# Non-Gaussian halo bias beyond the squeezed limit

Fabian Schmidt

*Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA*
(Received 29 April 2013; published 17 June 2013)

Primordial non-Gaussianity, in particular the coupling of modes with widely different wavelengths, can have a strong impact on the large-scale clustering of tracers through a scale-dependent bias with respect to matter. We demonstrate that the standard derivation of this non-Gaussian scale-dependent bias is in general valid only in the extreme squeezed limit of the primordial bispectrum, i.e. for clustering over very large scales. We further show how the treatment can be generalized to describe the scale-dependent bias on smaller scales, without making any assumptions on the nature of tracers apart from a dependence on the small-scale fluctuations within a finite region. If the leading scale-dependent bias $\Delta b \propto k^{\alpha}$, then the first subleading term will scale as $k^{\alpha+2}$. This correction typically becomes relevant as one considers clustering over scales $k \gtrsim 10^{-2} h \text{ Mpc}^{-1}$.

## I. INTRODUCTION

Primordial non-Gaussianity is one of the most promising probes of the physics and nature of inflation in the early Universe [1]. Currently the best constraints on non-Gaussianity come from the cosmic microwave background as observed by the Planck satellite [2]. However, it has become clear recently that observations of the clustering of large-scale structure (LSS) tracers will offer competitive constraining power on non-Gaussianity [3–9].

The key ingredient in describing the impact of non-Gaussianity on LSS statistics is the description of the biasing of tracers, that is, the relation to the matter density field. Dalal *et al.* [10] showed that in the presence of non-Gaussianity of the local type, where the non-Gaussian potential perturbation $\phi$ is given in terms of a Gaussian field $\hat{\phi}$ via

$$\phi(\mathbf{x}) = \hat{\phi}(\mathbf{x}) + f_{\text{NL}}(\hat{\phi}^2(\mathbf{x}) - \langle \hat{\phi}^2 \rangle), \qquad (1)$$

where $f_{\text{NL}}$ is a dimensionless parameter, leads to a strongly scale-dependent bias which increases towards large scales as $k^{-2}$ in Fourier space. That is, on large scales, tracers follow the potential, rather than matter, as in the case of Gaussian initial conditions. This effect can be understood as follows. Given a potential described by Eq. (1), one can easily show [10,11] that the power spectrum of small-scale matter fluctuations in a patch of size $R_L$ around $\mathbf{x}$ is rescaled as

$$P(k_s) = [1 + 4f_{\text{NL}}\phi_L(\mathbf{x})]\hat{P}(k_s), \qquad (2)$$

where $\phi_L(\mathbf{x})$ is the potential averaged over the patch, and $\hat{P}(k_s)$ is the matter power spectrum derived from the Gaussian field $\hat{\phi}$. Tracers in this patch then effectively form in a Universe with higher primordial power spectrum amplitude, which will clearly change their abundance relative to other patches resulting in a modulation of the tracer density by long wavelength potential perturbations.

The rescaling Eq. (2) is only valid in the large-scale limit $R_L \to \infty$, that is when gradients of $\phi$ can be neglected. As we will see, this is formally equivalent to only considering the leading contribution to the bispectrum $B_\phi$ in the squeezed limit, $\lim_{k \to 0} B_\phi(k, k_s, |\mathbf{k}_s + \mathbf{k}|)$. This is clearly not a good assumption for most current large-scale structure surveys, which probe Fourier modes $k \gtrsim 10^{-2} h \text{ Mpc}^{-1}$. Several treatments have gone beyond the approximation Eq. (2) [12–14]; however, they all assumed specific models for the tracers (thresholding or excursion set). It is likely that more refined models for tracers will be necessary in order to adequately describe the large samples of galaxies, quasars, and other tracers delivered by ongoing and upcoming surveys. The goal of this paper is to go beyond the limit described by Eq. (2) while keeping the treatment fully independent of detailed assumptions about the tracers.

Our treatment will be based on the approach developed in Schmidt *et al.* [15] (see also [16,17]). The underlying idea is that, when coarse-grained on a sufficiently large scale $R_L$, the abundance of tracers only depends locally on various coarse-grained properties of the density field. The properties considered in [15] were the coarse-grained matter density $\rho_L = \bar{\rho}(1 + \delta_L)$, the curvature of the matter density $\nabla^2 \rho_L$, and the amplitude of small-scale fluctuations $y_*$. The assumption of locality is valid as long as $R_L$ is much larger than the scale of nonlocality of the tracer considered. By defining renormalized bias parameters, it is then possible to absorb the dependence on the arbitrary scale $R_L$ in the expression for tracer correlations, which then only involves large-scale matter correlators. The renormalized bias parameters which, following conventional usage, we refer to as "peak-background split" (PBS) bias parameters, are given in terms of derivatives of the mean abundance of tracers with respect to the properties of the background Universe and initial conditions.

Let us briefly recap the renormalization procedure presented in [15] for tracer clustering in the presence of primordial non-Gaussianity. In this paper we will assume that the leading non-Gaussian contribution is given by the bispectrum; the generalization to higher $N$-point functions is straightforward. If we assume that, on large scales, tracer correlations are completely described by the dependence of the tracer density on the coarse-grained fractional matter density perturbation $\delta_L$,

$$\delta_L(\mathbf{x}) = \int d^3\mathbf{y} W_L(|\mathbf{x} - \mathbf{y}|)\delta(\mathbf{y}), \qquad (3)$$

where $W_L$ is an arbitrary spherically symmetric filter function on the scale $R_L$, then the tracer correlation function $\xi_h$ is to leading order given by

$$\xi_h(r) = b_1^2 \xi_L(r) + b_1 b_2 \langle\delta_L(1)\delta_L^2(2)\rangle + \cdots, \qquad (4)$$

where $\xi_L(r) = \langle\delta_L(1)\delta_L(2)\rangle$, and $b_N$ are the renormalized PBS bias parameters defined as derivatives of the mean tracer abundance with respect to a change in the background density $\bar{\rho}$ of the Universe.

The leading non-Gaussian modification is the second term $\propto b_1 b_2$. In terms of the matter bispectrum $B_m(k_1, k_2, k_3)$, this three-point correlator is given by

$$\langle\delta_L(1)\delta_L^2(2)\rangle = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} W_L(k) \int \frac{d^3k_1}{(2\pi)^3} W_L(k_1)$$
$$\times W_L(|\mathbf{k} + \mathbf{k}_1|) B_m(k, k_1, |\mathbf{k} + \mathbf{k}_1|). \qquad (5)$$

In the approach described in [15], Eq. (4) is only a valid description of tracer correlations as long as it is independent of the value of the coarse-graining scale $R_L$. In order for this to be satisfied, we need to write Eq. (5) in a form that is separable in $r$ and $R_L$. For realistic bispectra $B$, this can in general only be done approximately. Fortunately, in the case of primordial non-Gaussianity where the contribution in Eq. (5) typically becomes significant for large separations $r$, we can make use of a separation of scales: while the Fourier integral over $k$ roughly picks out scales of $k \sim 1/r$, the integral over $k_1$ peaks for $k_1 \sim 1/R_L$. We thus perform an expansion of Eq. (5) in powers of $k/k_1 \sim kR_L$.

Let us restrict to local primordial non-Gaussianity for the time being. As shown in [15], to leading order in powers of $k/k_1$,

$$\langle\delta_L(1)\delta_L^2(2)\rangle = 4f_{\rm NL}\sigma_L^2\xi_{\phi\delta}(r), \qquad (6)$$

where $\xi_{\phi\delta}$ is the potential-matter cross-correlation function. Clearly, this contribution is strongly $R_L$ dependent through the factor $\sigma_L^2$. The solution introduced in [15] is to explicitly account for the dependence of the tracer density on the local amplitude of small-scale matter fluctuations, $y_* = (\delta_s^2/\sigma_s^2 - 1)/2$, by generalizing the renormalized local bias parameters $b_N$ to a bivariate bias expansion $b_{NM}$ [18,19], where $b_{N0}$ are equal to the local biases $b_N$. The lowest-order new bias parameter $b_{01}$, which

corresponds to the response of the mean tracer abundance to a change in the matter power spectrum normalization, then absorbs the term Eq. (6), leading to an $R_L$-independent final expression for the two-point correlation,

$$\xi_h(r) = b_{10}^2 \xi_L(r) + 2b_{10}b_{01} 2f_{\rm NL}\xi_{\phi\delta}(r) + \cdots. \qquad (7)$$

The second term here corresponds to the scale-dependent bias identified in [10], where $b_{01}$ quantifies the amplitude and is given by the derivative of the tracer abundance with respect to a change in the amplitude of primordial fluctuations.

Unfortunately, Eq. (6) is only an accurate approximation to the full expression Eq. (5) on very large scales. Taking the expansion in $k/k_1$ to order $(k/k_1)^2$, we will show below that

$$\langle\delta_L(1)\delta_L^2(2)\rangle = 2f_{\rm NL}[\sigma_L^2\xi_{\phi\delta}(r) + \sigma_{X,L}^2\xi_{\varphi\delta}(r)], \qquad (8)$$

where $\varphi(\mathbf{x}) = -\nabla^2\phi(\mathbf{x})$, and the spectral moment $\sigma_{X,L}$ is defined below [Eq. (12) with Eq. (29)]. The crucial point is that the second term in Eq. (8) scales differently with $r$ and $R_L$ than the first term. Hence, there is no hope that it will be absorbed by the single bias parameter $b_{01}$ introduced in [15]. In fact, the physical interpretation of the two terms in Eq. (8) is quite different: while the first term quantifies the uniform rescaling of the local small-scale fluctuations by $\phi$ [Eq. (2)], the second term corresponds to a scale-dependent rescaling of small-scale fluctuations by the field $\varphi(\mathbf{x})$,

$$P(k_s) \rightarrow [1 + 4f_{\rm NL}\varphi(\mathbf{x})X(k_s)]P(k_s). \qquad (9)$$

The first effect is naturally captured by allowing for a dependence of the tracer density on the amplitude of small-scale fluctuations. On the other hand, to capture the effect of the second term in Eq. (8), we need to allow for a dependence of the tracer density on the shape of the power spectrum of small-scale fluctuations. As we will see, such a dependence along with the associated renormalized bias parameter is exactly what is needed to absorb the $R_L$ dependence introduced by the second term in Eq. (8).

The purpose of the following sections is to make these statements rigorous. Further, we will present all derivations for a general bispectrum of primordial non-Gaussianity, as the treatment is easily phrased to encompass the general case.

The outline of the paper is as follows. In Sec. II, we introduce notation and conventions used throughout the paper. Section III describes the squeezed-limit expansion of three-point correlators for general (separable) bispectra. The scale-dependent bias beyond the squeezed limit is derived in Sec. IV, while examples and numerical results are presented in Sec. V. We make the connection to previous results in Sec. VI and discuss other sources that become relevant in this regime in Sec. VII. We conclude in Sec. VIII.

## II. NOTATION

We assume that the primordial $N$-point functions are given in terms of the Bardeen potential during matter domination, $\phi$. Throughout, we will only deal with the statistics of $\phi$ and the initial (linear) density field, scaled to some redshift $z$. Note in particular that the density bias parameters correspondingly denote Lagrangian biases throughout. The relation between $\phi$ and the linear density field at redshift $z$ is written in Fourier space as

$$\delta(\mathbf{k}, z) = \mathcal{M}(k, z)\phi(\mathbf{k})$$

$$\mathcal{M}(k, z) = \frac{2}{3}\frac{k^2 T(k)g(z)}{\Omega_m H_0^2(1+z)}, \qquad (10)$$

where $T(k)$ is the matter transfer function normalized to unity as $k \to 0$, and $g(z)$ is the linear growth rate of the gravitational potential normalized to unity during the matter-dominated epoch. In the following, we will drop the argument $z$ in $\delta$ and $\mathcal{M}$ since it is not of relevance in the derivation. Further, we define $\mathcal{M}_Y(k) = \mathcal{M}(k)\tilde{W}_Y(k)$, where $Y$ stands for different filters such as $L$, $s$, $*$ which we will encounter below. We let $P_\phi(k)$ denote the power spectrum of $\phi$, and $P_m(k) = \mathcal{M}^2(k)P_\phi(k)$ the matter power spectrum (again dropping the argument $z$).

