\mathbf{y}_i)$,

$$2\Phi_{0i} = \nu_i^2 + \frac{(\gamma_1\nu_i - u_i)^2}{1-\gamma_1^2} + 5(3\upsilon_i^2 + w_i^2), \tag{A53}$$

$F(u_i, \upsilon_i, w_i)$ is the weight function defined as [36]

$$F(u_i,\upsilon_i,w_i) \equiv (u_i - 2w_i)[(u_i + w_i)^2 - 9\upsilon_i^2]\upsilon_i(\upsilon_i^2 - w_i^2), \tag{A54}$$

and $\mathcal{I}_\beta(\tilde{\Lambda}_1, \tilde{\Lambda}_2)$, with $\beta(\Upsilon_2) \equiv (15/2)\Upsilon_2/(1-\Upsilon_2^2)$, is the integral

$$\mathcal{I}_\beta(\tilde{\Lambda}_1,\tilde{\Lambda}_2) = \int_{\mathrm{SO(3)}} d\mathrm{R}\exp[\beta\,\mathrm{tr}(\tilde{\Lambda}_1\mathrm{R}\tilde{\Lambda}_2\mathrm{R}^\top)]. \tag{A55}$$

Here, the integration domain is $0 \leq \varphi \leq 2\pi$, $0 \leq \vartheta \leq \pi$, $0 \leq \psi < 2\pi$ and $d\mathrm{R} \equiv (1/8\pi^2)d\cos\vartheta d\varphi d\psi$ is the normalized Haar measure ($\int d\mathrm{R} = 1$) on the group SO(3). There is no analytic, closed-form solution to the integral $\mathcal{I}_\beta(\tilde{\Lambda}_1, \tilde{\Lambda}_2)$, although it can still be expressed as a hypergeometric series in the argument $\beta\Lambda_1$ and Λ_2 (see, e.g.,

[146], and references therein). We emphasize that Eq. (A52) is better than an approximation based on a second order Taylor expansion of the probability density $P_2(\mathbf{y}_1, \mathbf{y}_2; r)$ since it retains the isotropic part at all orders.

For convenience, we write the peak correlation up to second order as follows:

$$\xi_{\rm pk}(r) = \xi^{(1)}_{\rm pk}(r) + \xi^{(2)}_{\rm pk}(r) \equiv \xi^{(1)}_{\rm pk}(r) + \sum_{i=1}^{3} \xi^{(2i)}_{\rm pk}(r) \quad \text{(A56)}$$

where $\xi^{(1)}_{\rm pk}$ is the first order piece, Eq. (14), and $\xi^{(2i)}_{\rm pk}$ are distinct second order contributions depending on (i.) correlation functions only, (ii.) the peak height ν, and (iii.) linear and 2nd order bias parameters. We will now detail each of these contributions.

(i) The first terms, $\xi^{(21)}_{\rm pk}$, follow from expanding the determinant of the covariant matrix $\hat{\rm C}(r)$ at the second order. We have

$$\sqrt{\frac{1-\gamma_1^2}{\Delta_{\rm P}}} - 1 \approx \frac{\Upsilon_0^2 - 4\gamma_1^2\Upsilon_0\Upsilon_1 + 2\gamma_1^2\Upsilon_1^2 + 2\gamma_1^4\Upsilon_1^2 + 2\gamma_1^2\Upsilon_0\Upsilon_2 - 4\gamma_1^2\Upsilon_1\Upsilon_2 + \Upsilon_2^2}{2(1-\gamma_1^2)^2}. \quad \text{(A57)}$$

Including the contribution from the trace, Eq. (A44), and expanding the multiplicative factor $(1-\Upsilon_1^2)^{-3/2}(1-\Upsilon_2^2)^{-5/2} \approx 1 + (3/2)\Upsilon_1^2 + (5/2)\Upsilon_2^2$ at second order yields

$$\begin{aligned}\xi^{(21)}_{\rm pk}(r) = {} & \frac{\Upsilon_0^2 - 4\gamma_1^2\Upsilon_0\Upsilon_1 + 2\gamma_1^2\Upsilon_1^2 + 2\gamma_1^4\Upsilon_1^2 + 2\gamma_1^2\Upsilon_0\Upsilon_2 - 4\gamma_1^2\Upsilon_1\Upsilon_2 + \Upsilon_2^2}{2(1-\gamma_1^2)^2} + \frac{3}{2}\Upsilon_1^2 + \frac{5}{2}\Upsilon_2^2 + \frac{3}{\sigma_1^4}(\xi_2^{(1)})^2 \\ & + \frac{3}{\sigma_1^2\sigma_2^2}[3(\xi_3^{(3/2)})^2 + 2(\xi_1^{(3/2)})^2] + \left(\frac{3}{1-\gamma_1^2}\right)\left[\frac{1}{\sigma_0^2\sigma_1^2}(\xi_1^{(1/2)})^2 - 2\frac{\gamma_1^2}{\sigma_1^4}\xi_1^{(1/2)}\xi_1^{(3/2)} + \frac{1}{\sigma_1^2\sigma_2^2}(\xi_1^{(3/2)})^2\right] \\ & + \frac{5}{7\sigma_2^4}[5(\xi_2^{(2)})^2 + 9(\xi_4^{(2)})^2] + \left(\frac{5}{1-\gamma_1^2}\right)\left[\frac{1}{\sigma_0^2\sigma_2^2}(\xi_2^{(1)})^2 - 2\frac{\gamma_1^2}{\sigma_1^2\sigma_2^2}\xi_2^{(1)}\xi_2^{(2)} + \frac{1}{\sigma_2^4}(\xi_2^{(2)})^2\right]. \quad \text{(A58)}\end{aligned}$$

Notice that $\xi^{(21)}_{\rm pk}$ does not depend upon the peak height, though it depends on the filtering scale R_S at which the peaks are identified.

(ii) The second contribution, $\xi^{(22)}_{\rm pk}$, contains all the terms for which the ν dependence cannot be expressed as a polynomial in the linear and 2nd order bias parameters (to be defined shortly). For subsequent use, we introduce the auxiliary function

$$\begin{aligned} f(u,\alpha) \equiv {} & \frac{3^2 5^{5/2}}{\sqrt{2\pi}}\left\{\int_0^{u/4} d\upsilon \int_{-\upsilon}^{+\upsilon} dw + \int_{u/4}^{u/2} d\upsilon \right. \\ & \left. \times \int_{3\upsilon - w}^{\upsilon} dw\right\} F(u,\upsilon,w) e^{-(5\alpha/2)(3\upsilon^2+w^2)} \\ = {} & \frac{1}{\alpha^4}\left\{\frac{e^{-5\alpha u^2/2}}{\sqrt{10\pi}}\left(-\frac{16}{5} + \alpha u^2\right) + \frac{e^{-5\alpha u^2/8}}{\sqrt{10\pi}}\right. \\ & \times \left(\frac{16}{5} + \frac{31}{2}\alpha u^2\right) + \frac{\sqrt{\alpha}}{2}(\alpha u^3 - 3u) \\ & \left. \times \left[\mathrm{Erf}\left(\sqrt{\frac{5\alpha}{2}}\frac{u}{2}\right) + \mathrm{Erf}\left(\sqrt{\frac{5\alpha}{2}}u\right)\right]\right\} \quad \text{(A59)} \end{aligned}$$

and its integral over the nth power of the peak curvature u times the u-dependent part of the one-point probability distribution,

$$G_n^{(\alpha)}(\gamma_1, w) = \int_0^\infty dx x^n f(x,\alpha)\frac{e^{-(x-w)^2/2(1-\gamma_1^2)}}{\sqrt{2\pi(1-\gamma_1^2)}}. \quad \text{(A60)}$$

These functions are very similar, albeit more general than those defined in Eqs. (A15) and (A19) of [36]. With the above, moments of the peak curvature can now conveniently be written as $\bar{u}^n(\nu) = G_n^{(1)}(\gamma_1, \gamma_1\nu)/G_0^{(1)}(\gamma_1, \gamma_1\nu)$.

Next, we collect all second order terms that feature the product of binomials $(3\upsilon_1^2 + w_1^2)^{n_1}(3\upsilon_2^2 + w^2)^{n_2}$ with $n_1 + n_2 \leq 2$. For instance, expanding the integrand of $\mathcal{I}_\beta$ about $\beta = 0$ gives

$$\begin{aligned}\mathcal{I}_\beta(\tilde{\Lambda}_1, \tilde{\Lambda}_2) & \approx 1 + \frac{2\beta^2}{45}(3\upsilon_1^2 + w_1^2)(3\upsilon_2^2 + w_2^2) \\ & = 1 + \frac{5\Upsilon_2^2}{2}(3\upsilon_1^2 + w_1^2)(3\upsilon_2^2 + w_2^2). \quad \text{(A61)}\end{aligned}$$

Similarly,

$$\begin{aligned} e^{-(5/2)((3\upsilon_i^2+w_i^2)/(1-\Upsilon_2^2))} \approx {} & e^{-(5/2)(3\upsilon_i^2+w_i^2)} \\ & \times \left[1 - \tfrac{5}{2}\Upsilon_2^2(3\upsilon_i^2 + w_i^2)\right]. \quad \text{(A62)} \end{aligned}$$

On inserting these expansions into the integral over the asymmetry parameters and using Eq. (A59), we find

$$\frac{5^5 3^4}{2\pi}\int\prod_{i=1,2}\Big\{dv_i dw_i F(u_i,v_i,w_i)\Big[1-\frac{5}{2}\Upsilon_2^2(3v_i^2+w_i^2)\Big]e^{-(5/2)(3v_i^2+w_i^2)}\Big\}\Big[1+\frac{5}{2}\Upsilon_2^2(3v_1^2+w_1^2)(3v_2^2+w_2^2)\Big]$$
$$-f(u_1,1)f(u_2,1)$$
$$=\Upsilon_2^2\Big[f(u_1,1)\partial_\alpha f(u_2,1)+\partial_\alpha f(u_1,1)f(u_2,1)+\frac{2}{5}\partial_\alpha f(u_1,1)\partial_\alpha f(u_2,1)\Big], \tag{A63}$$

where $\partial_\alpha f(u_i,1)\equiv\partial_\alpha f(u_i,\alpha)|_{\alpha=1}$. We are subtracting the zeroth order contribution from the left-hand side of (A63), which would otherwise give unity upon an integration over the peak curvature. There are two additional terms proportional to $(3v_i^2+w_i^2)q(r)$ and a term like $(3v_1^2+w_1^2)(3v_2^2+w_2^2)h(r)$. For these terms, integrating out the asymmetry parameters yields

$$\frac{5^5 3^4}{4\pi}\int\prod_{i=1,2}\{dv_i dw_i F(u_i,v_i,w_i)e^{-(5/2)(3v_i^2+w_i^2)}\}\Big\{\Big[-\frac{2}{3}(3v_1^2+w_1^2)-\frac{2}{3}(3v_2^2+w_2^2)\Big]q(r)+\frac{4}{9}(3v_1^2+w_1^2)(3v_2^2+w_2^2)h(r)\Big\}$$
$$=\frac{2}{15}q(r)[f(u_1,1)\partial_\alpha f(u_2,1)+\partial_\alpha f(u_1,1)f(u_1,1)]+2\Big(\frac{2}{15}\Big)^2 h(r)\partial_\alpha f(u_1,1)\partial_\alpha f(u_2,1). \tag{A64}$$

Adding Eqs. (A63) and (A64), we can eventually express the second order contribution $\xi_{\rm pk}^{(22)}(r)$ to the peak correlation function as

$$\xi_{\rm pk}^{(22)}(r)=\frac{(1-\gamma_1^2)^{-1}}{(2\pi)^5\bar{n}_{\rm pk}^2R_1^6}\int_0^\infty du_1\int_0^\infty du_2\Big\{\Upsilon_2^2\Big[f(u_1,1)\partial_\alpha f(u_2,1)+\partial_\alpha f(u_1,1)f(u_2,1)+\frac{2}{5}\partial_\alpha f(u_1,1)\partial_\alpha f(u_2,1)\Big]$$
$$+\frac{2}{15}q(r)[f(u_1,1)\partial_\alpha f(u_2,1)+\partial_\alpha f(u_1,1)f(u_1,1)]+\frac{8}{45}h(r)\partial_\alpha f(u_1,1)\partial_\alpha f(u_2,1)\Big\}$$
$$\times\exp\Big[-\frac{u_1^2-2\gamma_1u_1\nu_1+\nu_1^2}{2(1-\gamma_1^2)}-\frac{u_2^2-2\gamma_1u_2\nu_2+\nu_2^2}{2(1-\gamma_1^2)}\Big]$$
$$=\Big(\Upsilon_2^2+\frac{2}{15}q(r)\Big)(\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_1)+\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2)|_{\alpha=1}+\frac{2}{5}\Big(\Upsilon_2^2+\frac{4}{45}h(r)\Big)$$
$$\times\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_1)|_{\alpha=1}\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2)|_{\alpha=1}. \tag{A65}$$

The last equality follows from the well-known relation Eq. (21) for the differential density of peaks of height ν. The logarithmic derivative of $G_0^{(\alpha)}$ with respect to α must be evaluated numerically. Nevertheless, it is worth noticing that $G_0^{(\alpha)}(\gamma_1,\omega)$ and $\partial_\alpha G_0^{(\alpha)}(\gamma_1,\omega)$ are sharply peaked around their maximum. For large values of ω, the former asymptotes to $G_0^{(\alpha)}\approx\alpha^{-5/2}\omega^3$. Hence, this implies $\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\omega)|_{\alpha=1}\approx-5/2$ in the limit $\omega\gg1$.