We will need various spectral moments of the density field. We define

$$\sigma_Y^2 = \int \frac{d^3k}{(2\pi)^3}P_m(k)\tilde{W}_Y^2(k) \qquad (11)$$

for any filter $W_Y$. Further, given a weighting function $f(k)$, we define

$$\sigma_{f,Y}^2 = \int \frac{d^3k}{(2\pi)^3}f(k)P_m(k)\tilde{W}_Y^2(k). \qquad (12)$$

It will also be useful to define nonlocal transformations of the density field. Again, given a function $f(k)$, we let

$$\delta_{f,Y}(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}e^{i\mathbf{k}\mathbf{x}}f(k)\tilde{W}_Y(k)\delta(\mathbf{k}). \qquad (13)$$

In particular, this yields

$$\langle\delta_{f,Y}\delta_Y\rangle = \sigma_{f,Y}^2. \qquad (14)$$

## III. BEYOND THE SQUEEZED LIMIT

We begin with deriving the correlator Eq. (5) and expanding it in the squeezed limit. More details are provided in the Appendix. Throughout, we will work to leading order in the dimensionless amplitude of non-Gaussianity $f_{\rm NL}$. At this order and for the models we consider, the only relevant $N$-point functions are the power spectrum and bispectrum.

Expressed in terms of the bispectrum of primordial perturbations $B_\phi$, defined through

$$\langle\phi_\mathbf{k}\phi_{\mathbf{k}_1}\phi_{\mathbf{k}_2}\rangle = (2\pi)^3\delta_D(\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2)B_\phi(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2), \qquad (15)$$

the term we are interested in becomes

$$\langle\delta_L(1)\delta_L^2(2)\rangle = \int\frac{d^3k}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{r}}\mathcal{M}_L(k)\int\frac{d^3k_1}{(2\pi)^3}\mathcal{M}_L(k_1)$$
$$\times \mathcal{M}_L(|\mathbf{k}+\mathbf{k}_1|)B_\phi(k, k_1, |\mathbf{k}+\mathbf{k}_1|). \qquad (16)$$

Our goal is to obtain an expression that is separable in $r$ and $R_L$. Neglecting the $R_L$ dependence introduced through $\mathcal{M}_L(k)$, which is irrelevant if $r \gg R_L$ (see Sec. VII), separability in $r$ and $R_L$ is equivalent to having an integrand separable in $\mathbf{k}$ and $\mathbf{k}_1$. As a necessary prerequisite, we assume that the bispectrum $B_\phi$ is given in separable form,

$$B_\phi(k_1, k_2, k_3) = \sum_\alpha A_\alpha[F_\alpha^{(1)}(k_1)F_\alpha^{(2)}(k_2)F_\alpha^{(3)}(k_3) + 5\text{ perm}], \qquad (17)$$

where $A_\alpha$ are constants and the six permutations guarantee the symmetry of $B_\phi$ in its arguments. The sum $\alpha$ runs over however many terms are necessary to adequately approximate the bispectrum in separable form. We further define

$$\tilde{F}_\alpha^{(i)}(k) = \mathcal{M}_L(k)F_\alpha^{(i)}(k). \qquad (18)$$

Equation (16) then becomes

$$\langle\delta_L(1)\delta_L^2(2)\rangle = \sum_\alpha A_\alpha\left\{\int\frac{d^3k}{(2\pi)^3}\tilde{F}^{(1)}(k)e^{i\mathbf{k}\cdot\mathbf{r}}\int\frac{d^3k_1}{(2\pi)^3}\tilde{F}_\alpha^{(2)}(k_1)\right.$$
$$\left.\times \tilde{F}_\alpha^{(3)}(|\mathbf{k}_1+\mathbf{k}|) + 5\text{ perm}\right\}. \qquad (19)$$

We now expand the $k_1$ integrand in powers of $k/k_1$ up to second order. As shown in the Appendix,

$$\langle\delta_L(1)\delta_L^2(2)\rangle$$
$$= \sum_\alpha A_\alpha\left\{\int\frac{d^3k}{(2\pi)^3}\tilde{F}_\alpha^{(1)}(k)e^{i\mathbf{k}\cdot\mathbf{r}}\int\frac{d^3k_1}{(2\pi)^3}\left[2\tilde{F}_\alpha^{(2)}(k_1)\tilde{F}_\alpha^{(3)}(k_1)\right.\right.$$
$$+\frac{1}{2}\frac{k^2}{k_1^2}\tilde{F}_\alpha^{(2)}(k_1)[(1-\mu^2)k_1\tilde{F}_\alpha^{\prime(3)}(k_1) + \mu^2 k_1^2\tilde{F}_\alpha^{\prime\prime(3)}(k_1)]$$
$$\left.+\frac{1}{2}\frac{k^2}{k_1^2}\tilde{F}_\alpha^{(3)}(k_1)[(1-\mu^2)k_1\tilde{F}_\alpha^{\prime(2)}(k_1) + \mu^2 k_1^2\tilde{F}_\alpha^{\prime\prime(2)}(k_1)]\right]$$
$$\left.+\{(123)\to(231)\} + \{(123)\to(312)\}\right\},$$

where primes denote derivatives with respect to $k_1$. This expression is now in the desired separable form. It is valid up to terms of order $k^4/k_1^4$ as the cubic terms drop out just like the linear terms did. We can make this result more obvious and compact by introducing the notation

$$\xi_{\alpha L}^{(i)}(r) \equiv \int\frac{d^3k}{(2\pi)^3}\mathcal{M}_L(k)F_\alpha^{(i)}(k)e^{i\mathbf{k}\cdot\mathbf{r}} \qquad (20)$$

$$\xi^{(i)}_{\nabla^2 \alpha L}(r) \equiv \int \frac{d^3k}{(2\pi)^3} \mathcal{M}_L(k)k^2 F^{(i)}_\alpha(k)e^{i\mathbf{k}\cdot\mathbf{r}} \tag{21}$$

$$\sigma^{2(ij)}_{\alpha L} \equiv \frac{1}{2\pi^2} \int_0^\infty dk_1 k_1^2 \mathcal{M}_L^2(k_1)F^{(i)}_\alpha(k_1)F^{(j)}_\alpha(k_1) \tag{22}$$

$$\sigma^{2(ij)}_{X\alpha L} \equiv \frac{1}{12\pi^2} \int_0^\infty dk_1 \{\tilde{F}^{(i)}_\alpha(k_1)[2k_1\tilde{F}'^{(j)}_\alpha(k_1) + k_1^2\tilde{F}''^{(j)}_\alpha(k_1)]$$
$$+ (i) \leftrightarrow (j)\}. \tag{23}$$

Note that for all spectral moments, $\sigma^{2(ij)} = \sigma^{2(ji)}$, and that the derivatives in Eq. (23) act on both $F^{(i)}_\alpha(k_1)$ and $\mathcal{M}_L(k_1)$. With this, we obtain

$$\langle \delta_L(1)\delta_L^2(2)\rangle = \sum_\alpha A_\alpha \{2\xi^{(1)}_{\alpha L}(r)\sigma^{2(23)}_{\alpha L} + \xi^{(1)}_{\nabla^2 \alpha L}(r)\sigma^{2(23)}_{X\alpha L}$$
$$+ \{(123) \rightarrow (231)\} + \{(123) \rightarrow (312)\}\}, \tag{24}$$

where the second line denotes the two remaining cyclic permutations. Apart from the residual $R_L$ dependence in $\xi_{\alpha L}, \xi_{\nabla^2 \alpha L}$, this expression is fully separable in $r$ and $R_L$ as desired. The expansion in $k/k_1$ can of course also be taken to higher order if necessary. We will discuss the necessity of this below.

Given the appearance of $\sigma^2_{\alpha L}$ and $\sigma^2_{X\alpha L}$ in Eq. (27), the non-Gaussian correction to $\xi_h$ is strongly $R_L$ dependent. In the next section, we will address this issue through renormalized bias parameters.

## A. Primordial non-Gaussianity of the local type

Equation (24) is general, but somewhat abstract. In order to clarify its physical significance, we now specialize to the case of local primordial non-Gaussianity, for which the bispectrum is given by

$$B_\phi(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2f_{NL}[P_\phi(k_1)P_\phi(k_2) + 2 \text{ perm}]. \tag{25}$$

Since $B_\phi$ is already separable, we only have one term in the sum Eq. (17), with $A_\alpha = A = f_{NL}$, and

$$F^{(1)}(k) = F^{(2)}(k) = P_\phi(k); \qquad F^{(3)}(k) = 1. \tag{26}$$

The permutation $\xi^{(3)}_{\alpha L}(r)\sigma^{2(12)}_{\alpha L}$ in Eq. (24) is suppressed (for scale-invariant $P_\phi$) by $(k/k_1)^3$ relative to the other two identical permutations. In keeping with our treatment up to $(k/k_1)^4$, we will retain the leading contribution from this term. We then obtain

$$\langle \delta_L(1)\delta_L^2(2)\rangle = 4f_{NL}\sigma^2_L \xi_{\phi\delta_L}(r) + 2f_{NL}\sigma^2_{X,L}\xi_{\varphi\delta_L}(r)$$
$$+ 2f_{NL}\sigma^{2(12)}_\alpha \xi^{(3)}_{\alpha L}(r), \tag{27}$$

where we have defined

$$\xi_{\phi\delta_L}(r) = \frac{d^3k}{(2\pi)^3} \mathcal{M}_L(k)P_\phi(k)e^{i\mathbf{k}\cdot\mathbf{r}} = \langle\phi(1)\delta_L(2)\rangle \tag{28}$$

$$\xi_{\varphi\delta_L}(r) = \frac{d^3k}{(2\pi)^3} \mathcal{M}_L(k)k^2 P_\phi(k)e^{i\mathbf{k}\cdot\mathbf{r}} = \langle\varphi(1)\delta_L(2)\rangle$$

$$X(k) = \frac{1}{6k^2}\{\mathcal{M}_L^{-1}(k)[2k\mathcal{M}'_L(k) + k^2\mathcal{M}''_L(k)]$$
$$+ (P_\phi\mathcal{M}_L)_k^{-1}[2k(P_\phi\mathcal{M}_L)'_k + k^2(P_\phi\mathcal{M}_L)''_k]\}, \tag{29}$$

and $\sigma^2_L, \sigma^2_{X,L}$ are defined through Eqs. (11) and (12). Here we have introduced $\varphi(\mathbf{x}) \equiv -\nabla^2\phi(\mathbf{x})$. The contribution from the third permutation is given by

$$\xi^{(3)}_{\alpha L}(r) \equiv \int \frac{d^3k}{(2\pi)^3} \mathcal{M}_L(k)e^{i\mathbf{k}\cdot\mathbf{r}} \tag{30}$$
$$\sigma^{2(12)}_{\alpha L} \equiv \frac{1}{2\pi^2} \int_0^\infty dk_1 k_1^2 P_\phi(k_1)P_m(k_1)\tilde{W}_L^2(k_1).$$

The first (leading) term in Eq. (27) is well known and agrees with that derived in [15]. This term can effectively be described as a rescaling of the local density field,

$$\delta(\mathbf{x}) \rightarrow [1 + 2f_{NL}\phi(\mathbf{x})]\delta(\mathbf{x}), \tag{31}$$

which leads to Eq. (2). On the other hand, the second (subleading) term in Eq. (27) can be seen as coupling the Laplacian of $\phi$ to the density field. However, this coupling is scale dependent, i.e., when performing a Fourier transform in a local patch where $\varphi$ can be considered constant, we have

$$\delta(\mathbf{k}) \rightarrow [1 + f_{NL}X(k)\varphi(\mathbf{x})]\delta(\mathbf{k}). \tag{32}$$

This effect is of course suppressed with respect to $\phi$ in the large-scale limit. In Fourier space on large scales $\mathcal{M}(k)P_\phi(k) \propto k^{-2}P_m(k)$, so that $\xi_{\phi\delta}(r)$ grows with respect to $\xi_L(r)$ on large scales, while $\xi_{\varphi\delta,L}(r) \sim \xi_L(r)$ (the nontrivial transfer function leads to departures on scales smaller than about $100h^{-1}$ Mpc).

## IV. SCALE-DEPENDENT BIAS

As shown in [15], one can absorb the leading term $\propto \sigma^2_L$ in Eq. (27) by introducing a dependence of the tracer density on the local variance of the small-scale density field. We define the small-scale density field as the local fluctuations around the coarse-grained field $\delta_L$,

$$\delta_s(\mathbf{x}) \equiv \delta_*(\mathbf{x}) - \delta_L(\mathbf{x})$$
$$= \int d^3\mathbf{y}[W_*(\mathbf{x} - \mathbf{y}) - W_L(\mathbf{x} - \mathbf{y})]\delta(\mathbf{y})$$
$$= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \tilde{W}_s(k)\delta(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{33}$$

$$\tilde{W}_s(k) = \tilde{W}_*(k) - \tilde{W}_L(k), \tag{34}$$

where we have introduced a fixed small smoothing scale $R_*$. Thus, we write

$$n_h[\delta_L(\mathbf{x})] \rightarrow n_h[\delta_L(\mathbf{x}), y_*(\mathbf{x})] \qquad (35)$$

$$y_*(\mathbf{x}) \equiv \frac{1}{2}\left(\frac{\delta_s^2(\mathbf{x})}{\sigma_s^2} - 1\right), \qquad (36)$$

where the subscript $*$ refers to the smoothing scale $R_*$, $\langle y_* \rangle = 0$, and the factor of $1/2$ is included to obtain expressions which conform to standard convention. A very similar derivation to that of Eq. (24) yields (Appendix)

$$\langle \delta_L(1) y_*(2) \rangle = \sum_\alpha A_\alpha \Big\{ \xi_{\alpha L}^{(1)}(r) \frac{\sigma_{\alpha s}^{2(23)}}{\sigma_s^2} \\ + \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \frac{\sigma_{X\alpha s}^{2(23)}}{\sigma_s^2} + 2 \text{ perm} \Big\}, \qquad (37)$$

where a subscript $s$ indicates that the filter function $\tilde{W}_L$ should be replaced with $\tilde{W}_s$ in Eqs. (22) and (23). In the case of local primordial non-Gaussianity, this again reduces to

$$\langle \delta_L(1) y_*(2) \rangle = 2 f_{\text{NL}} \Big\{ \xi_{\phi\delta_L}(r) + \frac{1}{2} \xi_{\varphi\delta_L}(r) \frac{\sigma_{Xs}^2}{\sigma_s^2} \\ + \frac{1}{2} \xi_{\alpha L}^{(3)} \frac{\sigma_{\alpha s}^{2(12)}}{\sigma_s^2} \Big\}. \qquad (38)$$

While the renormalization of the $R_L$ dependence in Eq. (24) works for the leading term as we will see now, we already notice two issues with Eq. (38) in comparison with Eq. (27). First, the second and third term in Eq. (38) depend on $R_L$, since $\tilde{W}_s$ is defined with respect to $\tilde{W}_L$ [Eq. (34)]. Thus, the final expression cannot simply involve $\langle \delta_L(1) y_*(2) \rangle$. Second, there is no reason why the relative magnitude of the three terms in Eq. (38) should be the same as for the three terms in Eq. (27), so that we cannot expect all terms in Eq. (27) to be absorbed by $\langle \delta_L(1) y_*(2) \rangle$. These issues will be resolved below.

### A. Bivariate bias expansion

We begin by reexamining the bivariate bias expansion of [15] (which in the local case at lowest order is equivalent to what was presented in previous papers [18,19]). Throughout this section and the next, we will restrict to a single separable contribution $\alpha$ and a single permutation (123) of Eq. (24) [note that this corresponds to two permutations in Eq. (17)]. The complete prescription for tracer clustering, presented in Sec. IV C below, will then involve a sum over the contributions from different $\alpha$ and permutations.

In terms of the "bare" bias parameters $c_{nm}$, that is, the coefficients of the Taylor series of $n_h$ in $\delta_L$, $y_*$, the tree-level expression for the tracer correlation function in the bivariate PBS expansion is [see Eq. (109) in [15]]

$$\xi_h^{\text{bare}}(r) = \frac{1}{\mathcal{N}^2} \{ c_{10}^2 \xi_L(r) + c_{10} c_{20} \langle \delta_L(1) \delta_L^2(2) \rangle \\ + 2 c_{10} c_{01} \langle \delta_L(1) y_*(2) \rangle \} \\ = \frac{1}{\mathcal{N}^2} \Big\{ c_{10}^2 \xi_L(r) + c_{10} c_{20} A_\alpha [2 \xi_{\alpha L}^{(1)}(r) \sigma_{\alpha L}^{2(23)} \\ + \xi_{\nabla^2 \alpha L}^{(1)}(r) \sigma_{X\alpha L}^{2(23)} ] \\ + 2 c_{10} c_{01} A_\alpha \Big[ \xi_{\alpha L}^{(1)}(r) \frac{\sigma_{\alpha s}^{2(23)}}{\sigma_s^2} + \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \frac{\sigma_{X\alpha s}^{2(23)}}{\sigma_s^2} \Big] \Big\}, \qquad (39)$$

where

$$\mathcal{N} \equiv \sum_{n,m=0}^{\infty} \frac{c_{nm}}{n! m!} \langle \delta_L^n y_*^m \rangle. \qquad (40)$$

Let us focus on the leading terms in the squeezed limit first. We need to generalize the definition of the renormalized bias parameters $b_{NM}$ to the case of the general separable bispectrum Eq. (17).

Consider a rescaling of the Fourier-space density field given by

$$\delta(\mathbf{k}_1) \rightarrow \delta(\mathbf{k}_1)[1 + \varepsilon F(k_1)]$$

$$F(k_1) = \frac{F_\alpha^{(2)}(k_1) F_\alpha^{(3)}(k_1)}{P_\phi(k_1)}. \qquad (41)$$

Further, we include a shift in the density perturbation $\delta \rightarrow \delta + D$, corresponding to adding a uniform matter density of $D\bar{\rho}$. We then define the bivariate bias parameters $b_{NM}$ as the response of the mean tracer density to this uniform matter density and the rescaling of the initial density field under Eq. (41) (both to be evaluated at fixed proper time),

$$b_{NM} \equiv \frac{1}{\langle n_h \rangle_{D=0,\varepsilon=0}} \frac{\partial^{N+M} \langle n_h \rangle_{D,\varepsilon}}{\partial D^N \partial \varepsilon^M} \Big|_{D=0,\varepsilon=0}. \qquad (42)$$

Note that $b_{NM}$ should really be denoted $b_{NM}^{\alpha(23)}$ here, since the definition refers to the specific rescaling in Eq. (41). However, since we are only dealing with a single $\alpha$ and permutation here and in Sec. IV B, we will drop this designation for notational clarity. Note further that the dimension of $b_{01}$ is such that the contribution to the correlation function is dimensionless. In particular,

$$\dim(b_{01}) = \dim[F(k_1)] = \frac{\text{Mpc}^3}{\dim(F_\alpha^{(1)})}, \qquad (43)$$

where we have assumed that the $A_\alpha$ are dimensionless so that $\dim(F_\alpha^{(1)} F_\alpha^{(2)} F_\alpha^{(3)}) = \text{Mpc}^6$ [Eq. (17)]. Under the rescaling Eq. (41), we have [using the definition Eq. (13)]

$$\delta_L(\mathbf{x}) \rightarrow \delta_L(\mathbf{x}) + \varepsilon \delta_{F,L}(\mathbf{x})$$

$$y_*(\mathbf{x}) \rightarrow y_*(\mathbf{x}) + \frac{\varepsilon}{\sigma_s^2} \delta_s(\mathbf{x}) \delta_{F,s}(\mathbf{x}) + \frac{\varepsilon^2}{2\sigma_s^2} \delta_{F,s}^2(\mathbf{x}). \qquad (44)$$

Note that

$$\langle\delta_Y(\mathbf{x})\delta_{F,Y}(\mathbf{x})\rangle = \sigma_{F,Y}^2 = \sigma_{\alpha Y}^{2(23)} \qquad (45)$$

for $Y = L, s, \dots$. Using Eq. (35), we obtain

$$\langle n_h\rangle(D, \varepsilon) = \langle n_h(0)\rangle\sum_{n,m}\frac{c_{nm}}{n!m!}\Big\langle[(\delta_L + \varepsilon\delta_{F,L})\delta_L + D]^n$$

$$\times\Big[y_* + \frac{\varepsilon}{\sigma_s^2}\delta_s\delta_{F,s} + \frac{\varepsilon^2}{\sigma_s^2}\delta_{F,s}^2\Big]^m\Big\rangle.$$

To the order we are interested in, we only need $b_{01}$, which is given by

$$b_{01} = \frac{1}{\mathcal{N}}\sum_{n,m}\frac{c_{nm}}{n!m!}\Big(n\langle\delta_{F,L}\delta_L^{n-1}y_*^m\rangle + \frac{m}{\sigma_s^2}\langle\delta_L^n\delta_s\delta_{F,s}y_*^{m-1}\rangle\Big)$$

$$= \frac{1}{\mathcal{N}}\Big(c_{01}\frac{\sigma_{\alpha s}^{2(23)}}{\sigma_s^2} + c_{20}\sigma_{\alpha L}^{2(23)} + \mathcal{O}(\delta^3, f_{\text{NL}}^2)\Big). \qquad (46)$$

Following the same reasoning as in [15], our guess for the tree-level tracer correlation function written in terms of renormalized bias parameters is

$$\xi_h^{\text{renorm}}(r) = b_{10}^2\xi_L(r) + 2b_{10}b_{01}\frac{\sigma_s^2}{\sigma_{\alpha s}^{2(23)}}\langle\delta_L(1)y_*(2)\rangle$$

$$= b_{10}^2\xi_L(r) + 2b_{10}b_{01}A_\alpha\Big\{\xi_{\alpha L}^{(1)}(r)$$

$$+ \frac{1}{2}\xi_{\nabla^2\alpha L}^{(1)}(r)\frac{\sigma_{X\alpha s}^{2(23)}}{\sigma_{\alpha s}^{2(23)}}\Big\}, \qquad (47)$$

where the prefactor in front of $\langle\delta_L(1)y_*(2)\rangle$ takes into account that the tree-level relation between $b_{01}$ and $c_{01}$ is $b_{01} = c_{01}\sigma_{\alpha s}^{2(23)}/\sigma_s^2$. In the second line we have used Eq. (37). The fact that renormalization is not successful at subleading order can already be seen from the last term in this equation. This depends on $R_L$ through the spectral moments, since the kernel $\tilde{W}_s$ depends on $R_L$. Thus, we assume that a proper renormalization will absorb this term as well and write our updated guess as

$$\xi_h^{\text{renorm}}(r) = b_{10}^2\xi_L(r) + 2b_{10}b_{01}A_\alpha\xi_{\alpha L}^{(1)}(r). \qquad (48)$$