(iii) The last contribution, $\xi_{\rm pk}^{(23)}$, is the sum of two parts: a term which arises from the exponential $e^{-\Phi}$ and, thus, involves the angle average correlations $\xi_0^{(n)}$; and a second part which involves the correlations $\xi_{j\neq0}^{(n)}$. Upon expanding $e^{-\Phi}$ at the second order and integrating over the shape parameters w_i, v_i and the peak curvature u_i, we obtain (after much tedious algebra)

$$\xi_{\rm pk}^{(23)}(r)=\frac{1}{2}\{\bar{b}_{\nu\nu1}\bar{b}_{\nu\nu2}(\xi_0^{(0)})^2+2(\bar{b}_{\nu\nu1}\bar{b}_{\nu\zeta2}+\bar{b}_{\nu\zeta1}\bar{b}_{\nu\nu2})\xi_0^{(0)}\xi_0^{(1)}+(\bar{b}_{\nu\nu1}\bar{b}_{\zeta\zeta2}+\bar{b}_{\zeta\zeta1}\bar{b}_{\nu\nu2})(\xi_0^{(1)})^2$$
$$+2\bar{b}_{\nu\zeta1}\bar{b}_{\nu\zeta2}[(\xi_0^{(1)})^2+\xi_0^{(0)}\xi_0^{(2)}]+2(\bar{b}_{\zeta\zeta1}\bar{b}_{\nu\zeta2}+\bar{b}_{\nu\zeta1}\bar{b}_{\zeta\zeta2})\xi_0^{(1)}\xi_0^{(2)}+\bar{b}_{\zeta\zeta1}\bar{b}_{\zeta\zeta2}(\xi_0^{(2)})^2\}$$
$$-\frac{1}{2(1-\gamma_1^2)}\Big\{2(\bar{b}_{\nu\zeta1}+\bar{b}_{\nu\zeta2})\Big[\frac{1}{\sigma_0^2}\xi_0^{(0)}\xi_0^{(1)}-\frac{\gamma_1^2}{\sigma_1^2}(\xi_0^{(0)}\xi_0^{(2)}+(\xi_0^{(1)})^2)+\xi_0^{(1)}\xi_0^{(2)}/\sigma_2^2\Big]+(\bar{b}_{\nu\nu1}+\bar{b}_{\nu\nu2})\Big[\frac{1}{\sigma_0^2}(\xi_0^{(0)})^2$$
$$-2\frac{\gamma_1^2}{\sigma_1^2}\xi_0^{(0)}\xi_0^{(1)}+\frac{1}{\sigma_2^2}(\xi_0^{(1)})^2\Big]+(\bar{b}_{\zeta\zeta1}+\bar{b}_{\zeta\zeta2})\Big[\frac{1}{\sigma_0^2}(\xi_0^{(1)})^2-2\frac{\gamma_1^2}{\sigma_1^2}\xi_0^{(1)}\xi_0^{(2)}+\frac{1}{\sigma_2^2}(\xi_0^{(2)})^2\Big]\Big\}-\Big\{\frac{5}{2\sigma_2^2}[\bar{b}_{\nu\nu1}(\xi_2^{(1)})^2$$
$$+2\bar{b}_{\nu\zeta1}\xi_2^{(1)}\xi_2^{(2)}+\bar{b}_{\zeta\zeta1}(\xi_2^{(2)})^2]^2\Big(1+\frac{2}{5}\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_1)|_{\alpha=1}\Big)+\frac{3}{2\sigma_1^2}[\bar{b}_{\nu\nu1}(\xi_1^{(1/2)})^2$$
$$+2\bar{b}_{\nu\zeta1}\xi_1^{(1/2)}\xi_1^{(3/2)}+\bar{b}_{\zeta\zeta1}(\xi_1^{(3/2)})^2]^2+1\leftrightarrow2\Big\}, \tag{A66}$$

where the second order bias parameters $\bar{b}_{\nu\nu i} = \bar{b}_{\nu\nu}(\nu_i, \gamma_1)$, $\bar{b}_{\nu\zeta i} = \bar{b}_{\nu\zeta}(\nu_i, \gamma_1)$, and $\bar{b}_{\zeta\zeta i} = \bar{b}_{\zeta\zeta}(\nu_i, \gamma_1)$ are constructed from products of the variables b_ν and b_ζ defined in Eq. (A49),

$$\bar{b}_{\nu\nu}(\nu_i, \gamma_1) \equiv \bar{b}^2_{\nu i} = \frac{\nu_i^2 - 2\gamma_1\nu_i\bar{u}(\nu_i) + \gamma_1^2\bar{u}^2(\nu_i)}{\sigma_0^2(1-\gamma_1^2)^2} \tag{A67}$$

$$\bar{b}_{\nu\zeta}(\nu_i, \gamma_1) \equiv \overline{b_{\nu i} b_{\zeta i}} = \frac{(1+\gamma_1^2)\nu_i\bar{u}(\nu_i) - \gamma_1[\nu_i^2 + \bar{u}^2(\nu_i)]}{\sigma_0\sigma_2(1-\gamma_1^2)^2} \tag{A68}$$

$$\bar{b}_{\zeta\zeta}(\nu_i, \gamma_1) \equiv \bar{b}^2_{\zeta i} = \frac{\bar{u}^2(\nu_i) - 2\gamma_1\nu_i\bar{u}(\nu_i) + \gamma_1^2\nu_i^2}{\sigma_2^2(1-\gamma_1^2)^2}. \tag{A69}$$

Here, the overline designates the average over the peak curvature. Note that the last line in the right-hand side of Eq. (A66) is the contribution from the correlations $\xi^{(n)}_{j\neq 0}$.

4. A compact expression

The second order contribution $\xi^{(2)}_{\rm pk} = \xi^{(21)}_{\rm pk} + \xi^{(22)}_{\rm pk} + \xi^{(23)}_{\rm pk}$ may be written down in compact form with the aid of the second order peak bias operator $\tilde{\mathfrak{b}}_{\rm IIi}$ defined through the Fourier space relation

$$\tilde{\mathfrak{b}}_{\rm IIi}(q_1, q_2) \equiv \overline{b_{\rm spki}(q_1) b_{\rm spki}(q_2)} - (1-\gamma_1^2)^{-1}\left[\frac{1}{\sigma_0^2} + \frac{(q_1 q_2)^2}{\sigma_2^2} - \frac{\gamma_1^2}{\sigma_1^2}(q_1^2 + q_2^2)\right], \tag{A70}$$

where $b_{\rm spki}(q) \equiv b_{\nu i} + b_{\zeta i} q^2$ and q_1 and q_2 are wave modes. By definition, its action on the functions $\xi^{(n_1)}_{\ell_1}(r)$ and $\xi^{(n_2)}_{\ell_2}(r)$ is

$$(\xi^{(n_1)}_{\ell_1} \tilde{\mathfrak{b}}_{\rm IIi} \xi^{(n_2)}_{\ell_2})(r) \equiv \frac{1}{4\pi^4}\int_0^\infty dq_1 \int_0^\infty dq_2 q_1^{2(n_1+1)} q_2^{2(n_2+1)} \times \tilde{\mathfrak{b}}_{\rm IIi}(q_1, q_2) P_{\delta_s}(q_1) P_{\delta_s}(q_2) \times j_{\ell_1}(q_1 r) j_{\ell_2}(q_2 r). \tag{A71}$$

The second order terms can be rearranged so as to recast the two-point correlation of peaks of height ν_1 and ν_2 into the more compact form

$$\begin{aligned}\xi_{\rm pk}(\nu_1, \nu_2, R_S, r) = {} & (\tilde{\mathfrak{b}}_{\rm I1}\tilde{\mathfrak{b}}_{\rm I2}\xi^{(0)}_0) + \frac{1}{2}(\xi^{(0)}_0 \tilde{\mathfrak{b}}_{\rm II1}\tilde{\mathfrak{b}}_{\rm II2}\xi^{(0)}_0) - \frac{3}{2\sigma_1^2}(\xi^{(1/2)}_1 \tilde{\mathfrak{b}}_{\rm II1}\xi^{(1/2)}_1) - \frac{3}{2\sigma_1^2}(\xi^{(1/2)}_1 \tilde{\mathfrak{b}}_{\rm II2}\xi^{(1/2)}_1) \\ & - \frac{5}{2\sigma_2^2}(\xi^{(1)}_2 \tilde{\mathfrak{b}}_{\rm II1}\xi^{(1)}_2)\left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu_2)|_{\alpha=1}\right) - \frac{5}{2\sigma_2^2}(\xi^{(1)}_2 \tilde{\mathfrak{b}}_{\rm II2}\xi^{(1)}_2)\left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu_1)|_{\alpha=1}\right) \\ & + \frac{5}{2\sigma_2^4}\left[(\xi^{(2)}_0)^2 + \frac{10}{7}(\xi^{(2)}_2)^2 + \frac{18}{7}(\xi^{(2)}_4)^2\right]\left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu_1)|_{\alpha=1}\right)\left(1 + \frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu_2)|_{\alpha=1}\right) \\ & + \frac{3}{\sigma_1^2\sigma_2^2}[3(\xi^{(3/2)}_3)^2 + 2(\xi^{(3/2)}_1)^2] + \frac{3}{2\sigma_1^4}[(\xi^{(1)}_0)^2 + 2(\xi^{(1)}_2)^2].\end{aligned} \tag{A72}$$

Notice that the first term of $\xi^{(21)}_{\rm pk}$ combines with the first curly bracket of $\xi^{(23)}_{\rm pk}$ to give the second term in the right-hand side of Eq. (A72), whereas the last term in Eq. (A72) is the sum of $(3/2)\Upsilon_1^2$ and $(3/\sigma_1^4)(\xi^{(1)}_2)^2$ in $\xi^{(21)}_{\rm pk}$. We recover Eq. (22) in the particular case $\nu_1 = \nu_2 = \nu$. For sake of illustration, the function $1 + (2/5)\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)$ is shown in Fig. 9 for several values of γ_1. Note that it decreases monotonically and vanishes in the limit $\nu \to \infty$.

In the local bias model, the N-order bias parameters are related to the Nth-order derivative of the mass function $n(\nu)$ through a peak-background split argument (see Sec. III B for details). Setting $\nu_1 = \nu_2 = \nu$ for simplicity and collecting the second order terms proportional to $(\xi^{(0)}_0)^2$ that are present in $\xi^{(21)}_{\rm pk}$ and $\xi^{(23)}_{\rm pk}$, we find their sum is

$$\frac{1}{2}\bar{b}^2_{\nu\nu}(\xi^{(0)}_0)^2 - \frac{\bar{b}_{\nu\nu}}{\sigma_0^2(1-\gamma_1^2)}(\xi^{(0)}_0)^2 + \frac{1}{2}\left(\frac{\xi^{(0)}_0}{\sigma_0^2(1-\gamma_1^2)}\right)^2 \equiv \frac{1}{2}b^2_{\rm II}(\xi^{(0)}_0)^2. \tag{A73}$$

Here, $b_{\rm II}$ is the second order peak-background split bias. Even though we do not calculate the peak correlation at the third order, it is straightforward to compute the coefficient multiplying $(\xi^{(0)}_0)^3$. This term arises from $e^{-\Phi}$ at order $\mathcal{O}(\xi^3)$ in the correlation functions, and $e^{-\Phi}$ at $\mathcal{O}(\xi)$ times $\sqrt{(1-\gamma_1^2/\Delta_{\rm P})}$ at $\mathcal{O}(\xi^2)$. Adding these two contributions yields

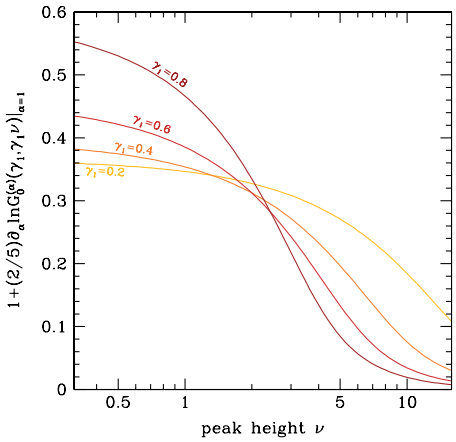

FIG. 9 (color online). The function $1 + (2/5)\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}$ for several values of the correlation strength $\gamma_1 = 0.2$, 0.4, 0.6, and 0.8. This function vanishes in the limit $\nu \to \infty$ since the logarithmic derivative of $G_0^{(\alpha)}$ tends towards $-5/2$.

$$\left[\frac{1}{6}\bar{b}_{\nu\nu\nu} - \frac{\bar{b}_\nu \bar{b}_{\nu\nu\nu}}{\sigma_0^2(1-\gamma_1^2)} + \frac{\bar{b}_\nu^2}{\sigma_0^4(1-\gamma_1^2)^2}\right](\xi_0^{(0)})^3 + \frac{1}{2}\frac{\bar{b}_\nu^2}{\sigma_0^4(1-\gamma_1)^2}(\xi_0^{(0)})^3 = \frac{1}{6}\left[\bar{b}_{\nu\nu\nu} - \frac{3\bar{b}_\nu}{\sigma_0^2(1-\gamma_1^2)}\right]^2(\xi_0^{(0)})^3 \equiv \frac{1}{2}b_{\rm III}^2(\xi_0^{(0)})^3, \tag{A74}$$

where $b_{\rm III}$ is the third order peak-background split bias and the coefficient $\bar{b}_{\nu\nu\nu} \equiv \bar{b}_\nu^3$ is defined in Eq. (29). This demonstrates that the peak-background split derivation of the scale-independent peak bias factors holds at least up to third order. We speculate that this remains true at higher order.

APPENDIX B: NONLINEAR EVOLUTION OF THE PEAK CORRELATION FUNCTION

In this Appendix, we provide technical details regarding the nonlinear evolution of the correlation of initial density peaks under gravitational instabilities. Here again, we will consider two different populations of density peaks of height ν_1 and ν_2 identified on the same filtering scale R_S.

1. The joint velocity distribution

Computing the redshift evolution of the peak correlation function requires knowledge of the conditional probability density $P_2(\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2|\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}, z_i)$ which, in light of the SVT decomposition described in the previous section, is

$$P_2(\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2|\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}', z_i) = \prod_{m=0,\pm1} P_2(\upsilon_1^{(m)}, \bar{\upsilon}_2^{(m)}|\mathbf{y}_1^{(m)}, \bar{\mathbf{y}}_2^{(m)}; \mathbf{r}', z_i). \tag{B1}$$

Here and henceforth, we omit writing the dependence on the initial redshift z_i for brevity. Furthermore, the velocities $\boldsymbol{\upsilon}$ are in units of $aHF\sigma_{-1}$ (so they are dimensionless). Clearly, the line-of-sight velocity $\upsilon^{(0)} = \boldsymbol{\upsilon} \cdot \hat{\mathbf{r}}'$ is a spin-0 component that couples only to the other spin-0 variables $\mathbf{y}_i^{(0)} \equiv \mathbf{y}^{(0)}(\mathbf{x}_i) = (\eta_i^{(0)}, \nu_i, u_i, \zeta_i^{(0)})$, while the components $(\upsilon^{(\pm1)}) = \boldsymbol{\upsilon} \cdot \mathbf{e}_\pm$ of the transverse vector correlate only with $\mathbf{y}_1^{(\pm1)} \equiv \mathbf{y}^{(\pm1)}(\mathbf{x}_1) = (\eta_1^{(\pm1)}, \zeta_1^{(\pm1)})$ and the corresponding conjugates at $\mathbf{x}_2$. Using Schur's identities, we can write the conditional covariance matrix $\Sigma^{(m)}$ and mean value $\Xi^{(m)}$ for the velocity components as

$$\Sigma^{(m)} = \mathrm{W}^{(m)} - \mathrm{V}^{(m)\mathsf{T}}(\mathrm{C}^{(m)})^{-1}\mathrm{V}^{(m)} \tag{B2}$$

$$\Xi^{(m)} = \mathrm{V}^{(m)\mathsf{T}}(\mathrm{C}^{(m)})^{-1}\mathbf{y}^{(m)}. \tag{B3}$$

Here, $W^{(m)}$ is the covariance of the spin-m velocity components $\upsilon^{(m)}$, $V^{(m)}$ is the cross covariance between $\upsilon^{(m)}$ and the vector $\mathbf{y}^{(m)} = (\mathbf{y}_1^{(m)}, \bar{\mathbf{y}}_2^{(m)})$, and $\mathrm{C}^{(m)} = \langle \mathbf{y}^{(m)}\mathbf{y}^{(m)\dagger}\rangle$ is the covariance matrix of $\mathbf{y}^{(m)}$. Like the 20-dimensional matrix $\mathrm{C}(r)$, $\mathrm{C}^{(m)}$ can also be partitioned into four block matrices, with the autocovariance $\mathrm{M}^{(m)}$ along the diagonal and the cross covariance $\mathrm{B}^{(m)}$ and its transpose in the bottom left and top right corners, respectively. The matrices $\mathrm{M}^{(m)}$ and $\mathrm{B}^{(m)}$ are defined in Eqs. (A20) and (A21). To calculate the mode-coupling power, it is quite convenient to work with the Fourier transform of $\Sigma^{(m)}$ and $\Xi^{(m)}$ rather than their real space counterparts (this allows us to circumvent the addition of three angular momenta). To this purpose, we will Fourier transform the entries of $\mathrm{W}^{(m)}$, $\mathrm{B}^{(m)}$, and $\mathrm{C}^{(m)}$ so that $\mathrm{X}^{(m)}(\mathbf{r}) = (1/8\pi^3)\int d^3\mathbf{q}\,\mathrm{X}^{(m)}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}}$. For the spin-0 variables, we obtain

$$\mathrm{B}^{(0)}(\mathbf{q}) = \begin{pmatrix} \frac{q^2}{\sigma_1^2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 & \frac{iq}{\sigma_0\sigma_1}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}) & \frac{iq^3}{\sigma_1\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}) & -\frac{iq^3}{2\sigma_1\sigma_2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^3 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})] \\ -\frac{iq}{\sigma_0\sigma_1}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}) & 1/\sigma_0^2 & q^2/\sigma_0\sigma_2 & -\frac{q^2}{2\sigma_0\sigma_2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 - 1] \\ -\frac{iq^3}{\sigma_1\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}) & q^2/\sigma_0\sigma_2 & q^4/\sigma_2^2 & -\frac{q^4}{2\sigma_2^2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 - 1] \\ \frac{iq^3}{2\sigma_1\sigma_2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^3 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})] & -\frac{q^2}{2\sigma_0\sigma_2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 - 1] & -\frac{q^4}{2\sigma_2^2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 - 1] & \frac{q^4}{4\sigma_2^2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 - 1]^2 \end{pmatrix} \tag{B4}$$

and

$$\mathrm{V}^{(0)\mathsf{T}}(\mathbf{q}) = \begin{pmatrix} \gamma_0/3 & 0_{1\times3} & \frac{1}{\sigma_{-1}\sigma_1}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 & -\mathcal{V}^\mathsf{T} \\ \frac{1}{\sigma_{-1}\sigma_1}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 & \mathcal{V}^\mathsf{T} & \gamma_0/3 & 0_{1\times3} \end{pmatrix}, \qquad \mathrm{W}^{(0)}(\mathbf{q}) = \begin{pmatrix} 1/3 & \frac{q^{-2}}{\sigma_{-1}^2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 \\ \frac{q^{-2}}{\sigma_{-1}^2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2 & 1/3 \end{pmatrix}, \tag{B5}$$

where $0_{1\times3} \equiv (0, 0, 0)$ and

$$\mathcal{V}^{\top} = \frac{1}{\sigma_{-1}}\left(\frac{iq^{-1}}{\sigma_0}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}), \frac{iq}{\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}), -\frac{iq}{2\sigma_2}[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^3 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})]\right). \tag{B6}$$

For the spin-1 variables, the cross covariances are

$$\mathrm{B}^{(\pm 1)}(\mathbf{q}) = \begin{pmatrix} \frac{q^2}{\sigma_1^2}(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) & -\frac{\sqrt{3}iq^3}{\sigma_1\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) \\ \frac{\sqrt{3}iq^3}{\sigma_1\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) & \frac{3q^4}{\sigma_2^2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) \end{pmatrix} \tag{B7}$$

whereas

$$\mathrm{V}^{(\pm 1)\top}(\mathbf{q}) = \begin{pmatrix} \gamma_0/3 & 0 & \frac{1}{\sigma_{-1}\sigma_1}(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) & \frac{\sqrt{3}iq}{\sigma_{-1}\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) \\ \frac{1}{\sigma_{-1}\sigma_1}(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) & -\frac{\sqrt{3}iq}{\sigma_{-1}\sigma_2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) & \gamma_0/3 & 0 \end{pmatrix} \tag{B8}$$

$$\mathrm{W}^{(\pm 1)}(\mathbf{q}) = \begin{pmatrix} 1/3 & \frac{q^{-2}}{\sigma_{-1}^2}(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) \\ \frac{q^{-2}}{\sigma_{-1}^2}(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}) & 1/3 \end{pmatrix}. \tag{B9}$$

We have omitted a factor of $P_\delta(q)$ in all matrix elements for shorthand purposes. Furthermore, we have not substituted $\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}$ by $\hat{q}^{(0)}$ and $\mathbf{e}_\pm\cdot\hat{\mathbf{q}}$ by $\hat{q}^{(\pm 1)}$ to avoid heavy notation for their complex conjugates. We also note that entries involving a odd power of $\hat{\mathbf{r}}$ change sign under the space reflection $\hat{\mathbf{r}} \to -\hat{\mathbf{r}}$.

With the SVT decomposition introduced above, the integral over the peak velocities can be written

$$\int d^N\mathbf{y} P_2(\boldsymbol{v}_1, \boldsymbol{v}_2|\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}') e^{i\sigma_v \mathbf{k}\cdot\Delta\boldsymbol{v}_{12}} = \exp\left(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J} + i\mathbf{J}^\dagger\Xi\right) \tag{B10}$$

where $\Sigma = \mathrm{diag}(\Sigma^{(0)}, \Sigma^{(+1)}, \Sigma^{(-1)})$ and $\Xi = (\Xi^{(0)}, \Xi^{(+1)}, \Xi^{(-1)})$ are the covariance and mean of the multivariate Gaussian; and $\mathbf{J}$, $\mathbf{J}^\dagger$ are the six-dimensional vector $\sigma_v(k^{(0)}, -k^{(0)}, k^{(+1)}, -k^{(+1)}, k^{(-1)}, -k^{(-1)})$ and its conjugate transpose, respectively. We will now Taylor expand the right-hand side of Eq. (B10) around the zeroth order contribution to $\Sigma^{(m)}$ and $\Xi^{(m)}$.