Inserting the expression for $b_{01}$ [Eq. (46)] yields

$$\xi_h^{\text{renorm}}(r) = \frac{1}{\mathcal{N}^2}\Big\{c_{10}^2\xi_L(r) + 2c_{10}c_{01}\frac{\sigma_{\alpha s}^{2(23)}}{\sigma_s^2}A_\alpha\xi_{\alpha L}^{(1)}(r)$$

$$+ 2c_{10}c_{20}\sigma_{\alpha L}^{2(23)}A_\alpha\xi_{\alpha L}^{(1)}(r)\Big\}. \qquad (49)$$

If the renormalized bias $b_{01}$ defined with respect to the transformation Eq. (41) properly removes all $R_L$ dependence of the tracer two-point function, Eq. (48) should match the bare bias expansion Eq. (39). Let us thus take the difference:

$$\xi_h^{\text{bare}}(r) - \xi_h^{\text{renorm}}(r)$$

$$= A_\alpha\xi_{\nabla^2\alpha L}^{(1)}(r)c_{10}\Big[c_{20}\sigma_{X\alpha L}^{2(23)} + c_{01}\frac{\sigma_{X\alpha s}^{2(23)}}{\sigma_s^2}\Big]. \qquad (50)$$

All terms here are subleading in the squeezed limit, that is, at leading order in the squeezed limit the renormalization

works in the same way as shown in [15]. However, when going beyond the squeezed limit for general primordial non-Gaussianity, the description of the tracer density as a bivariate function $n_h[\delta_L, y_*]$ is not sufficient.

### B. Trivariate bias expansion

We are seeking a third local parameter (besides $\delta_L$ and $y_*$) which the tracer density in general depends on, and which will allow us to absorb the $R_L$-dependent terms in Eq. (50). In the local model, the residual $R_L$ dependence that we encounter when going beyond the squeezed limit is induced by the fact that in this case, the effect of long-wavelength modes on small-scale modes is not simply to rescale them uniformly, but to also change the shape of the small-scale power spectrum [Eq. (29)]. More generally, small-scale modes are rescaled differently as function of the correlation scale $r$, as the relative importance of the terms given by $\xi_{\alpha L}^{(1)}(r)$ and $\xi_{\nabla^2\alpha L}^{(1)}(r)$ changes. Thus, we have to allow for a dependence of the tracer density not only on the local variance of small-scale fluctuations, but also on their shape. As a proxy for the power spectrum shape, we will use

$$\mu_*(\mathbf{x}) \equiv \frac{d}{d\ln R_*}y_*(\mathbf{x}). \qquad (51)$$

This choice is useful because transformations of the density field of the form

$$\delta_L \rightarrow \delta_L + D; \qquad \delta(\mathbf{x}) \rightarrow [1 + \iota]\delta(\mathbf{x}) \qquad (52)$$

leave $\mu_*(\mathbf{x})$ invariant. Using Eqs. (51) and (37) we obtain, again for one of the three cyclic permutation of the contribution $\alpha$,

$$\langle\delta_L(1)\mu_*(2)\rangle = A_\alpha\Big\{\xi_{\alpha L}^{(1)}(r)\frac{d}{d\ln R_*}\Big(\frac{\sigma_{\alpha s}^{2(23)}}{\sigma_s^2}\Big)$$

$$+ \frac{1}{2}\xi_{\nabla^2\alpha L}^{(1)}(r)\frac{d}{d\ln R_*}\Big(\frac{\sigma_{X\alpha s}^{2(23)}}{\sigma_s^2}\Big)\Big\}. \qquad (53)$$

For local non-Gaussianity, $\sigma_{\alpha s}^2 = \sigma_s^2$, so that the first term drops out, and we obtain (restoring the number of permutations)

$$\langle\delta_L(1)\mu_*(2)\rangle \overset{\text{local}}{=} 4f_{\text{NL}}\frac{1}{2}\xi_{\varphi L}^{(1)}(r)\frac{d}{d\ln R_*}\Big(\frac{\sigma_{Xs}^2}{\sigma_s^2}\Big)$$

$$= 2f_{\text{NL}}\xi_{\varphi L}^{(1)}(r)\frac{\sigma_{Xs}^2}{\sigma_s^2}\Big(\frac{d\ln\sigma_{Xs}^2}{d\ln R_*} - \frac{d\ln\sigma_s^2}{d\ln R_*}\Big).$$

Let us now adopt $\mu_*(\mathbf{x})$ as third local parameter:

$$n_h[\delta_L(\mathbf{x}), y_*(\mathbf{x})] \rightarrow n_h[\delta_L(\mathbf{x}), y_*(\mathbf{x}), \mu_*(\mathbf{x})]. \qquad (54)$$

We again consider the scale-dependent transformation Eq. (41), and in addition a further, independent scale-dependent transformation of the density field

$$\delta(\mathbf{k}) \rightarrow [1 + \iota f(k)]\delta(\mathbf{k}) \qquad (55)$$

Here we leave the function $f(k)$ free for the moment. This leads to

$$\delta_L(\mathbf{x}) \rightarrow \delta_L(\mathbf{x}) + \iota \delta_{f,L}(\mathbf{x})$$

$$y_*(\mathbf{x}) \rightarrow y_*(\mathbf{x}) + \frac{\iota}{\sigma_s^2} \delta_s(\mathbf{x}) \delta_{f,s}(\mathbf{x}) + \frac{\iota^2}{\sigma_s^2} \delta_{f,s}^2(\mathbf{x})$$

$$\mu_*(\mathbf{x}) \rightarrow \mu_*(\mathbf{x}) + \frac{d}{d \ln R_*}\left[ \frac{\iota}{\sigma_s^2} \delta_s(\mathbf{x}) \delta_{f,s}(\mathbf{x}) + \frac{\iota^2}{\sigma_s^2} \delta_{f,s}^2(\mathbf{x}) \right]. \tag{56}$$

In keeping with the leading order treatment in $f_{\rm NL}$, we will only consider the first derivative of the mean tracer density with respect to $\iota$. Thus, we can drop the $\iota^2$ terms in Eq. (56). We define

$$b_{NML} = \frac{1}{\langle n_h(0) \rangle} \frac{\partial^{N+M+L} \langle n_h(D, \varepsilon, \iota) \rangle}{\partial D^N \varepsilon^M \iota^L}\bigg|_0, \tag{57}$$

where the mean tracer number density is given by

$$\langle n_h \rangle (D, \varepsilon, \iota)$$
$$= \langle n_h(0) \rangle \sum_{nml} \frac{c_{nml}}{n!m!l!} \bigg\langle \left[ (1 + \varepsilon F(k) + \iota f(k)) \delta_L + D \right]^n$$
$$\times \left[ y_* + \varepsilon \frac{\delta_s \delta_{F,s}}{\sigma_s^2} + \iota \frac{\delta_s \delta_{f,s}}{\sigma_s^2} \right]^m \left[ \mu_* + \varepsilon \frac{d}{d \ln R_*}\left( \frac{\delta_s \delta_{F,s}}{\sigma_s^2} \right) \right.$$
$$\left. + \iota \frac{d}{d \ln R_*}\left( \frac{\delta_s \delta_{f,s}}{\sigma_s^2} \right) \right]^l \bigg\rangle. \tag{58}$$

We thus obtain

$$b_{N00} = b_{N0}$$

$$b_{010} = \frac{1}{\langle n_h \rangle(0)} \frac{\partial \langle n_h \rangle(D, \iota, \varepsilon)}{\partial \varepsilon}\bigg|_0$$
$$= \frac{1}{\mathcal{N}}\left[ c_{200} \sigma_{F,L}^2 + c_{010} \frac{\sigma_{F,s}^2}{\sigma_s^2} + c_{001} \frac{d}{d \ln R_*}\left( \frac{\sigma_{F,s}^2}{\sigma_s^2} \right) \right], \tag{59}$$

and

$$b_{001} = \frac{1}{\langle n_h \rangle(0)} \frac{\partial \langle n_h \rangle(D, \iota, \varepsilon)}{\partial \iota}\bigg|_0$$
$$= \frac{1}{\mathcal{N}}\left[ c_{200} \sigma_{f,L}^2 + c_{010} \frac{\sigma_{f,s}^2}{\sigma_s^2} + c_{001} \frac{d}{d \ln R_*}\left( \frac{\sigma_{f,s}^2}{\sigma_s^2} \right) \right], \tag{60}$$

where we have expanded to order $\delta_L^2$. Note the nontrivial coefficient multiplying $c_{001}$ in Eq. (60). We thus have to divide $b_{001}$ by this coefficient when writing down the contributions to $\xi_h$ that correlate $\mu_*$ with other perturbations.

We now update our ansatz for the renormalized tracer correlation [Eq. (48)] to the trivariate case. Since we found in the previous section that the bivariate case (neglecting the dependence of $n_h$ on $\mu_*$) successfully produced an $R_L$-independent expression for the leading squeezed-limit contribution, we expect that only the subleading contribution of $\langle \delta_L(1) \mu_*(2) \rangle$ [Eq. (53)] will contribute. That is, we expect the leading contribution in Eq. (53) to be absorbed by $b_{010}$. Thus, our expectation is

$$\xi_h^{\rm tri}(r) = b_{100}^2 \xi_L(r) + 2b_{100} b_{010} A_\alpha \xi_{\alpha L}^{(1)}(r)$$
$$+ 2b_{100} b_{001} A_\alpha \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \frac{d(\sigma_{X\alpha s}^{2(23)}/\sigma_s^2)/d \ln R_*}{d(\sigma_{f,s}^2/\sigma_s^s)/d \ln R_*}. \tag{61}$$

The last term is in general only guaranteed to be $R_L$-independent if $\sigma_{X\alpha s}^{2(23)} = \sigma_{f,s}^2$, that is, if

$$k_1^2 P_m(k_1) f(k_1) = \frac{1}{6} \tilde{F}_\alpha^{(2)}(k_1)[2k_1 \tilde{F}_\alpha^{\prime(3)}(k_1) + k_1^2 \tilde{F}_\alpha^{\prime\prime(3)}(k_1)]$$
$$+ (2) \leftrightarrow (3). \tag{62}$$