2. Zeroth order: Diffusion damping due to random velocities

At the zeroth order, the covariance matrices of the spin-0 and spin-1 conditional velocity distributions reduce to

$$\Sigma^{(0)} = \Sigma^{(\pm 1)} \approx \tfrac{1}{3}(1-\gamma_0^2)\mathbf{I}_2, \tag{B11}$$

whereas there is no net transverse velocity Ξ at this order. Inserting this result in Eq. (B10) yields

$$\int d^N\mathbf{y} P_2(\mathbf{v}_1, \mathbf{v}_2|\mathbf{y}_1, \mathbf{y}_2; \mathbf{r}') e^{i\sigma_v \mathbf{k}\cdot\Delta\mathbf{v}_{12}} \approx e^{-(1/3)k^2\sigma_v^2(1-\gamma_0^2)} = e^{-(1/3)k^2\sigma_{\mathrm{vpk}}^2}, \tag{B12}$$

where $\sigma_{\mathrm{vpk}}^2(z) = \sigma_v^2(z)(1-\gamma_0^2)$ is the three-dimensional velocity dispersion of peaks identified on the filtering scale R_S [Eq. (15)]. Hence, at the zeroth order, the evolved correlation of density peaks is simply obtained through a convolution of the initial correlation $\xi_{\mathrm{pk}}(R_S, \nu, r)$ with the diffusion kernel $\exp[-(1/3)k^2\sigma_{\mathrm{vpk}}^2(z)]$.

3. First order: Linear growth due to coherent motions

At the first order, the Fourier transform of the covariance matrices $\Sigma^{(m)}(\mathbf{q})$ receive a contribution $\delta\Sigma^{(m)}(\mathbf{q})$ given by

$$\delta\Sigma^{(0)}(\mathbf{q}) = \frac{1}{\sigma_{-1}^2}\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2}{q^2} b_{\mathrm{vpk}}^2(q) P_{\delta_s}(q)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\delta\Sigma^{(\pm 1)}(\mathbf{q}) = \frac{1}{\sigma_{-1}^2}\frac{(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}})}{q^2} b_{\mathrm{vpk}}^2(q) P_{\delta_s}(q)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{B13}$$

The velocity bias factor $b_{\mathrm{vpk}}(q)$, Eq. (16), is the same for the two peak populations because it depends only on the filtering scale R_S. In addition, there is a nonzero mean velocity with line-of-sight components

$$\delta\Xi_1^{(0)}(\mathbf{q}) = -\frac{i}{\sigma_{-1}}\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})}{q} b_{\mathrm{vpk}}(q) b_{\mathrm{spk2}}(q) P_{\delta_s}(q) + \frac{5i}{2\sigma_{-1}\sigma_2} q[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^3 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})] b_{\mathrm{vpk}}(q)\zeta_2^{(0)} P_{\delta_s}(q)$$
$$\delta\Xi_2^{(0)}(\mathbf{q}) = \frac{i}{\sigma_{-1}}\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})}{q} b_{\mathrm{vpk}}(q) b_{\mathrm{spk1}}(q) P_{\delta_s}(q) - \frac{5i}{2\sigma_{-1}\sigma_2} q[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^3 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})] b_{\mathrm{vpk}}(q)\zeta_1^{(0)} P_{\delta_s}(q), \tag{B14}$$

and transverse components

$$\delta\Xi_1^{(\pm1)}(\mathbf{q}) = \frac{5\sqrt{3}i}{\sigma_{-1}\sigma_2} \times q(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}})b_{\rm vpk}(q)\zeta_2^{(\pm1)}P_{\delta_s}(q)$$
$$\delta\Xi_2^{(\pm1)}(\mathbf{q}) = -\frac{5\sqrt{3}i}{\sigma_{-1}\sigma_2} \times q(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}})b_{\rm vpk}(q)\zeta_1^{(\pm1)}P_{\delta_s}(q). \quad \text{(B15)}$$

Here, $b_{\rm spki}(q) = b_{\nu i} + b_{\zeta i}q^2$ is the linear bias for peaks of height ν_i and curvature u_i. Even if the peaks had all the same height, we would have to distinguish between the linear bias of the two populations because the peak curvatures would not necessarily be the same. We also remark that the mean transverse velocity components $\Xi^{(\pm1)}$ vanish upon averaging over the orientations of the principal axis frames. With the aid of these results, we now expand the right-hand side of Eq. (B10) up to first order and impose the peak constraint to arrive at

$$\frac{1}{\bar{n}_{\rm pk}^2}\int d^6\zeta_1 d^6\zeta_2 n_{\rm pk}(\mathbf{x}_1')n_{\rm pk}(\mathbf{x}_2')P_2(\mathbf{y}_1,\mathbf{y}_2;\mathbf{r}')\exp\left(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J}+i\mathbf{J}^\dagger\Xi\right)$$
$$\approx e^{-(1/3)k^2\sigma_{\rm vpk}^2}\left\{1+\bar{b}_{\nu1}\bar{b}_{\nu2}\xi_0^{(0)}+(\bar{b}_{\nu1}\bar{b}_{\zeta2}+\bar{b}_{\nu2}\bar{b}_{\zeta1})\xi_0^{(1)}+\bar{b}_{\zeta1}\bar{b}_{\zeta2}\xi_0^{(2)}+\prod_{i=1,2}\left[\frac{1}{\bar{n}_{\rm pk}}\int d^6\zeta_i n_{\rm pk}(\mathbf{x}_i')\hat{P}_1(\mathbf{y}_i)\right]\right.$$
$$\left.\times\left(-\frac{1}{2}\mathbf{J}^\dagger\delta\Sigma\mathbf{J}+i\mathbf{J}^\dagger\delta\Xi\right)\right\} = e^{-(1/3)k^2\sigma_{\rm vpk}^2}\left\{1+\int\frac{d^3\mathbf{q}}{(2\pi)^3}\left[\tilde{\mathfrak{b}}_{\rm I1}(q)\tilde{\mathfrak{b}}_{\rm I2}(q)+\left(\frac{\sigma_v}{\sigma_{-1}}\right)\frac{k}{q}(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}})(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})[\tilde{\mathfrak{b}}_{\rm I1}(q)\right.\right.$$
$$\left.\left.+\tilde{\mathfrak{b}}_{\rm I2}(q)]\tilde{\mathfrak{b}}_{\rm vpk}(q)+\left(\frac{\sigma_v}{\sigma_{-1}}\right)^2\frac{k^2}{q^2}\left[(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}})^2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})^2+\sum_{a=\pm}(\mathbf{e}_a'\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_a'\cdot\hat{\mathbf{q}})(\mathbf{e}_a'\cdot\hat{\mathbf{k}})(\bar{\mathbf{e}}_a'\cdot\hat{\mathbf{k}})\right]\tilde{\mathfrak{b}}_{\rm vpk}^2(q)\right]P_{\delta_s}(q)e^{i\mathbf{q}\cdot\mathbf{r}'}\right\} \quad \text{(B16)}$$

upon an integration over the peak asymmetry parameters υ_i, w_i and the peak curvatures u_i (see Appendix A). In the second term of this equation, $\hat{P}_1(\mathbf{y}_i)$ is a one-point probability density and, in the third, $\hat{\mathbf{r}}'\cdot\mathbf{e}_\pm' = \hat{\mathbf{r}}'\cdot\bar{\mathbf{e}}_\pm' = 0$ where $\hat{\mathbf{r}}' = \mathbf{r}'/r'$. It is important to bear in mind that the spectral moments σ_n are evaluated at initial redshift z_i, e.g., $\sigma_{-1} = \sigma_{-1}(z_i)$. Hence, the ratio σ_v/σ_{-1} equals $D(z)/D(z_i)$. To perform the integral over $\mathbf{r}'$, we utilize the following relations

$$\int d^3\mathbf{r}\hat{r}_i\hat{r}_j e^{i(\mathbf{q}-\mathbf{k})\cdot\mathbf{r}} = (2\pi)^3\delta^{(3)}(\mathbf{q}-\mathbf{k})\hat{k}_i\hat{k}_j,$$
$$\int d^3\mathbf{r}e_{\pm i}\bar{e}_{\pm j}e^{i(\mathbf{q}-\mathbf{k})\cdot\mathbf{r}} = (2\pi)^3\delta^{(3)}(\mathbf{q}-\mathbf{k})\mathfrak{e}_{\pm i}\bar{\mathfrak{e}}_{\pm j}, \quad \text{(B17)}$$

where $\mathfrak{e}_+$ and $\mathfrak{e}_-$ denote the unit vectors orthogonal to the wave vector $\hat{\mathbf{k}}$, i.e., $\hat{\mathbf{k}}\cdot\mathfrak{e}_\pm = 0$. As a consequence, all the terms involving $\mathbf{e}_\pm'\cdot\mathbf{k}$ or $\bar{\mathbf{e}}_\pm'\cdot\mathbf{k}$ vanish. This cancellation reflects the fact that, on average, only streaming motions along the separation vector $\mathbf{r}'$ of a peak pair can affect the peak correlation $\xi_{\rm pk}$. On employing the identity $(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}})^2 + \sum_{a=\pm}(\mathbf{e}_a'\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_a'\cdot\hat{\mathbf{q}}) = 1$ [which follows from Eq. (A28)], we thus obtain

$$\frac{1}{\bar{n}_{\rm pk}^2}\int d^3\mathbf{r}'e^{-i\mathbf{k}\cdot\mathbf{r}'}\int d^6\zeta_1 d^6\zeta_2 n_{\rm pk}(\mathbf{x}_1')n_{\rm pk}(\mathbf{x}_2')P_2(\mathbf{y}_1,\mathbf{y}_2;\mathbf{r}')$$
$$\times\exp\left(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J}+i\mathbf{J}^\dagger\Xi\right)$$
$$\approx e^{-(1/3)k^2\sigma_{\rm vpk}^2}\left\{\tilde{\mathfrak{b}}_{\rm I1}(q)\tilde{\mathfrak{b}}_{\rm I2}(q)+\left(\frac{D(z)}{D(z_i)}\right)[\tilde{\mathfrak{b}}_{\rm I1}(q)\right.$$
$$\left.+\tilde{\mathfrak{b}}_{\rm I2}(q)]\tilde{\mathfrak{b}}_{\rm vpk}(q)+\left(\frac{D(z)}{D(z_i)}\right)^2\tilde{\mathfrak{b}}_{\rm vpk}^2(q)\right\}P_{\delta_s}(q). \quad \text{(B18)}$$

After some further manipulation, this yields the desired result Eq. (63) where it is assumed $\nu_1 = \nu_2 = \nu$.

4. Second order: Lagrangian and gravity mode coupling

The calculation of the 2nd order term in the Taylor expansion of Eq. (B10), which is the lowest order contribution to the mode-coupling power, is long and fastidious. For clarity, we can decompose this second order mode-coupling power into three distinct pieces which we will compute successively : (i.) the second order contribution to the initial peak correlation $\xi_{\rm pk}(\nu_1,\nu_2,R_S,r)$ convolved with the diffusion kernel, (ii.) $\xi_{\rm pk}^{(1)}(\nu_1,\nu_2,R_S,r)$ times the first order term in the expansion of $\exp(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J}+i\mathbf{J}^\dagger\Xi)$, and (iii.) the second order term in the expansion of $\exp(-\frac{1}{2}\mathbf{J}^\dagger\Sigma\mathbf{J}+i\mathbf{J}^\dagger\Xi)$.

(i) The first piece reflects the fact that the initial, second order contribution is progressively smeared out by the peak motions. It is trivially

$$\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{-(1/3)k^2\sigma_{\rm vpk}^2}P_{\rm pk}^{(2)}(\nu_1,\nu_2,R_S,k)e^{i\mathbf{k}\cdot\mathbf{r}}, \quad \text{(B19)}$$

where $P_{\rm pk}^{(2)}(\nu_1,\nu_2,R_S,k)$ is the Fourier transform of the second order correlation of initial density peaks, $\xi_{\rm pk}^{(2)}(\nu_1,\nu_2,R_S,r)$.

(ii) Taking the exponential damping factor out of the integral, the second part of the mode coupling can be written as

$$e^{-(1/3)k^2\sigma_{\rm vpk}^2}\prod_{i=1,2}\left[\frac{1}{\bar{n}_{\rm pk}}\int d^6\zeta_i n_{\rm pk}(\mathbf{x}_i')\hat{P}_1(\mathbf{y}_i)\right](\mathbf{y}_1^\dagger \mathrm{M}^{-1}\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_2)\left(-\frac{1}{2}\mathbf{J}^\dagger\delta\Sigma\mathbf{J}+i\mathbf{J}^\dagger\delta\Xi\right), \tag{B20}$$

where the Fourier transform of $\mathbf{y}_2^\dagger \mathrm{M}^{-1}\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1$ is given by

$$\begin{aligned}\int d^3\mathbf{r}(\mathbf{y}_2^\dagger \mathrm{M}^{-1}\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1)e^{-i\mathbf{q}\cdot\mathbf{r}} = \Big\{& b_{\rm spk1}(q)b_{\rm spk2}(q) - \frac{5}{2\sigma_2}q^2[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2-1](b_{\rm spk1}(q)\zeta_2^{(0)}+b_{\rm spk2}(q)\zeta_1^{(0)}) \\ &+\frac{25}{4\sigma_2^2}q^4[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2-1]^2\zeta_1^{(0)}\zeta_2^{(0)}+\frac{75}{\sigma_2^2}q^4(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2\times\sum(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}})\zeta_1^{(\pm1)}\bar{\zeta}_2^{(\pm1)} \\ &+\frac{3}{2\sigma_2^2}q^4\sum(\mathbf{e}_\pm\cdot\hat{\mathbf{q}})^2(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}})^2\zeta_1^{(\pm2)}\bar{\zeta}_2^{(\pm2)}\Big\}P_{\delta_s}(q).\end{aligned} \tag{B21}$$

As we can see from Eq. (B14), there are also terms linear in $\zeta_1^{(0)}$ and $\zeta_2^{(0)}$ in the first order mean velocity $\delta\Xi$. On angle averaging over the variables which define the orientation of the principal axes, product of the form $\zeta_i^{(0)}\zeta_i^{(0)}$ reduce to $(3/10)\,\mathrm{tr}(\tilde{\zeta}_i^2)$ as already shown in Appendix A. Hence, the Fourier transform of $\mathbf{y}_2^\dagger \mathrm{M}^{-1}\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1$ and $i\mathbf{J}^\dagger\delta\Xi$ becomes

$$\begin{aligned}\langle(\mathbf{y}_2^\dagger \mathrm{M}^{-1}\mathrm{B}\mathrm{M}^{-1}\mathbf{y}_1)(i\mathbf{J}^\dagger\delta\Xi)\rangle = \Big\{&\left(\frac{\sigma_v}{\sigma_{-1}}\right)(\hat{\mathbf{r}}\cdot\mathbf{k})\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)}{q_2}b_{\rm spk1}(q_1)b_{\rm spk2}(q_1)[b_{\rm spk1}(q_2)+b_{\rm spk2}(q_2)]b_{\rm vpk}(q_2) \\ &+\frac{15}{8\sigma_2^2}\left(\frac{\sigma_v}{\sigma_{-1}}\right)q_1^2q_2(\hat{\mathbf{r}}\cdot\mathbf{k})[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)^3-(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)][b_{\rm spk1}(q_1)\,\mathrm{tr}(\tilde{\zeta}_2^2)+b_{\rm spk2}(q_1)\,\mathrm{tr}(\tilde{\zeta}_1^2)]b_{\rm vpk}(q_2)\Big\}P_{\delta_s}(q_1)P_{\delta_s}(q_2).\end{aligned} \tag{B22}$$

Upon substituting this result in Eq. (B20) and integrating over the other variables, the second piece (ii.) can eventually be expressed as

$$\begin{aligned}e^{-(1/3)k^2\sigma_{\rm vpk}^2}\int\frac{d^3\mathbf{q}_1}{(2\pi)^3}\int\frac{d^3\mathbf{q}_2}{(2\pi)^3}\Big\{&\left(\frac{\sigma_v}{\sigma_{-1}}\right)^2\frac{k^2}{q_2^2}\Big[(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)^2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})^2+\sum_{a=\pm}(\mathbf{e}_a'\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}_a'\cdot\hat{\mathbf{q}}_2)(\mathbf{e}_a'\cdot\hat{\mathbf{k}})(\bar{\mathbf{e}}_a'\cdot\hat{\mathbf{k}})\Big]\tilde{\mathfrak{b}}_{\rm I1}(q_1)\tilde{\mathfrak{b}}_{\rm I2}(q_1)\tilde{\mathfrak{b}}_{\rm vpk}^2(q_2) \\ &+\left(\frac{\sigma_v}{\sigma_{-1}}\right)(\hat{\mathbf{r}}'\cdot\mathbf{k})\frac{(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)}{q_2}[\overline{b_{\rm spk1}(q_1)b_{\rm spk1}(q_2)}\tilde{\mathfrak{b}}_{I2}(q_1)+\overline{b_{\rm spk2}(q_1)b_{\rm spk2}(q_2)}\tilde{\mathfrak{b}}_{\rm I1}(q_1)]\tilde{\mathfrak{b}}_{\rm vpk}(q_2) \\ &-\frac{1}{2\sigma_2^2}\left(\frac{\sigma_v}{\sigma_{-1}}\right)q_1^2q_2(\hat{\mathbf{r}}'\cdot\mathbf{k})[3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)^3-(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)](\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2)\tilde{\mathfrak{b}}_{\rm I1}(q_1) \\ &+\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2)\tilde{\mathfrak{b}}_{\rm I2}(q_1))|_{\alpha=1}\tilde{\mathfrak{b}}_{\rm vpk}(q_2)\Big\}P_{\delta_s}(q_1)P_{\delta_s}(q_2)e^{i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{r}'},\end{aligned} \tag{B23}$$

where we have replaced $\mathrm{tr}(\tilde{\zeta}_i^2)$ by $(2/3)(3\nu_i^2+w_i^2)$ before proceeding with the integration over the asymmetry parameters (see Appendix A). Equation (A69) can be employed to rewrite the average $\overline{b_{\rm pki}(q_1)b_{\rm pki}(q_2)}$ in terms of the second order peak bias factors $\tilde{\mathfrak{b}}_{20}$, $\tilde{\mathfrak{b}}_{11}$ and $\tilde{\mathfrak{b}}_{02}$.

(iii) The third part of the mode-coupling power is the most difficult to compute. It can be written as

$$e^{-(1/3)k^2\sigma_{\rm vpk}^2}\prod_{i=1,2}\left[\frac{1}{\bar{n}_{\rm pk}}\int d^6\zeta_i n_{\rm pk}(\mathbf{x}_i')\hat{P}_1(\mathbf{y}_i)\right]\left\{\frac{1}{2}\left(-\frac{1}{2}\mathbf{J}^\dagger\delta\Sigma\mathbf{J}+i\mathbf{J}^\dagger\delta\Xi\right)^2-\frac{1}{2}\mathbf{J}^\dagger\delta^2\Sigma\mathbf{J}+i\mathbf{J}^\dagger\delta^2\Xi\right\}. \tag{B24}$$

Let us first concentrate on the integral over the terms quadratic in $\delta\Sigma$ and $\delta\Xi$. We must proceed carefully with the calculation of $\mathbf{J}^\dagger\delta\Xi$ because, once again, there are terms of the form $\zeta_i^{(\pm s)}\zeta_i^{(\mp s)}$ which do not vanish upon averaging over the angular variables. A straightforward computation yields

$$\langle(\mathbf{J}^\dagger\delta\Xi)^2\rangle = \Big\{-\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2(\hat{\mathbf{r}}\cdot\mathbf{k})^2\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)}{q_1}\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)}{q_2}[b_{\rm spk1}(q_1)+b_{\rm spk2}(q_1)][b_{\rm spk1}(q_2)+b_{\rm spk2}(q_2)]b_{\rm vpk}(q_1)b_{\rm vpk}(q_2)$$
$$-\frac{15}{8\sigma_2^2}\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2 q_1q_2(\hat{\mathbf{r}}\cdot\mathbf{k})^2[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)^3-(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)][3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)^3-(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)][{\rm tr}(\tilde{\zeta}_1^2)+{\rm tr}(\tilde{\zeta}_2^2)]b_{\rm vpk}(q_1)b_{\rm vpk}(q_2)$$
$$+\frac{45}{2\sigma_2^2}\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}_+\cdot\mathbf{k})(\bar{\mathbf{e}}_-\cdot\mathbf{k})\Big[\sum_{a=\pm}(\mathbf{e}_a\cdot\hat{\mathbf{q}}_1)(\bar{\mathbf{e}}_a\cdot\hat{\mathbf{q}}_1)(\mathbf{e}_{-a}\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}_{-a}\cdot\hat{\mathbf{q}}_2)\Big][{\rm tr}(\tilde{\zeta}_1^2)$$
$$+{\rm tr}(\tilde{\zeta}_2^2)]b_{\rm vpk}(q_1)b_{\rm vpk}(q_2)\Big\}P_{\delta_s}(q_1)P_{\delta_s}(q_2). \tag{B25}$$

Adding the contribution proportional to $(\mathbf{J}^\dagger\delta\Sigma\mathbf{J})^2$ and $(\mathbf{J}^\dagger\delta\Sigma\mathbf{J})(\mathbf{J}^\dagger\delta\Xi)$ and integrating out the asymmetry parameters, we finally obtain

$$\frac{1}{2}e^{-(1/3)k^2\sigma_{\rm vpk}^2}\int\frac{d^3\mathbf{q}_1}{(2\pi)^3}\int\frac{d^3\mathbf{q}_2}{(2\pi)^3}\Big\{\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^4\frac{k^4}{q_1^2q_2^2}\Big[(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)^2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})^2+\sum_{a=\pm}(\mathbf{e}'_a\cdot\hat{\mathbf{q}}_1)(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{q}}_1)(\mathbf{e}'_a\cdot\hat{\mathbf{k}})(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{k}})\Big]\Big[(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)^2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})^2$$
$$+\sum_{a=\pm}(\mathbf{e}'_a\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{q}}_2)(\mathbf{e}'_a\cdot\hat{\mathbf{k}})(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{k}})\Big]\tilde{\mathfrak{b}}_{\rm vpk}^2(q_1)\tilde{\mathfrak{b}}_{\rm vpk}^2(q_2)+2\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^3\frac{k^3}{q_1^2}(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})\frac{(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)}{q_2}\Big[(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)^2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{k}})^2$$
$$+\sum_{a=\pm}(\mathbf{e}'_a\cdot\hat{\mathbf{q}}_1)(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{q}}_1)(\mathbf{e}'_a\cdot\hat{\mathbf{k}})(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{k}})\Big][\tilde{\mathfrak{b}}_{\rm I1}(q_2)+\tilde{\mathfrak{b}}_{\rm I2}(q_2)]\tilde{\mathfrak{b}}_{\rm vpk}^2(q_1)\tilde{\mathfrak{b}}_{\rm vpk}(q_2)+\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2(\hat{\mathbf{r}}'\cdot\mathbf{k})^2\frac{(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)}{q_1}\frac{(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)}{q_2}$$
$$\times[b_{\rm spk1}(q_1)b_{\rm spk1}(q_2)+\tilde{\mathfrak{b}}_{\rm I1}(q_1)\tilde{\mathfrak{b}}_{\rm I2}(q_2)+\tilde{\mathfrak{b}}_{\rm I1}(q_2)\tilde{\mathfrak{b}}_{\rm I2}(q_1)+\overline{b_{\rm spk2}(q_1)b_{\rm spk2}(q_2)}]\tilde{\mathfrak{b}}_{\rm vpk}(q_1)\tilde{\mathfrak{b}}_{\rm vpk}(q_2)$$
$$-\frac{1}{\sigma_2^2}\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)\Big(q_1q_2(\hat{\mathbf{r}}'\cdot\mathbf{k})^2[3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)^2-1]-12(\bar{\mathbf{e}}'_+\cdot\mathbf{k})(\bar{\mathbf{e}}'_-\cdot\mathbf{k})$$
$$\times\Big[\sum_{a=\pm}(\mathbf{e}'_a\cdot\hat{\mathbf{q}}_1)(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{q}}_1)(\mathbf{e}'_{-a}\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}'_{-a}\cdot\hat{\mathbf{q}}_2)\Big]\Big)(\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_1)$$
$$+\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2))|_{\alpha=1}\tilde{\mathfrak{b}}_{\rm vpk}(q_1)\tilde{\mathfrak{b}}_{\rm vpk}(q_2)\Big\}P_{\delta_s}(q_1)P_{\delta_s}(q_2)e^{i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{r}'}, \tag{B26}$$