Thus, we will fix this as our choice of $f(k)$, and in the following derivation set $\sigma_{X\alpha Y}^{2(23)} = \sigma_{f,Y}^2$, where $Y = s, L$. Note that the dimension of $b_{001}$ is equal to the dimension of $f$ and is given by

$$\dim(b_{001}) = \dim\left( \frac{F_\alpha^{(2)}(k_1) F_\alpha^{(3)}(k_1)}{k_1^2 P_m(k_1)} \right) = \frac{{\rm Mpc}^5}{\dim(F_\alpha^{(1)})}, \tag{63}$$

that is, the dimension of $b_{01}$ times ${\rm Mpc}^2$. The bare expansion then becomes [using $\sigma_{\alpha Y}^{2(23)} = \sigma_{F,Y}^2$ from Eq. (45)]

$$\xi_h^{\rm bare}(r) = \frac{1}{\mathcal{N}^2}\{c_{100}^2 \xi_L(r) + c_{100} c_{200} \langle \delta_L(1) \delta_L^2(2) \rangle + 2c_{100} c_{010} \langle \delta_L(1) y_*(2) \rangle + 2c_{100} c_{001} \langle \delta_L(1) \mu_*(2) \rangle\}$$

$$= \frac{1}{\mathcal{N}^2}\left\{ c_{100}^2 \xi_L(r) + c_{100} c_{200} A_\alpha [2\xi_{\alpha L}^{(1)}(r) \sigma_{F,L}^2 + \xi_{\nabla^2 \alpha L}^{(1)}(r) \sigma_{f,L}^2] + 2c_{100} c_{010} A_\alpha \left[ \xi_{\alpha L}^{(1)}(r) \frac{\sigma_{F,s}^2}{\sigma_s^2} + \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \frac{\sigma_{f,s}^2}{\sigma_s^2} \right] \right.$$

$$\left. + 2c_{100} c_{001} A_\alpha \left[ \xi_{\alpha L}^{(1)}(r) \frac{d}{d \ln R_*}\left( \frac{\sigma_{F,s}^2}{\sigma_s^2} \right) + \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \frac{d}{d \ln R_*}\left( \frac{\sigma_{f,s}^2}{\sigma_s^2} \right) \right] \right\}$$

$$= \frac{1}{\mathcal{N}^2}\left\{ c_{100}^2 \xi_L(r) + 2A_\alpha \xi_{\alpha L}^{(1)}(r) c_{100}\left[ c_{200} \sigma_{F,L}^2 + c_{010} \frac{\sigma_{F,s}^2}{\sigma_s^2} + c_{001} \frac{d}{d \ln R_*}\left( \frac{\sigma_{F,s}^2}{\sigma_s^2} \right) \right] \right.$$

$$\left. + A_\alpha \xi_{\nabla^2 \alpha L}^{(1)}(r) c_{100}\left[ c_{200} \sigma_{f,L}^2 + c_{010} \frac{\sigma_{f,s}^2}{\sigma_s^2} + c_{001} \frac{d}{d \ln R_*}\left( \frac{\sigma_{f,s}^2}{\sigma_s^2} \right) \right] \right\}. \tag{64}$$

On the other hand, inserting Eqs. (59) and (60) into Eq. (61) yields

$$
\xi_h^{\mathrm{tri}}(r) = \frac{1}{\mathcal{N}^2} \Bigg\{ c_{100}^2 \xi_L(r) + 2c_{100} \Bigg[ c_{200} \sigma_{F,L}^2 + c_{010} \frac{\sigma_{F,s}^2}{\sigma_s^2} 
$$
$$
+ c_{001} \frac{d}{d \ln R_*} \Big( \frac{\sigma_{F,s}^2}{\sigma_s^2} \Big) \Bigg] A_\alpha \xi_{\alpha L}^{(1)}(r)
$$
$$
+ 2c_{100} \Bigg[ c_{200} \sigma_{f,L}^2 + c_{010} \frac{\sigma_{f,s}^2}{\sigma_s^2} 
$$
$$
+ c_{001} \frac{d}{d \ln R_*} \Big( \frac{\sigma_{f,s}^2}{\sigma_s^2} \Big) \Bigg] A_\alpha \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \Bigg\}. \quad (65)
$$

We thus find that the trivariate renormalized expression,

$$
\xi_h^{\mathrm{tri}}(r) = b_{100}^2 \xi_L(r) + 2b_{100} b_{010} A_\alpha \xi_{\alpha L}^{(1)}(r)
$$
$$
+ 2b_{100} b_{001} A_\alpha \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \quad (66)
$$

has successfully absorbed all $R_L$ dependence from the bare expansion including subleading terms in the squeezed limit (that is, apart from the residual dependence through $\xi_L$, $\xi_{\alpha L}^{(1)}$, $\xi_{\nabla^2 \alpha L}^{(1)}$, which is negligible on large scales).

### C. Total contribution

We are now ready to derive the full expression for the two-point tracer correlation including subleading corrections in the squeezed limit for the general separable bispectrum Eq. (17). For each $\alpha$, and each of the three cyclic permutations, we define general trivariate bias parameters through

$$
b_{NML}^{(\alpha,ij)} = \frac{1}{\langle n_h \rangle (0)} \frac{\partial^{N+M+L} \langle n_h \rangle (D, \varepsilon^{(\alpha,ij)}, \iota^{(\alpha,ij)})}{\partial D^N \partial (\varepsilon^{(\alpha,ij)})^M \partial (\iota^{(\alpha,ij)})^L} \Bigg|_0, \quad (67)
$$

where $\varepsilon^{(\alpha,ij)}$, $\iota^{(\alpha,ij)}$ are scale-dependent rescalings of the density field given by

$$
\delta(\mathbf{k}_1) \rightarrow \delta(\mathbf{k}_1) \Bigg[ 1 + \varepsilon^{(\alpha,ij)} \frac{F_\alpha^{(i)}(k_1) F_\alpha^{(j)}(k_1)}{P_\phi(k_1)} \Bigg]
$$

$$
\delta(\mathbf{k}_1) \rightarrow \delta(\mathbf{k}_1) \Bigg\{ 1 + \iota^{(\alpha,ij)} \frac{1}{6k_1^2 P_m(k_1)} 
$$
$$
\times [\tilde{F}_\alpha^{(i)}(k_1) [2k_1 \tilde{F}_\alpha'^{(j)}(k_1) + k_1^2 \tilde{F}_\alpha''^{(j)}(k_1)] + (i \leftrightarrow j) ] \Bigg\}. \quad (68)
$$

Note that $b_{NML}^{(\alpha,ij)} = b_{NML}^{(\alpha,ji)}$. Then, the renormalized tracer two-point correlation at tree level is given by

$$
\xi_h^{\mathrm{tri}}(r) = b_{100}^2 \xi_L(r) + 2b_{100} \sum_\alpha A_\alpha \Bigg[ b_{010}^{(\alpha,23)} \xi_{\alpha L}^{(1)}(r)
$$
$$
+ \frac{1}{2} b_{001}^{(\alpha,23)} \xi_{\nabla^2 \alpha L}^{(1)}(r) + \{(123) \rightarrow (231)\}
$$
$$
+ \{(123) \rightarrow (312)\} \Bigg]. \quad (69)
$$

Note that when going beyond tree level, one in general also has to include mixed contributions simultaneously involving rescalings Eq. (68) for different $\alpha$ and/or permutations. Equation (69) is accurate up to terms of order $(k/k_1)^4$ in the squeezed limit. We will make this statement more precise in the following sections. We will discuss the expected relative amplitude of $b_{010}$ and $b_{001}$ below in Sec. V B. Note again that both $b_{010}$ and $b_{001}$ are in general dimensionful [Eqs. (43) and (63)].

### D. Tracer correlations in Fourier space

In Fourier space, the tree-level result Eq. (69) can be phrased in terms of a generalized scale-dependent bias:

$$
\frac{P_h(k)}{P_m(k)} = b_{100}^2 + 2b_{100} \Delta b(k) \quad (70)
$$

$$
\Delta b(k) \equiv \sum_\alpha A_\alpha \Bigg[ b_{010}^{(\alpha,23)} \mathcal{S}_\alpha^{(1)}(k) + \frac{1}{2} b_{001}^{(\alpha,23)} k^2 \mathcal{S}_\alpha^{(1)}(k)
$$
$$
+ \{(123) \rightarrow (231)\} + \{(123) \rightarrow (312)\} \Bigg]
$$

$$
\mathcal{S}_\alpha^{(i)}(k) = \frac{F_\alpha^{(i)}(k)}{\mathcal{M}_L(k) P_\phi(k)}. \quad (71)
$$

Thus for each contribution $\alpha$ in the separable bispectrum Eq. (17), and for each of the three cyclic permutations, we have in general two independent contributions to the scale-dependent bias. The leading term in the large-scale limit $(k \rightarrow 0)$ scales as $\mathcal{S}(k)$, while the subleading term scales as $k^2 \mathcal{S}(k)$.

Equation (70) now allows us to estimate the relative magnitude of the subleading term. For each $\alpha$ and permutation,

$$
\frac{\Delta b_{\mathrm{subleading}}(k)}{\Delta b_{\mathrm{leading}}(k)} = \frac{k^2 b_{001}^{(\alpha,23)}}{2 b_{010}^{(\alpha,23)}}. \quad (72)
$$

On the other hand, the contributions from different $\alpha$ and permutations $(ijk)$, $(lmn)$ scale as

$$
\frac{\Delta b^{(\beta,l)}(k)}{\Delta b^{(\alpha,i)}(k)} = \frac{b_{010}^{\beta(mn)} \mathcal{S}_\beta^{(l)}(k)}{b_{010}^{\alpha(jk)} \mathcal{S}_\alpha^{(i)}(k)}, \quad (73)
$$

and similarly for the subleading term. Since Eqs. (69) and (70) are derived neglecting terms that, for each contribution $\alpha$ and permutation, are suppressed by $k^4$ relative to the leading contribution, only contributions which are suppressed by less than $k^4$ in Eq. (73) relative to the overall leading contribution $(i, \alpha)$ should be included in Eq. (70).

Note further that we have not included the stochasticity in Eq. (70) which in some models of primordial non-Gaussianity can become important on large scales as well [20,21].

## V. EXAMPLES AND NUMERICAL ESTIMATES

### A. Local non-Gaussianity

As before, for local non-Gaussianity we consider the two leading cylic permutations as well as the leading contribution from the third permutation. Thus,

$$\xi_h^{\text{tri}}(r) = b_{100}^2 \xi_L(r) + 2b_{100} 2f_{\text{NL}} \Bigg[ b_{010}^{\text{loc}} \xi_{\phi\delta_L}(r)$$
$$+ \frac{1}{2} b_{001}^{\text{loc}} \xi_{\varphi\delta_L}(r) + \frac{1}{2} b_{010}^{(12)\text{loc}} \xi_{\alpha L}^{(3)}(r) \Bigg]. \quad (74)$$

Here, the bias parameters are derived with respect to the rescalings

$$b_{010}^{\text{loc}}: \; \delta(\mathbf{k}_1) \to [1 + \varepsilon]\delta(\mathbf{k}_1)$$
$$b_{001}^{\text{loc}}: \; \delta(\mathbf{k}_1) \to [1 + \varepsilon X(k_1)]\delta(\mathbf{k}_1) \quad (75)$$
$$b_{010}^{(12)\text{loc}}: \; \delta(\mathbf{k}_1) \to [1 + \varepsilon P_\phi(k_1)]\delta(\mathbf{k}_1),$$

where $X(k_1)$ is defined in Eq. (29). Note that the leading term agrees with [15]. In Fourier space, this becomes

$$\frac{P_h(k)}{P_m(k)} = b_{100}^2 + 2b_{100}\Delta b(k)$$
$$\Delta b(k) = 2f_{\text{NL}}\Bigg[ b_{010}^{\text{loc}} \mathcal{M}_L^{-1}(k) + \frac{1}{2} b_{001}^{\text{loc}} k^2 \mathcal{M}_L^{-1}(k)$$
$$+ \frac{1}{2} b_{010}^{(12)\text{loc}} \mathcal{M}_L^{-1}(k) P_\phi^{-1}(k) \Bigg]. \quad (76)$$

Following Sec. IV D, we can estimate the relative magnitude of the subleading correction and the contribution from the third permutation relative to the leading term as

$$\frac{\Delta b_{\text{subleading}}^{\text{loc}}(k)}{\Delta b_{\text{leading}}^{\text{loc}}(k)} = \frac{b_{001}^{\text{loc}}}{2b_{010}^{\text{loc}}} k^2$$
$$\frac{\Delta b^{(12)\text{loc}}(k)}{\Delta b_{\text{leading}}^{\text{loc}}(k)} = \frac{b_{010}^{(12)\text{loc}}}{b_{010}^{\text{loc}}} \frac{1}{2P_\phi(k)} = \frac{k_0^{-3} b_{010}^{(12)\text{loc}}}{2\mathcal{A}_s b_{010}^{\text{loc}}} k^3, \quad (77)$$

where in the last line we have assumed a scale-invariant spectrum of $\phi$ with amplitude $\mathcal{A}_s$ defined at the pivot scale $k_0$,

$$P_\phi(k) = \mathcal{A}_s \left(\frac{k}{k_0}\right)^{-3}. \quad (78)$$

Thus, the subleading term is suppressed by a factor of $k^2$ with respect to the leading term, while the contribution from the third permutation is suppressed by $k^3$.