where, for the sake of completeness, we have included all the terms involving $(\bar{\mathbf{e}}'_\pm\cdot\hat{\mathbf{k}})$ even though they will vanish when we carry out the integral over $\mathbf{r}'$. Next, we consider the integral over the second order terms $\delta^2\Sigma$ and $\delta^2\Xi$. The Fourier transform of the second order contribution to the covariance matrices $\Sigma^{(m)}$ are

$$\delta^2\Sigma^{(0)}(\mathbf{q}_1,\mathbf{q}_2)=\frac{1}{\sigma_{-1}^2}\Big\{-\frac{3}{\sigma_1^2}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)+\frac{5}{4\sigma_2^2}q_1q_2[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)^2-1]+(1-\gamma_1^2)^{-1}\Big[\frac{1}{\sigma_0^2}(q_1q_2)^{-1}$$
$$+\frac{1}{\sigma_2^2}q_1q_2-\frac{2\gamma_1}{\sigma_0\sigma_2}q_1^{-1}q_2\Big]\Big\}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)b_{\rm vpk}(q_1)b_{\rm vpk}(q_2)P_{\delta_s}(q_1)P_{\delta_s}(q_2)\mathbf{I}_2$$
$$\delta^2\Sigma^{(\pm1)}(\mathbf{q}_1,\mathbf{q}_2)=\frac{1}{\sigma_{-1}^2}\Big[-\frac{3}{\sigma_1^2}+\frac{15}{\sigma_2^2}q_1q_2(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)\Big](\mathbf{e}_\pm\cdot\hat{\mathbf{q}}_1)(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}_1)(\mathbf{e}_\pm\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}_\pm\cdot\hat{\mathbf{q}}_2)b_{\rm vpk}(q_1)b_{\rm vpk}(q_2)P_{\delta_s}(q_1)P_{\delta_s}(q_2)\mathbf{I}_2, \tag{B27}$$

whereas, for the line-of-sight components of the mean velocity, we arrive at

$$\delta^2\Xi_1^{(0)}(\mathbf{q}_1,\mathbf{q}_2)=-\frac{i}{\sigma_{-1}}\Big\{(1-\gamma_1^2)^{-1}\Big(\frac{1}{\sigma_0^2}q_1^{-1}+\frac{1}{\sigma_2^2}q_1q_2^2-\frac{\gamma_1}{\sigma_0\sigma_2}q_1^{-1}q_2^2-\frac{\gamma_1}{\sigma_0\sigma_2}q_1\Big)+\frac{3}{\sigma_1^2}q_2(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)$$
$$-\frac{5}{4\sigma_2^2}q_1q_2^2[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)^2-1]\Big\}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)b_{\rm spk1}(q_2)b_{\rm vpk}(q_1)P_{\delta_s}(q_1)P_{\delta_s}(q_2)$$
$$\delta^2\Xi_2^{(0)}(\mathbf{q}_1,\mathbf{q}_2)=\frac{i}{\sigma_{-1}}\Big\{(1-\gamma_1^2)^{-1}\Big(\frac{1}{\sigma_0^2}q_1^{-1}+\frac{1}{\sigma_2^2}q_1q_2^2-\frac{\gamma_1}{\sigma_0\sigma_2}q_1^{-1}q_2^2-\frac{\gamma_1}{\sigma_0\sigma_2}q_1\Big)+\frac{3}{\sigma_1^2}q_2(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)$$
$$-\frac{5}{4\sigma_2^2}q_1q_2^2[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_2)^2-1]\Big\}(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}}_1)b_{\rm spk2}(q_2)b_{\rm vpk}(q_1)P_{\delta_s}(q_1)P_{\delta_s}(q_2). \tag{B28}$$

We have ignored all terms linear in $\zeta_i^{(0)}$ because these will cancel out when we integrate over the angular variables. Furthermore, the second order contribution to the transverse velocity component, $\delta^2\Xi_i^{(\pm1)}$, can be ignored because it

vanishes upon averaging over the orientations of the principal axis frames. On integrating over the asymmetry parameters and the peak curvature, the third piece (iii.) can eventually be cast into the form

$$e^{-(1/3)k^2\sigma_{\rm vpk}^2}\int\frac{d^3\mathbf{q}_1}{(2\pi)^3}\int\frac{d^3\mathbf{q}_2}{(2\pi)^3}\Big\{\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2\Big[\frac{3}{\sigma_1^2}(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)-\frac{5}{4\sigma_2^2}q_1q_2[3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)^2-1]-(1-\gamma_1^2)^{-1}$$
$$\times(q_1q_2)^{-1}\Big(\frac{1}{\sigma_0^2}+\frac{(q_1q_2)^2}{\sigma_2^2}-2\frac{\gamma_1^2}{\sigma_1^2}q_2^2\Big)\Big](\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)(\hat{\mathbf{r}}'\cdot\mathbf{k})^2\tilde{\mathfrak{b}}_{\rm vpk}(q_1)\tilde{\mathfrak{b}}_{\rm vpk}(q_2)+\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)^2\Big(\frac{3}{\sigma_1^2}-\frac{15}{\sigma_2^2}q_1q_2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)\Big)$$
$$\times\Big(\sum_{a=\pm}(\mathbf{e}'_a\cdot\hat{\mathbf{q}}_1)(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{q}}_1)(\mathbf{e}'_a\cdot\hat{\mathbf{q}}_2)(\bar{\mathbf{e}}'_a\cdot\hat{\mathbf{q}}_2)(\mathbf{e}'_a\cdot\mathbf{k})(\bar{\mathbf{e}}'_a\cdot\mathbf{k})\Big)\tilde{\mathfrak{b}}_{\rm vpk}(q_1)\tilde{\mathfrak{b}}_{\rm vpk}(q_2)+\Big(\frac{\sigma_v}{\sigma_{-1}}\Big)\Big[\frac{3}{\sigma_1^2}q_2(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)$$
$$-\frac{5}{4\sigma_2^2}q_1q_2^2[3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_2)^2-1]-(1-\gamma_1^2)^{-1}q_1^{-1}\Big(\frac{1}{\sigma_0^2}+\frac{(q_1q_2)^2}{\sigma_2^2}-\frac{\gamma_1^2}{\sigma_1^2}(q_1^2+q_2^2)\Big)\Big](\hat{\mathbf{r}}'\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{r}}'\cdot\mathbf{k})^2$$
$$\times[\tilde{\mathfrak{b}}_{\rm I1}(q_2)+\tilde{\mathfrak{b}}_{\rm I2}(q_2)]\tilde{\mathfrak{b}}_{\rm vpk}(q_1)\Big\}P_{\delta_s}(q_1)P_{\delta_s}(q_2)e^{i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{r}'}. \tag{B29}$$

To derive the mode-coupling power $P_{\rm MC}(k)$, we must now add the contributions Eqs. (B23), (B26), and (B29) and perform the integration over $\mathbf{r}'$. At this point, it is convenient to express the results in terms of quantities at the collapse redshift z_0 rather than the initial redshift $z_i \gg 1$. This change of fiducial redshift is readily achieved by making the replacement $z_i \to z_0$.

Again, all the terms involving the multiplicative factors $(\mathbf{e}'_{\pm}\cdot\mathbf{k})$ or $(\bar{\mathbf{e}}'_{\pm}\cdot\mathbf{k})$ cancel out and, as for the correlation of initial density peaks, the mode-coupling power can be drastically simplified upon substituting the expression of the second order peak bias $\tilde{\mathfrak{b}}_{\rm II}(q_1, q_2, z_0)$, Eq. (A69). Adding Eqs. (B23), (B26), and (B29), the mode-coupling power can be written

$$P_{\rm MC}(\nu, R_S, k, z) = e^{-(1/3)k^2\sigma_{\rm vpk}^2(z)}P_{\rm pk}^{(2)}(R_S, \nu, k)+\frac{e^{-(1/3)k^2\sigma_{\rm vpk}^2(z)}}{2(2\pi)^3}\int d^3\mathbf{q}_1\int d^3\mathbf{q}_2\Big\{\Big(\frac{D(z)}{D(z_0)}\Big)\mathcal{F}_1(\mathbf{q}_1)[\tilde{\mathfrak{b}}_{\rm II1}(q_1,q_2)\tilde{\mathfrak{b}}_{\rm I2}(q_2)$$
$$+\tilde{\mathfrak{b}}_{\rm II2}(q_1,q_2)\tilde{\mathfrak{b}}_{\rm I1}(q_2)\Big]+\Big(\frac{D(z)}{D(z_0)}\Big)\mathcal{F}_1(\mathbf{q}_2)[\tilde{\mathfrak{b}}_{\rm II1}(q_1,q_2)\tilde{\mathfrak{b}}_{\rm I2}(q_1)+\tilde{\mathfrak{b}}_{\rm II2}(q_1,q_2)\tilde{\mathfrak{b}}_{\rm I1}(q_1)]$$
$$+\Big(\frac{D(z)}{D(z_0)}\Big)^2[\mathcal{F}_1(\mathbf{q}_2)]^2\tilde{\mathfrak{b}}_{\rm I1}(q_1)\tilde{\mathfrak{b}}_{\rm I2}(q_1)+\Big(\frac{D(z)}{D(z_0)}\Big)^2[\mathcal{F}_1(\mathbf{q}_1)]^2\tilde{\mathfrak{b}}_{\rm I1}(q_2)\tilde{\mathfrak{b}}_{\rm I2}(q_2)$$
$$+\Big(\frac{D(z)}{D(z_0)}\Big)^2\mathcal{F}_1(\mathbf{q}_1)\mathcal{F}_1(\mathbf{q}_2)[\tilde{\mathfrak{b}}_{\rm II1}(q_1,q_2)+\tilde{\mathfrak{b}}_{\rm II2}(q_1,q_2)]+\Big(\frac{D(z)}{D(z_0)}\Big)^2\mathcal{F}_1(\mathbf{q}_1)\mathcal{F}_1(\mathbf{q}_2)[\tilde{\mathfrak{b}}_{\rm I1}(q_1)\tilde{\mathfrak{b}}_{\rm I2}(q_2)$$
$$+\tilde{\mathfrak{b}}_{\rm I1}(q_2)\tilde{\mathfrak{b}}_{\rm I2}(q_1)]+\Big(\frac{D(z)}{D(z_0)}\Big)^3[\mathcal{F}_1(\mathbf{q}_1)]^2\mathcal{F}_1(\mathbf{q}_2)[\tilde{\mathfrak{b}}_{\rm I1}(q_2)+\tilde{\mathfrak{b}}_{\rm I2}(q_2)\Big]+\Big(\frac{D(z)}{D(z_0)}\Big)^3\mathcal{F}_1(\mathbf{q}_1)[\mathcal{F}_1(\mathbf{q}_2)]^2[\tilde{\mathfrak{b}}_{\rm I1}(q_1)$$
$$+\tilde{\mathfrak{b}}_{\rm I2}(q_1)]+\Big(\frac{D(z)}{D(z_0)}\Big)^4[\mathcal{F}_1(\mathbf{q}_1)]^2[\mathcal{F}_1(\mathbf{q}_2)]^2-\frac{1}{2\sigma_2^2}\Big(\frac{D(z)}{D(z_0)}\Big)q_1^2q_2^2\mathcal{F}_1(\mathbf{q}_2)[3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2$$
$$-1](\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2)\tilde{\mathfrak{b}}_{I1}(q_1)+\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_1)\tilde{\mathfrak{b}}_{\rm I2}(q_1))|_{\alpha=1}$$
$$-\frac{1}{2\sigma_2^2}\Big(\frac{D(z)}{D(z_0)}\Big)^2\mathcal{F}_1(\mathbf{q}_1)\mathcal{F}_1(\mathbf{q}_2)[3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2-1](\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_1)$$
$$+\partial_\alpha\ln G_0^{(\alpha)}(\gamma_1,\gamma_1\nu_2))|_{\alpha=1}+\Big(\frac{D(z)}{D(z_0)}\Big)^2q_1q_2\mathcal{F}_1(\mathbf{q}_1)\mathcal{F}_1(\mathbf{q}_2)\Big(\frac{3}{\sigma_1^2}(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)$$
$$-\frac{5}{4\sigma_2^2}q_1q_2[3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2-1]\Big)+\Big(\frac{D(z)}{D(z_0)}\Big)q_1q_2\mathcal{F}_1(\mathbf{q}_1)\Big(\frac{3}{\sigma_1^2}(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)$$
$$-\frac{5}{4\sigma_2^2}q_1q_2[3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1)^2-1][3(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_2)^2-1]\Big)[\tilde{\mathfrak{b}}_{\rm I1}(q_2)+\tilde{\mathfrak{b}}_{\rm I2}(q_2)]\Big\}P_{\delta_s}(q_1)P_{\delta_s}(q_2)\delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2), \tag{B30}$$

where, unless otherwise specified, all quantities are evaluated at redshift z_0. In analogy with standard PT, we have defined the kernels $\mathcal{F}_n$ as

$$\mathcal{F}_n(\mathbf{q}_1, \cdots \mathbf{q}_n) \equiv \frac{1}{n!}\frac{(\mathbf{k}\cdot\hat{\mathbf{q}}_1)}{q_1}\cdots\frac{(\mathbf{k}\cdot\hat{\mathbf{q}}_n)}{q_n} \times \tilde{\mathfrak{b}}_{\rm vpk}(q_1)\cdots\tilde{\mathfrak{b}}_{\rm vpk}(q_n). \quad \text{(B31)}$$