### B. Universal mass function

In order to quantitatively assess the importance of the subleading term in Eqs. (69) and (70), we need to estimate the magnitude of $b_{001}$. For dark matter halos this can be done accurately through N-body simulations with modified initial conditions following Eq. (68) [or Eq. (75) for local non-Gaussianity]. However, a detailed comparison with

N-body simulations is beyond the scope of this paper. Instead, we make use of a generalization of the universal mass function prescription as discussed in Sec. IV E of [15]. We write the mean abundance of tracers as

$$\bar{n}_h = \bar{n}_h(\bar{\rho}, \sigma_*, J_*) \quad (79)$$

$$J_* \equiv \frac{d \ln \sigma_*}{d \ln R_*}, \quad (80)$$

where $\sigma_*$ is the variance of the linear matter density field on scale $R_*$, and $R_*$ is related to the mass $M_*$ through $M_* = 4\pi/3\bar{\rho}R_*^3$. The Jacobian $J_*$ is present in order to convert from an interval in $\sigma_*$ to a mass interval. In this approximation, $\bar{n}_h$ is given as a function of the mean density of the Universe and the variance of the density field smoothed on a single scale $R_*$, as well as its derivative with respect to scale.

Let us again consider an individual contribution $\alpha$ and permutation (123). Under the rescaling Eq. (68), $\sigma_*$ transforms to lowest order as

$$\sigma_* \overset{\varepsilon}{\to} \sigma_* \left[1 + \varepsilon \frac{\sigma_{\alpha,*}^{2(23)}}{\sigma_*^2}\right] \qquad \sigma_* \overset{\iota}{\to} \sigma_* \left[1 + \iota \frac{\sigma_{f,*}^2}{\sigma_*^2}\right], \quad (81)$$

where $f(k)$ is defined through Eq. (62). As shown in [15] (see also [13]), the scale-dependent biases are then given by

$$b_{010} = \left[\frac{1}{\bar{n}_h} \frac{\partial \bar{n}_h}{\partial \ln \sigma_*} + 2\left(\frac{d \ln \sigma_{\alpha,*}^{2(23)}}{d \ln \sigma_*^2} - 1\right)\right] \frac{\sigma_{\alpha,*}^{2(23)}}{\sigma_*^2}$$
$$= \left[b_{010}^{\text{loc}} + 2\left(\frac{d \ln \sigma_{\alpha,*}^{2(23)}}{d \ln \sigma_*^2} - 1\right)\right] \frac{\sigma_{\alpha,*}^{2(23)}}{\sigma_*^2}, \quad (82)$$

and

$$b_{001} = \left[b_{010}^{\text{loc}} + 2\left(\frac{d \ln \sigma_{f,*}^2}{d \ln \sigma_*^2} - 1\right)\right] \frac{\sigma_{f,*}^2}{\sigma_*^2}. \quad (83)$$

Here, $b_{010}^{\text{loc}}$ is the leading scale-dependent bias parameter for local primordial non-Gaussianity, for a tracer following Eq. (79). We have assumed that the tracer density scales linearly with the Jacobian as expected physically. For such tracers, the leading and subleading bias parameters quantifying the response to general nonlocal non-Gaussianity are thus directly related to the leading bias parameter for local non-Gaussianity. If we further specialize to a universal mass function, $\bar{n}_h = \bar{n}_h(\bar{\rho}, \nu = \delta_c/\sigma_*, J_*)$, then $b_{010}^{\text{loc}} = b_{100}\delta_c$ (recall that $b_{100}$ is the Lagrangian bias).

The precise magnitude of the bias coefficients $b_{010}$, $b_{001}$ depends on the exact rescaling Eq. (68). Generally, if the main contribution to $\sigma_{\alpha,*}^{2(23)}$ comes from Fourier modes $k \sim k_*$, then

$$\sigma_{f,*}^2 \sim k_*^{-2} \sigma_{\alpha,*}^{2(23)}. \quad (84)$$

Note that depending on the shape of the rescaling $k_*$ does not necessarily have to be of order $1/R_*$. Thus, for tracers following Eqs. (79) and (72) simplifies to

$$\frac{\Delta b_{\text{subleading}}(k)}{\Delta b_{\text{leading}}(k)} \sim \left(\frac{k}{k_*}\right)^2. \qquad (85)$$

More generally, Eq. (79) implies that for scale-free bispectra for which $\mathcal{S}_\alpha^{(i)}(k) \propto k^{n_\alpha^{(i)}}$, $k_*$ is the only scale involved, so that other contributions as in Eq. (73) will all be suppressed by powers of $k/k_*$. Conversely, $k \sim k_*$ indicates the breakdown of the perturbative expansion in the squeezed limit.

### C. Numerical results

Fig. 1 shows the leading and subleading scale-dependent bias contributions for local non-Gaussianity assuming tracers following a universal mass function. Specifically, we assume a halo mass $M = 2 \times 10^{13} h^{-1} M_\odot$, so that $b_{010}^{\text{loc}} = b_{100} \delta_c = 0.44$. We have multiplied the scale-dependent bias by $(k/H_0)^2$ in order to obtain a weakly scale-dependent result. Note that even the leading term has a residual scale dependence due to the transfer function contained in $\mathcal{M}(k)$. We also show the contribution $\Delta b^{(12)\text{loc}}(k)$ from the third permutation of the local bispectrum [Eq. (77)]. The higher order contributions become important as $k \gtrsim 0.05 h$ Mpc$^{-1}$. Note in particular that the term $\Delta b^{(12)\text{loc}}$ grows rapidly towards smaller scales. We will discuss this issue in the context of the relation to previous approaches in Sec. VI.
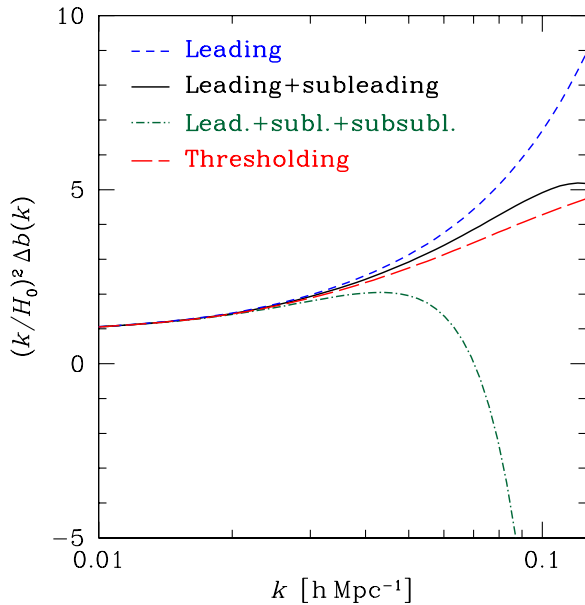


FIG. 1 (color online). Contributions to the scale-dependent bias from local non-Gaussianity ($f_{\text{NL}} = 1$) for halos with $M = 2 \times 10^{13} h^{-1} M_\odot$ at $z = 0$ ($b_{100} \simeq 0.44$) and assuming a universal mass function from the Sheth-Tormen prescription [25], scaled by $(k/H_0)^2$ to yield a scale-independent value on large scales. The dashed line shows the leading term, the solid the leading plus subleading (order $(k/k_*)^2$) term, while the dash-dotted line includes the order $(k/k_*)^3$ term. The red long-dashed line shows the prediction from the conditional PS mass function (Sec. VI).
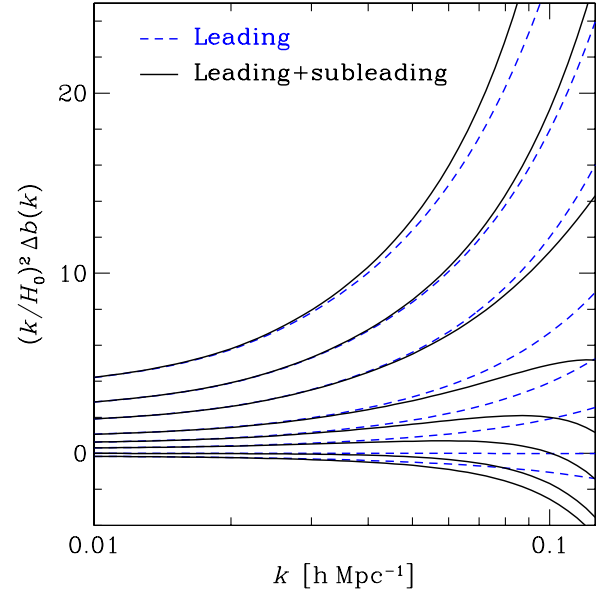


FIG. 2 (color online). Leading and sub-leading contributions to the scale-dependent bias from local non-Gaussianity, as in Fig. 1, but for different masses at $z = 0$. The curves shown correspond to, from top to bottom, $M = 2 \times 10^{14}$, $10^{14}$, $5 \times 10^{13}$, $2 \times 10^{13}$, $10^{13}$, $5 \times 10^{12}$, $2 \times 10^{12}$, and $10^{12} h^{-1} M_\odot$, respectively.

Fig. 2 shows the leading and leading + subleading contributions to the scale-dependent bias for a range of masses from $10^{12}$ to $2 \times 10^{14} h^{-1} M_\odot$. Clearly, the typical scale at which the subleading correction becomes important does not depend sensitively on the mass. Generally, the correction is more important at lower masses, specifically around $M_*$ where the Lagrangian bias $b_{100}$ vanishes. For vanishing $b_{100}$, the leading term vanishes, whereas the subleading term does not disappear entirely due to the nonvanishing derivative of $\ln \sigma_X^2$ with respect to $\ln \sigma_*^2$. Since $\sigma_{X,*}^2$ scales more weakly with $R_*$ than $\sigma_*^2$, this derivative is less than one leading to a suppression of the scale-dependent bias when including the subleading correction.

In principle, a general tracer could lead to very different numerical results, i.e. much larger or smaller subleading corrections. However, given the accuracy of universal mass functions of 10%–20% at least for dark matter halos, we expect the magnitude of the corrections as shown in Fig. 2 to be typical.