We have also introduced the symmetric function $\tilde{\mathfrak{b}}^{\rm E}_{\rm IIi}(\mathbf{q}_1, \mathbf{q}_2, z)$,

$$\tilde{\mathfrak{b}}^{\rm E}_{\rm IIi}(q_1, q_2, z) \equiv \mathcal{F}_2(\mathbf{q}_1, \mathbf{q}_2) + \frac{1}{2}\left(\frac{D(z_0)}{D(z)}\right)[\mathcal{F}_1(\mathbf{q}_1)\tilde{\mathfrak{b}}_{\rm Ii}(q_2, z_0) + \mathcal{F}_1(\mathbf{q}_2)\tilde{\mathfrak{b}}_{\rm Ii}(q_1, z_0)] + \frac{1}{2}\left(\frac{D(z_0)}{D(z)}\right)^2\tilde{\mathfrak{b}}_{\rm IIi}(q_1, q_2, z_0), \quad \text{(B32)}$$

which represents the evolved (Eulerian), second order bias of initial density maxima in the Zel'dovich approximation. The action of $\tilde{\mathfrak{b}}^{\rm E}_{\rm II}(q_1, q_2, z)$ on fields and correlation functions is identical to that of $\tilde{\mathfrak{b}}_{\rm II}(q_1, q_2, z_0)$ (see Sec. III A). With the aid of these auxiliary functions, we can rearrange the Fourier transform of $(1/2)\xi^{(0)}_0\tilde{\mathfrak{b}}_{\rm II1}\tilde{\mathfrak{b}}_{\rm II2}\xi^{(0)}_0$ times the exponential damping with the first nine terms in the curly brackets of Eq. (B30) into the compact expression

$$\frac{2}{(2\pi)^3}\left(\frac{D(z)}{D(z_0)}\right)^4 e^{-(1/3)k^2\sigma^2_{\rm vpk}(z)}\int d^3\mathbf{q}_1\int d^3\mathbf{q}_2\tilde{\mathfrak{b}}^{\rm E}_{\rm II1}(\mathbf{q}_1, \mathbf{q}_2, z) \times \tilde{\mathfrak{b}}^{\rm E}_{\rm II2}(\mathbf{q}_1, \mathbf{q}_2, z)P_{\delta_S}(q_1, z_0)P_{\delta_S}(q_2, z_0)\delta^{(3)}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2). \quad \text{(B33)}$$

Observing that $-iq(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})$ and $-(1/2)q^2[3(\hat{\mathbf{r}}\cdot\hat{\mathbf{q}})^2-1]$ are the Fourier transform of $\xi^{(1/2)}_1$ and $\xi^{(1)}_2$, the four last terms in the curly brackets of Eq. (B30) can be combined in a similar way with some of the terms present in the initial peak correlation $\xi_{\rm pk}(R_S, \nu, r)$ and eventually arrive at Eq. (72). Upon changing to the variables $x \equiv q_1/k$ and $\mu = \hat{\mathbf{k}}\cdot\hat{\mathbf{q}}_1$, the mode-coupling power spectrum can be explicitly written as

$$\begin{aligned}
P_{\rm MC}(\nu, R_S, k, z) = {} & \frac{k^3}{(2\pi)^2}e^{-(1/3)k^2\sigma^2_{\rm vpk}(z)}\Bigg\{2\left(\frac{D(z)}{D(z_0)}\right)^4\int_0^\infty dx x^2P_\delta(kx, z_0)\int_{-1}^{+1}d\mu[\tilde{\mathfrak{b}}^{\rm E}_{\rm II}(k, x, \mu, z)]^2P_\delta(k\sqrt{1+x^2-2x\mu}, z_0) \\
& + \frac{6k^2}{\sigma_1^2(z_0)}\left(\frac{D(z)}{D(z_0)}\right)^2\int_0^\infty dx x^3P_\delta(kx, z_0)\int_{-1}^{+1}d\mu\,\mu(1-x\mu)\tilde{\mathfrak{b}}^{\rm E}_{\rm II}(k, x, \mu, z)P_\delta(k\sqrt{1+x^2-2x\mu}, z_0) \\
& - \frac{5k^4}{2\sigma_2^2(z_0)}\left(\frac{D(z)}{D(z_0)}\right)^3\left(1+\frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}\right)\int_0^\infty dx x^4P_\delta(kx, z_0) \\
& \times\int_{-1}^{+1}d\mu(3\mu^2-1)[2-x^2(1-3\mu^2)-4x\mu]\tilde{\mathfrak{b}}^{\rm E}_{\rm II}(k, x, \mu, z)P_\delta(k\sqrt{1+x^2-2x\mu}, z_0) \\
& + \frac{25k^8}{64\sigma_2^4(z_0)}\left(1+\frac{2}{5}\partial_\alpha \ln G_0^{(\alpha)}(\gamma_1, \gamma_1\nu)|_{\alpha=1}\right)^2\int_0^\infty dx x^6P_\delta(kx, z_0) \\
& \times\int_{-1}^{+1}d\mu[(11-30\mu^2+27\mu^4)(1+x^2-2x\mu)^2 - 6(5-42\mu^2+45\mu^4)(1+x^2-2x\mu)(1-x\mu)^2 \\
& + 9(3-30\mu^2+35\mu^4)(1-x\mu)^4]P_\delta(k\sqrt{1+x^2-2x\mu}, z_0) + \frac{27k^4}{8\sigma_1^4(z_0)}\int_0^\infty dx x^4P_\delta(kx) \\
& \times\int_{-1}^{+1}d\mu[(3\mu^2-1)(1-x\mu)^2 + (1-\mu^2)(1+x^2-2x\mu)]P_\delta(k\sqrt{1+x^2-2x\mu}, z_0) \\
& - \frac{15k^5}{4\sigma_1^2(z_0)\sigma_2^2(z_0)}\int_0^\infty dx x^5P_\delta(kx)\int_{-1}^{+1}d\mu\,\mu(1-x\mu)[15\mu^2(1-x\mu)^2 - 9[\mu^2(1+x^2-2x\mu) \\
& + (1-x\mu)^2] + 7(1+x^2-2x\mu)]P_\delta\left(k\sqrt{1+x^2-2x\mu}, z_0\right)\Bigg\}, \quad \text{(B34)}
\end{aligned}$$

where the second order Eulerian peak bias is

$$\begin{aligned}
\tilde{\mathfrak{b}}^{\rm E}_{\rm II}(k, x, \mu, z) = {} & \frac{1}{2}\Bigg\{\frac{\mu(1-x\mu)}{x(1+x^2-2x\mu)}\left(1-\frac{\sigma_0^2}{\sigma_1^2}k^2x^2\right)\left[1-\frac{\sigma_0^2}{\sigma_1^2}k^2(1+x^2-2x\mu)\right] + \left(\frac{D(z_0)}{D(z)}\right)\frac{\mu}{x} \\
& \times\left(1-\frac{\sigma_0^2}{\sigma_1^2}k^2x^2\right)[\tilde{\mathfrak{b}}_{10}+\tilde{\mathfrak{b}}_{01}k^2(1+x^2-2x\mu)] + \left(\frac{D(z_0)}{D(z)}\right)\frac{(1-x\mu)}{(1+x^2-2x\mu)}\left[1-\frac{\sigma_0^2}{\sigma_1^2}k^2(1+x^2-2x\mu)\right] \\
& \times(\tilde{\mathfrak{b}}_{10}+\tilde{\mathfrak{b}}_{01}k^2x^2) + \left(\frac{D(z_0)}{D(z)}\right)^2[\tilde{\mathfrak{b}}_{20}+\tilde{\mathfrak{b}}_{11}k^2(1+2x^2-2x\mu)+\tilde{\mathfrak{b}}_{02}k^4x^2(1+x^2-2x\mu)]\Bigg\}. \quad \text{(B35)}
\end{aligned}$$

To calculate the mode-coupling in configuration space, we simply Fourier transform $P_{\rm MC}(\nu, R_S, k, z_0)$.