## VI. RELATION TO PREVIOUS RESULTS

We now make the connection to previous results on the scale-dependent bias for general non-Gaussianity beyond the squeezed limit. In Desjacques *et al.* [13], the scale-dependent bias was derived by applying a conditional mass function approach to the Press-Schechter (PS) mass function [22]. The non-Gaussianity was taken into account by applying an Edgeworth expansion to the Gaussian PDF of

the density field. As shown in [13], the scale-dependent bias defined through Eq. (70) is in this case given by

$$\Delta b_{\rm PS}(k) = \left[b_1 \delta_c + 2\frac{\partial \ln \mathcal{F}_*^{(3)}(k)}{\partial \ln \sigma_*^2}\right] 2\mathcal{F}_*^{(3)}(k)\mathcal{M}_L^{-1}(k). \quad (86)$$

Here, a subscript $*$ denotes filtering with a tophat of radius $R_*$, the Lagrangian radius corresponding to the mass scale of the tracer as in Sec. V B (this was denoted as $R_s$ in [13]). Note the derivative with respect to $\ln \sigma_*^2$ rather than $\ln \sigma_*$ as written in [13]. Here, the function $\mathcal{F}_*^{(3)}(k)$ is given by an integral over the bispectrum,

$$\mathcal{F}_*^{(3)}(k) = \frac{1}{4\sigma_*^2 P_\phi(k)} \int \frac{d^3 k_1}{(2\pi)^3} \mathcal{M}_*(k_1)\mathcal{M}_*(|\mathbf{k} + \mathbf{k}_1|)$$
$$\times B_\phi(\mathbf{k}, \mathbf{k}_1, -\mathbf{k} - \mathbf{k}_1),$$

where $\mathcal{M}_*(k) = \mathcal{M}(k)\tilde{W}_* k$. The function $\mathcal{F}_*^{(3)}$ can then be directly related to the squeezed three-point function

$$\langle \delta_L(1)\delta_*^2(2)\rangle = 4\sigma_*^2 \int \frac{d^3 k}{(2\pi)^3} \mathcal{M}_L(k)P_\phi(k)\mathcal{F}_*^{(3)}(k)e^{i\mathbf{k}\mathbf{r}}$$
$$= \sum_\alpha A_\alpha \{2\xi_{\alpha L}^{(1)}(r)\sigma_{\alpha *}^{2(23)}$$
$$+ \xi_{\nabla^2 \alpha L}^{(1)}(r)\sigma_{X\alpha *}^{2(23)} + 2 \text{ perm}\},$$

again up to order $(k/k_*)^4$, see for example the derivation leading up to Eq. (A7). In Fourier space, this relation becomes

$$2\mathcal{F}_*^{(3)}(k) = \frac{1}{P_\phi(k)}\sum_\alpha A_\alpha \left\{F_\alpha^{(1)}(k)\frac{\sigma_{\alpha *}^{2(23)}}{\sigma_*^2} + \frac{1}{2}k^2 F_\alpha^{(1)}(k)\frac{\sigma_{X\alpha *}^{2(23)}}{\sigma_*^2}\right.$$
$$\left.+ \{(123) \to (231)\} + \{(123) \to (312)\}\right\}. \quad (87)$$

Equation (86) then becomes

$$\Delta b_{\rm PS}(k) = \frac{1}{P_\phi(k)\mathcal{M}_L(k)}\sum_\alpha A_\alpha \left\{F_\alpha^{(1)}(k)\left[b_1\delta_c + 2\frac{\partial}{\partial \ln \sigma_*^2}\right]\right.$$
$$\times \frac{\sigma_{\alpha *}^{2(23)}}{\sigma_*^2} + \frac{1}{2}k^2 F_\alpha^{(1)}(k)\left[b_1\delta_c + 2\frac{\partial}{\partial \ln \sigma_*^2}\right]$$
$$\left.\times \frac{\sigma_{X\alpha *}^{2(23)}}{\sigma_*^2} + 2 \text{ perm}\right\}$$
$$= \sum_\alpha A_\alpha \left\{\mathcal{S}_\alpha^{(1)}(k)\left[b_1\delta_c + 2\frac{\partial}{\partial \ln \sigma_*^2}\right]\frac{\sigma_{\alpha *}^{2(23)}}{\sigma_*^2}\right.$$
$$+ \frac{1}{2}k^2 \mathcal{S}_\alpha^{(1)}(k)\left[b_1\delta_c + 2\frac{\partial}{\partial \ln \sigma_*^2}\right]$$
$$\left.\times \frac{\sigma_{X\alpha *}^{2(23)}}{\sigma_*^2} + 2 \text{ perm}\right\}, \quad (88)$$

using the definitions after Eq. (70). Comparing with Eq. (70), we can now read off the bias parameters

$$b_{010}^{(\alpha, 23)} = \left[b_1\delta_c + 2\frac{\partial}{\partial \ln \sigma_*^2}\right]\frac{\sigma_{\alpha *}^{2(23)}}{\sigma_*^2}$$
$$= \left[b_1\delta_c + 2\left(\frac{\partial \ln \sigma_{\alpha *}^{2(23)}}{\partial \ln \sigma_*^2} - 1\right)\right]\frac{\sigma_{\alpha *}^{2(23)}}{\sigma_*^2}, \quad (89)$$

$$b_{001}^{(\alpha, 23)} = \left[b_1\delta_c + 2\left(\frac{\partial \ln \sigma_{X\alpha *}^{2(23)}}{\partial \ln \sigma_*^2} - 1\right)\right]\frac{\sigma_{X\alpha *}^{2(23)}}{\sigma_*^2}. \quad (90)$$

We see that both the leading and subleading bias parameters derived from the conditional PS mass function agree with those expected from a general universal mass function [Eqs. (82) and (83), with $b_{010}^{\rm loc} = b_{100}\delta_c$]. Fundamentally, this is a consequence of the fact that in the Press-Schechter approach, as in general for universal mass functions, there is only one scale $R_*$ that enters the description of tracer statistics. We thus expect this result to hold at higher order in $k/k_*$ as well.

Since Eq. (86) does not involve a perturbative expansion in the ratio of wave numbers $k/k_*$, it can serve as a useful guide as to where this expansion breaks down. Fig. 3 shows the residuals when including the leading and subleading terms from Eq. (76). We see a residual which scales as $k^3$ for small $k$. When including the term from the last line of
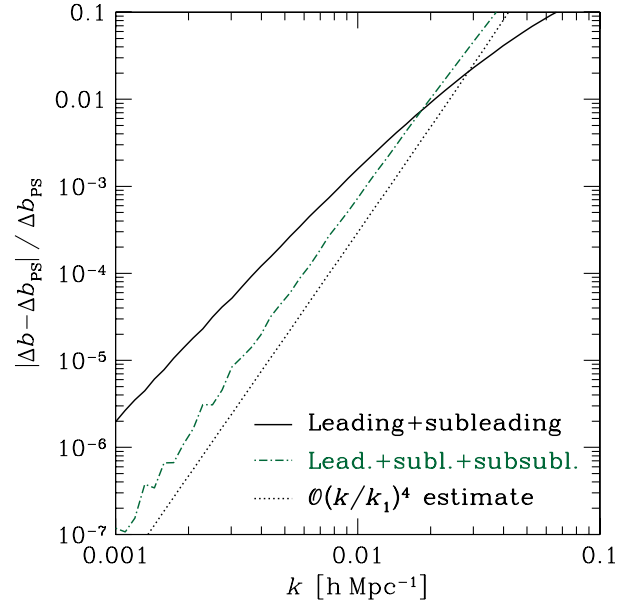


FIG. 3 (color online). Fractional difference between the contributions in Eq. (76), evaluated for a universal mass function, and the result for the conditional PS mass function Eq. 86 from [13] for local primordial non-Gaussianity. The black line solid line shows the residuals when including the leading and subleading (order $(k/k_*)^2$) contributions, while the green dash-dotted line also includes the order $(k/k_*)^3$ contribution from the last line of Eq. (76). The dotted line shows a rough estimate of the order $(k/k_*)^4$ correction (see text). We have again assumed $M = 2 \times 10^{13} h^{-1} M_\odot$ and $z = 0$, although the results are essentially independent of the mass.

Eq. (76), we see that the residuals become even smaller and scale as $k^4$ as expected. Some numerical artifacts are visible when the residuals become of order $10^{-6}$ or smaller. These are due to the numerical derivative performed when evaluating Eq. (86). We can in fact perform a rough estimate of the expected correction order $(k/k_*)^4$ in Eq. (76), via

$$\Delta b_{\text{NNLO est}} = f_{\text{NL}} b_{100} \delta_c \frac{\sigma_{-4,*}^2}{\sigma_*^2} k^4 \mathcal{M}_L^{-1}(k)$$

$$\sigma_{-4,*}^2 = \frac{1}{2\pi^2} \int_{k_{\text{min}}}^\infty k^2 dk k^{-4} P_m(k) \tilde{W}_L^2(k).$$

(91)

Note that due to the logarithmic divergence we need to introduce a low-$k$ cutoff in $\sigma_{-4,*}^2$. This is likely to be an artefact of the universal mass function prescription, where the leading effect of a change in the small-scale power spectrum shape on the tracer density is given by this formally divergent spectral moment. In reality, tracers will have a finite response to such a change. Here we choose $k_{\text{min}} = 0.01h \text{ Mpc}^{-1}$, corresponding to the turn-over in $P_m(k)$. Equation (91) is in any case only to be seen as a very rough estimate. This contribution is shown as dotted line in Fig. 3, making clear that the residual, after taking into account all terms in Eq. (76), indeed scales as $k^4$ on large scales.

Figure 3 shows that for the local model, the order $(k/k_*)^3$ correction becomes comparable to the lower order corrections at $k \sim 0.02h \text{ Mpc}^{-1}$, signaling a breakdown of the perturbative expansion there, even though the fractional deviation from the full result Eq. (86) when including terms up to order $(k/k_*)^2$ remain at 10% or less all the way to $k \sim 0.1h \text{ Mpc}^{-1}$.

In summary, the conditional PS mass function results derived in [13] are consistent for tracers following a universal mass function, in the sense that they match the results from the general renormalization approach when restricted to universal mass functions. Note that this holds once the polynomials in $\delta_c/\sigma_*$ are replaced with bias parameters, as described in [13]. In this context, the key advantage of the PS approach is that it sums over all powers of $k/k_*$, without relying on a perturbative expansion in this parameter.

On the other hand, realistic tracers will not simply depend on the variance of the density field on a single scale, thus breaking the relation between the bias parameters $b_{001}$, $b_{010}$, and $b_{001}$. Furthermore, there are other scale-dependent biases which contribute at the same order as the subleading correction $b_{001}$. We will turn to this issue next.

## VII. SCALE-DEPENDENT BIAS BEYOND THE LARGE-SCALE LIMIT

We have seen that beyond the squeezed limit, there is a subleading correction to the scale-dependent bias from primordial non-Gaussianity that scales as $k^2$ relative to

the leading term. In addition to this correction however, we expect two additional contributions that are leading order in $f_{\text{NL}}$ (i.e., in the primordial bispectrum), and have the same scaling with $k$.

First, as shown in [15,23], nonlocality in the formation of tracers generically induces a dependence on the curvature of the density field, leading to a contribution of

$$\xi_h(r) \supset b_{\nabla^2\delta} b_{010} \langle \nabla^2 \delta_L(1) y_*(2) \rangle. \quad (92)$$

This contribution will also serve to absorb the residual $R_L$ dependence present in $\xi_{\alpha L}^{(i)}(r)$, $\xi_{\nabla^2\alpha L}^{(i)}(r)$ in a similar way as discussed in [15]. In Fourier space, Eq. (92) corresponds to a contribution to the scale-dependent bias of the form

$$\frac{P_h(k)}{P_m(k)} \supset 2 b_{\nabla^2\delta} k^2 \Delta b_{\text{lead}}(k), \quad (93)$$

where $\Delta b_{\text{lead}}$ is the leading contribution to the non-Gaussian scale-dependent bias from Eq. (71). If $L_\delta$ is the scale of nonlocality of the tracer (in terms of its dependence on the matter density), then $b_{\nabla^2\delta} \sim L_\delta^2$, so that this additional contribution scales as $(kL_\delta)^2$ relative to the leading term.