[1] G. Efstathiou, W. J. Sutherland, and S. J. Maddox, Nature (London) **348**, 705 (1990).
[2] W. E. Ballinger, A. F. Heavens, and A. N. Taylor, Mon. Not. R. Astron. Soc. **276**, L59 (1995).
[3] H. Tadros *et al.*, Mon. Not. R. Astron. Soc. **305**, 527 (1999).
[4] W. J. Percival *et al.*, Mon. Not. R. Astron. Soc. **327**, 1297 (2001).
[5] M. Tegmark *et al.*, Astrophys. J. **606**, 702 (2004).
[6] S. Cole *et al.*, Mon. Not. R. Astron. Soc. **362**, 505 (2005).
[7] D. J. Eisenstein *et al.*, Astrophys. J. **633**, 560 (2005).
[8] M. Tegmark *et al.*, Phys. Rev. D **74**, 123507 (2006).
[9] G. Hütsi, Astron. Astrophys. **459**, 375 (2006).
[10] W. J. Percival *et al.*, Mon. Not. R. Astron. Soc. **381**, 1053 (2007).
[11] A. G. Sánchez, M. Crocce, A. Cabré, C. M. Baugh, and E. Gaztañaga, Mon. Not. R. Astron. Soc. **400**, 1643 (2009).
[12] A. Cabré and E. Gaztañaga, Mon. Not. R. Astron. Soc. **393**, 1183 (2009).
[13] V. J. Martínez *et al.*, Astrophys. J. Lett. **696**, L93 (2009).
[14] B. A. Reid *et al.*, Mon. Not. R. Astron. Soc. **404**, 60 (2010).
[15] W. Hu and M. White, Astrophys. J. **471**, 30 (1996).
[16] D. J. Eisenstein, W. Hu, and M. Tegmark, Astrophys. J. Lett. **504**, L57 (1998).
[17] A. Cooray, W. Hu, D. Huterer, and M. Joffre, Astrophys. J. Lett. **557**, L7 (2001).
[18] W. Hu and Z. Haiman, Phys. Rev. D **68**, 063004 (2003).
[19] C. Blake and K. Glazebrook, Astrophys. J. **594**, 665 (2003).
[20] E. V. Linder, Phys. Rev. D **68**, 083504 (2003).
[21] T. Matsubara, Astrophys. J. **615**, 573 (2004).
[22] L. Amendola, C. Quercellini, and E. Giallongo, Mon. Not. R. Astron. Soc. **357**, 429 (2005).
[23] C. Blake and S. Bridle, Mon. Not. R. Astron. Soc. **363**, 1329 (2005).
[24] K. Glazebrook and C. Blake, Astrophys. J. **631**, 1 (2005).
[25] D. Dolney, B. Jain, and M. Takada, Mon. Not. R. Astron. Soc. **366**, 884 (2006).
[26] H. Zhan and L. Knox, Astrophys. J. **644**, 663 (2006).
[27] C. Blake *et al.*, Mon. Not. R. Astron. Soc. **365**, 255 (2006).
[28] N. Padmanabhan and M. White, Phys. Rev. D **77**, 123540 (2008).
[29] E. Gaztañaga, A. Cabré, and L. Hui, Mon. Not. R. Astron. Soc. **399**, 1663 (2009).
[30] M. Shoji, D. Jeong, and E. Komatsu, Astrophys. J. **693**, 1404 (2009).
[31] J. Yoo, D. H. Weinberg, J. L. Tinker, Z. Zheng, and M. S. Warren, Astrophys. J. **698**, 967 (2009).
[32] E. A. Kazin *et al.*, Astrophys. J. **710**, 1444 (2010).
[33] H. J. Mo and S. D. M. White, Mon. Not. R. Astron. Soc. **282**, 347 (1996).
[34] R. K. Sheth and G. Lemson, Mon. Not. R. Astron. Soc. **304**, 767 (1999).
[35] N. Kaiser, Astrophys. J. Lett. **284**, L9 (1984).
[36] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
[37] S. Cole and N. Kaiser, Mon. Not. R. Astron. Soc. **237**, 1127 (1989).
[38] T. Matsubara, Astrophys. J. **525**, 543 (1999).
[39] J. N. Fry and E. Gaztanaga, Astrophys. J. **413**, 447 (1993).
[40] A. S. Szalay, Astrophys. J. **333**, 21 (1988).
[41] P. McDonald and A. Roy, J. Cosmol. Astropart. Phys. 8 (2009) 20.
[42] H. J. Mo, Y. P. Jing, and S. D. M. White, Mon. Not. R. Astron. Soc. **284**, 189 (1997).
[43] R. Scoccimarro, R. K. Sheth, L. Hui, and B. Jain, Astrophys. J. **546**, 20 (2001).
[44] M. Manera, R. K. Sheth, and R. Scoccimarro, Mon. Not. R. Astron. Soc. **402**, 589 (2010).
[45] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).
[46] A. F. Heavens, S. Matarrese, and L. Verde, Mon. Not. R. Astron. Soc. **301**, 797 (1998).
[47] P. Catelan, F. Lucchin, S. Matarrese, and C. Porciani, Mon. Not. R. Astron. Soc. **297**, 692 (1998).
[48] C. Porciani, S. Matarrese, F. Lucchin, and P. Catelan, Mon. Not. R. Astron. Soc. **298**, 1097 (1998).
[49] A. Taruya, Astrophys. J. **537**, 37 (2000).
[50] A. Buchalter, M. Kamionkowski, and A. H. Jaffe, Astrophys. J. **530**, 36 (2000).
[51] P. McDonald, Phys. Rev. D **74**, 103512 (2006).
[52] R. E. Smith, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **75**, 063512 (2007).
[53] E. Sefusatti and E. Komatsu, Phys. Rev. D **76**, 083004 (2007).
[54] T. Nishimichi *et al.*, Publ. Astron. Soc. Jpn. **59**, 93 (2007).
[55] T. Matsubara, Phys. Rev. D **78**, 083519 (2008).
[56] D. Jeong and E. Komatsu, Astrophys. J. **691**, 569 (2009).
[57] S. Saito, M. Takada, and A. Taruya, Phys. Rev. D **80**, 083528 (2009).
[58] P. Valageas, arXiv:1009.1131.
[59] U. Pen, Astrophys. J. **504**, 601 (1998).
[60] A. Dekel and O. Lahav, Astrophys. J. **520**, 24 (1999).
[61] A. Taruya and Y. Suto, Astrophys. J. **542**, 559 (2000).
[62] C. S. Frenk, S. D. M. White, M. Davis, and G. Efstathiou, Astrophys. J. **327**, 507 (1988).
[63] N. Katz, T. Quinn, and J. M. Gelb, Mon. Not. R. Astron. Soc. **265**, 689 (1993).
[64] C. Porciani, A. Dekel, and Y. Hoffman, Mon. Not. R. Astron. Soc. **332**, 339 (2002).
[65] J. A. Peacock and A. F. Heavens, Mon. Not. R. Astron. Soc. **217**, 805 (1985).
[66] Y. Hoffman and J. Shaham, Astrophys. J. **297**, 16 (1985).
[67] S. Otto, H. D. Politzer, and M. B. Wise, Phys. Rev. Lett. **56**, 1878 (1986).
[68] H. M. P. Couchman, Mon. Not. R. Astron. Soc. **225**, 777 (1987).
[69] P. Coles, Mon. Not. R. Astron. Soc. **238**, 319 (1989).
[70] S. L. Lumsden, A. F. Heavens, and J. A. Peacock, Mon. Not. R. Astron. Soc. **238**, 293 (1989).
[71] J. A. Peacock and A. F. Heavens, Mon. Not. R. Astron. Soc. **243**, 133 (1990).
[72] E. Regös and A. S. Szalay, Mon. Not. R. Astron. Soc. **272**, 447 (1995).
[73] A. Manrique *et al.*, Astrophys. J. **499**, 548 (1998).
[74] V. Desjacques, Phys. Rev. D **78**, 103503 (2008).
[75] V. Desjacques and R. K. Sheth, Phys. Rev. D **81**, 023526 (2010).
[76] N. Kaiser and M. Davis, Astrophys. J. **297**, 365 (1985).

[77] J. R. Bond, in *Large-Scale Motions in the Universe: A Vatican Study Week*, edited by V. C. Rubin and G. V.Coyne (Princeton University Press, New Jersey, 1988), pp. 419–435.
[78] S. Colafrancesco, F. Lucchin, and S. Matarrese, Astrophys. J. **345**, 3 (1989).
[79] R. Cen, Astrophys. J. **509**, 494 (1998).
[80] R. K. Sheth, in *The Onset of Nonlinearity in Cosmology*, New York Academy Sciences Annals Vol. 927, edited by J. N. Fry, J. R. Buchler, and H. Kandrup (Wiley, New York, 2001), pp. 1.
[81] R. A. C. Croft and E. Gaztanaga, Astrophys. J. **495**, 554 (1998).
[82] S. De and R. A. C. Croft, Mon. Not. R. Astron. Soc. **401**, 1989 (2010).
[83] W. J. Percival and B. M. Schäfer, Mon. Not. R. Astron. Soc. **385**, L78 (2008).
[84] N. Dalal, M. White, J. R. Bond, and A. Shirokov, Astrophys. J. **687**, 12 (2008).
[85] H. D. Politzer and M. B. Wise, Astrophys. J. Lett. **285**, L1 (1984).
[86] B. A. Bassett and R. Hlozek, arXiv:0910.5224.
[87] T. Giannantonio and C. Porciani, Phys. Rev. D **81**, 063530 (2010).
[88] A. S. Szalay and L. G. Jensen, Acta Phys. Hung. **62**, 263 (1987).
[89] E. Komatsu *et al.*, arXiv:1001.4538.
[90] J. E. Gunn and J. R. I. Gott, Astrophys. J. **176**, 1 (1972).
[91] W. H. Press and P. Schechter, Astrophys. J. **187**, 425 (1974).
[92] A. G. Doroshkevich, Astrofiz. **6**, 581 (1970).
[93] J. R. Bond and S. T. Myers, Astrophys. J. Suppl. Ser. **103**, 1 (1996).
[94] R. K. Sheth, H. J. Mo, and G. Tormen, Mon. Not. R. Astron. Soc. **323**, 1 (2001).
[95] V. Desjacques, Mon. Not. R. Astron. Soc. **388**, 638 (2008).
[96] M. Kac, Bull. Am. Math. Soc. **49**, 938 (1943).
[97] S. O. Rice, Bell Syst. Tech. J. **25**, 46 (1945).
[98] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999).
[99] A. Taruya, H. Magara, Y. P. Jing, and Y. Suto, Publ. Astron. Soc. Jpn. **53**, 155 (2001).
[100] J. L. Tinker *et al.*, arXiv:1001.3162.
[101] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063519 (2006).
[102] Y. B. Zel'dovich, Astron. Astrophys. **5**, 84 (1970).
[103] P. J. E. Peebles, in *Research Supported by the National Science Foundation*, edited by P. J. E. Peebles (Princeton University Press, Princeton, New Jersey, 1980), p. 435.
[104] S. Bharadwaj, Astrophys. J. **472**, 1 (1996).
[105] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063520 (2006).
[106] J. N. Fry, Astrophys. J. **279**, 499 (1984).
[107] M. H. Goroff, B. Grinstein, S. Rey, and M. B. Wise, Astrophys. J. **311**, 6 (1986).
[108] B. Grinstein and M. B. Wise, Astrophys. J. **320**, 448 (1987).
[109] T. Matsubara, Phys. Rev. D **77**, 063530 (2008).
[110] Y. B. Zel'dovich, Adv. Astron. Astrophys. **3**, 241 (1965).
[111] P. J. E. Peebles, Astron. Astrophys. **32**, 391 (1974).
[112] P. Coles, Mon. Not. R. Astron. Soc. **262**, 1065 (1993).
[113] R. J. Scherrer and D. H. Weinberg, Astrophys. J. **504**, 607 (1998).
[114] J. N. Fry, Astrophys. J. Lett. **461**, L65+ (1996).
[115] M. Tegmark and P. J. E. Peebles, Astrophys. J. Lett. **500**, L79 (1998).
[116] L. Hui and K. P. Parfrey, Phys. Rev. D **77**, 043527 (2008).
[117] P. Catelan, C. Porciani, and M. Kamionkowski, Mon. Not. R. Astron. Soc. **318**, L39 (2000).
[118] R. Casas-Miranda, H. J. Mo, R. K. Sheth, and G. Boerner, Mon. Not. R. Astron. Soc. **333**, 730 (2002).
[119] N. Hamaus, U. Seljak, V. Desjacques, R. E. Smith, and T. Baldauf, Phys. Rev. D **82**, 043515 (2010).
[120] J. Beltrán Jiménez and R. Durrer, arXiv:1006.2343.
[121] M. Crocce and R. Scoccimarro, Phys. Rev. D **77**, 023533 (2008).
[122] R. E. Smith, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **77**, 043525 (2008).
[123] N. Padmanabhan and M. White, Phys. Rev. D **80**, 063508 (2009).
[124] R. E. Angulo, C. M. Baugh, C. S. Frenk, and C. G. Lacey, Mon. Not. R. Astron. Soc. **383**, 755 (2007).
[125] D. J. Eisenstein, H. Seo, and M. White, Astrophys. J. **664**, 660 (2007).
[126] P. Fosalba, E. Gaztañaga, F. J. Castander, and M. Manera, Mon. Not. R. Astron. Soc. **391**, 435 (2008).
[127] M. Crocce, P. Fosalba, F. J. Castander, and E. Gaztañaga, Mon. Not. R. Astron. Soc. **403**, 1353 (2010).
[128] http://www.ice.cat/mice.
[129] M. Crocce, P. Folsalba, F. J. Castander, and E. Gaztañaga (unpublished).
[130] S. D. Landy and A. S. Szalay, Astrophys. J. **412**, 64 (1993).
[131] J. Kim, C. Park, J. R. Gott, and J. Dubinski, Astrophys. J. **701**, 1547 (2009).
[132] L. Appel and B. J. T. Jones, Mon. Not. R. Astron. Soc. **245**, 522 (1990).
[133] A. Manrique and E. Salvador-Sole, Astrophys. J. **453**, 6 (1995).
[134] H. Seo *et al.*, Astrophys. J. **720**, 1650 (2010).
[135] E. M. Lifshiftz, J. Phys. (Moscow) **10**, 116 (1946).
[136] P. J. E. Peebles and J. T. Yu, Astrophys. J. **162**, 815 (1970).
[137] J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).
[138] H. Kodama and M. Sasaki, Prog. Theor. Phys. Suppl. **78**, 1 (1984).
[139] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).
[140] R. Dürrer, Fundam. Cosm. Phys. **15**, 209 (1994).
[141] C. Ma and E. Bertschinger, Astrophys. J. **455**, 7 (1995).
[142] W. Hu, arXiv:astroph/0402060.
[143] K. Gorski, Astrophys. J. Lett. **332**, L7 (1988).
[144] R. Dürrer and M. Kunz, Phys. Rev. D **57**, R3199 (1998).
[145] L. D. Landau and E. M. Lifshitz, in *The Classical Theory of Fields*, (Butterworth-Heinemann, Oxford, 1975), 4th ed., Vol. 2.
[146] V. Desjacques and R. E. Smith, Phys. Rev. D **78**, 023527 (2008).