Throughout the discussion of the non-Gaussian case in [15] and here, we have assumed that the tracer density is a purely local function of $y_*$, the parameter which quantifies the amplitude of small-scale fluctuations (in this paper, we have also introduced a local dependence on $\mu_* = dy_*/d\ln R_*$). In general, however, one also expects that tracers depend on the amplitude of small-scale fluctuations in some finite region of size $L_y$. Then, in straightforward analogy with the density case, we also need to allow for a bias $b_{\nabla^2 y}$ with respect to $\nabla^2 y_*$, where $b_{\nabla^2 y} \sim L_y^2$. That is, strictly speaking we need to generalize each of the rescalings in Eq. (68) to be spatially dependent, $\varepsilon \to \varepsilon \mathbf{x}^2$. Schematically, this leads to a contribution to the tracer correlation of

$$\xi_h(r) \supset b_{100} b_{\nabla^2 y} \langle \delta_L(1) \nabla^2 y_*(2) \rangle, \quad (94)$$

which, in Fourier space, becomes

$$\frac{P_h(k)}{P_m(k)} \supset 2 b_{100} b_{\nabla^2 y} k^2 \frac{\Delta b_{\text{lead}}(k)}{b_{010}}. \quad (95)$$

This contribution thus scales as $(kL_y)^2$ relative to the leading term.

Typically, one might expect $L_\delta \sim L_y \sim R_*$, where $R_*$ is the Lagrangian radius of the region that collapses to form the tracer. On the other hand, the length scale we found for the subleading terms in the squeezed limit is $1/k_* \sim 50h^{-1} \text{ Mpc}$, suggesting that this correction is somewhat more important than the other contributions described in this section for typical tracers for which $R_* = 1 - 10h^{-1} \text{ Mpc}$. However, the value of $k_*$ depends on the specific type of non-Gaussianity considered, and in general all three subleading contributions to the scale-dependent

halo bias can be comparable in magnitude. Thus, if one of them is included (even implicitly as for example in the conditional PS mass function result), then all of them should be included for consistency, unless one can show that the subleading contribution dominates over the other contributions described in this section.

## VIII. CONCLUSIONS

We have derived the subleading contributions to the scale-dependent bias $\Delta b(k)$ of large-scale structure tracers for a general separable primordial bispectrum. The leading contribution is given by the scaling of the bispectrum in the squeezed limit, $\lim_{k \to 0} B_\phi(k, k_s, |\mathbf{k}_s + \mathbf{k}|)$, and the scale-dependence of the subleading contribution is suppressed by a factor of $k^2$ relative to this term. This subleading contribution is important to quantify, since it tells us at which $k$ the usual squeezed-limit result ceases to be accurate. For local non-Gaussianity and tracers following a universal mass function, we found that this happens at $k \sim 0.02 h$ Mpc$^{-1}$, although the first two leading terms provide an excellent approximation to the result from the conditional PS mass function up to $k \lesssim 0.1 h$ Mpc$^{-1}$. Our approach is independent of any assumptions on the tracers apart from a finite scale of nonlocality.

Throughout, our results have been phrased in terms of the three-point function of the primordial perturbations, which is the standard result of computations performed for particular inflationary models, thus allowing for a direct application of the results of this paper to models of inflation. In contrast, several previous papers [11,14,19] have employed a fictitious Gaussian field mapped to the physical field via a quadratic kernel. In this latter approach, the effect on large-scale structure tracers is mediated by modulated spectral moments of the density field, e.g. $\sigma_*^2|_\phi$. This approach is complicated by the fact that the kernel is not uniquely determined by the bispectrum, so that additional constraints need to be imposed [14]. However, the derivation in this paper can also be applied to the kernel approach in a straightforward way. Note that while the bispectrum specifies the kernel uniquely in the squeezed limit [11] (for nondivergent kernels), this is no longer the case when including subleading terms. Thus, different kernels which yield the same bispectrum are expected to make different predictions for the subleading contribution to the scale-dependent bias. We leave this issue for future work.

In addition to the subleading contribution in the squeezed limit, we have pointed out two other contributions that will contribute at linear order in $f_{NL}$ and with the same tree-level $k$ dependence (Sec. VII). These contributions should be included (or at least carefully considered) when putting constraints on non-Gaussianity using large-scale structure statistics on intermediate and small scales, i.e. for $k \gtrsim 0.05 h$ Mpc$^{-1}$.

It is straightforward to extend the squeezed limit expansion presented here to higher order in $k$. In that case, one needs to add a dependence of the tracer density on another property of the density field further quantifying the sensitivity to the amplitude of small-scale fluctuations as a function of scale. One possible choice would be $d^2 y_*/d(\ln R_*)^2$.

Another straightforward extension is the inclusion of higher primordial $N$-point functions. For example, in the presence of a primordial four-point function both the linear and quadratic bias become scale-dependent [13,24], and one needs to take into account the dependence of the tracer density on the skewness of the density field. No conceptually new issue arises, and the calculation will closely follow the one presented here.

Finally, we have shown how the coefficient of both leading and subleading terms in the scale-dependent bias for a general bispectrum can be derived for tracers identified in N-body simulations, by running simulations with modified initial conditions [Eq. (68)]. This will allow for a precise test of the accuracy of the universal mass function prediction for dark matter halos in the context of non-Gaussian halo bias. We leave this for future work.

## ACKNOWLEDGMENTS

## APPENDIX: DERIVATION OF EQS. (24) AND (37)

In this Appendix we derive the next-to-leading order squeezed-limit expressions Eqs. (24) and (37). As described in Sec. III, we assume that the bispectrum $B_\phi$ is given in separable form,

$$
\begin{aligned}
&B_\phi(k_1, k_2, k_3) \\
&= \sum_\alpha [F_\alpha^{(1)}(k_1) F_\alpha^{(2)}(k_2) F_\alpha^{(3)}(k_3) + F_\alpha^{(1)}(k_1) F_\alpha^{(3)}(k_3) F_\alpha^{(2)}(k_2) \\
&\quad + F_\alpha^{(2)}(k_1) F_\alpha^{(3)}(k_2) F_\alpha^{(1)}(k_3) + F_\alpha^{(2)}(k_1) F_\alpha^{(1)}(k_3) F_\alpha^{(3)}(k_2) \\
&\quad + F_\alpha^{(3)}(k_1) F_\alpha^{(1)}(k_2) F_\alpha^{(2)}(k_3) + F_\alpha^{(3)}(k_1) F_\alpha^{(3)}(k_3) F_\alpha^{(1)}(k_2)],
\end{aligned}
$$
(A1)

where the six permutations guarantee the symmetry of $B_\phi$ in its arguments. This leads to Eq. (19). We now expand the $k_1$ integrand in powers of $k/k_1$ up to second order. For notational simplicity, we only consider the contribution from a single term $\alpha$ and two permutations $(2) \leftrightarrow (3)$, and drop the subscript $\alpha$ for the moment. This yields

$$\int \frac{d^3 k_1}{(2\pi)^3} [\tilde{F}^{(2)}(\mathbf{k}_1)\tilde{F}^{(3)}(-\mathbf{k}_1 - \mathbf{k}) + (2) \leftrightarrow (3)]$$

$$= \int \frac{d^3 k_1}{(2\pi)^3} \left[ \tilde{F}^{(2)}(\mathbf{k}_1) \left( 1 - k^i \partial_i + \frac{1}{2} k^i k^j \partial_i \partial_j \right) \right.$$

$$\left. \times \tilde{F}^{(3)}(-\mathbf{k}_1) + (2) \leftrightarrow (3) \right]$$

$$= \int \frac{d^3 k_1}{(2\pi)^3} \left[ \tilde{F}^{(2)}(\mathbf{k}_1)\tilde{F}^{(3)}(-\mathbf{k}_1) - \tilde{F}^{(3)}(-\mathbf{k}_1)k^i \partial_i \tilde{F}^{(2)}(\mathbf{k}_1) \right.$$

$$\left. + \frac{1}{2} \tilde{F}^{(3)}(-\mathbf{k}_1)k^i k^j \partial_i \partial_j \tilde{F}^{(2)}(\mathbf{k}_1) + (2) \leftrightarrow (3) \right], \quad (A2)$$

where all derivatives are with respect to $\mathbf{k}_1$. We now perform an integration by parts for the second and last terms. The former term (linear in $\mathbf{k}$) cancels with its permutation $(2) \leftrightarrow (3)$, yielding

$$\int \frac{d^3 k_1}{(2\pi)^3} \left[ 2\tilde{F}^{(2)}(\mathbf{k}_1)\tilde{F}^{(3)}(-\mathbf{k}_1) \right.$$

$$+ \frac{1}{2} \tilde{F}^{(3)}(-\mathbf{k}_1)k^i k^j \partial_i \partial_j \tilde{F}^{(2)}(\mathbf{k}_1)$$

$$\left. + \frac{1}{2} \tilde{F}^{(2)}(\mathbf{k}_1)k^i k^j \partial_i \partial_j \tilde{F}^{(3)}(-\mathbf{k}_1) \right]. \quad (A3)$$

Note that we have not used that $\tilde{F}^{(i)}(\mathbf{k}_1) = \tilde{F}^{(i)}(k_1)$ so far. We now use this fact however to obtain, defining $\mu = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}_1$,

$$k^i k^j \partial_i \partial_j F(k_1) = \frac{k^2}{k_1^2} [(1 - \mu^2)k_1 F'(k_1) + \mu^2 k_1^2 F''(k_1)],$$

$$(A4)$$

where we have denoted derivatives with respect to $k_1$ with primes. We thus obtain, up to terms of order $k^4/k_1^4$ (cubic terms drop out just like the linear terms did)

$$\langle \delta_L(1)\delta_L^2(2) \rangle$$

$$= \sum_\alpha \left\{ \left[ \int \frac{d^3 k}{(2\pi)^3} \tilde{F}_\alpha^{(1)}(k)e^{i\mathbf{k} \cdot \mathbf{r}} \int \frac{d^3 k_1}{(2\pi)^3} \left[ 2\tilde{F}_\alpha^{(2)}(k_1)\tilde{F}_\alpha^{(3)}(k_1) \right. \right. \right.$$

$$+ \frac{1}{2} \frac{k^2}{k_1^2} \tilde{F}_\alpha^{(2)}(k_1)[(1 - \mu^2)k_1 \tilde{F}_\alpha^{\prime(3)}(k_1) + \mu^2 k_1^2 \tilde{F}_\alpha^{\prime\prime(3)}(k_1)]$$

$$\left. \left. + \frac{1}{2} \frac{k^2}{k_1^2} \tilde{F}_\alpha^{(3)}(k_1)[(1 - \mu^2)k_1 \tilde{F}_\alpha^{\prime(2)}(k_1) + \mu^2 k_1^2 \tilde{F}_\alpha^{\prime\prime(2)}(k_1)] \right] \right.$$

$$\left. + \{(123) \rightarrow (231)\} + \{(123) \rightarrow (312)\} \right\}. \quad (A5)$$

This expression is now in the desired separable form, and the $\mu$ integral becomes trivial. We now introduce some notation

$$\xi_{\alpha L}^{(i)}(r) \equiv \int \frac{d^3 k}{(2\pi)^3} \mathcal{M}_L(k)F_\alpha^{(i)}(k)e^{i\mathbf{k} \cdot \mathbf{r}}$$

$$\xi_{\nabla^2 \alpha L}^{(i)}(r) \equiv \int \frac{d^3 k}{(2\pi)^3} \mathcal{M}_L(k)k^2 F_\alpha^{(i)}(k)e^{i\mathbf{k} \cdot \mathbf{r}}$$

$$\sigma_{\alpha L}^{2(ij)} \equiv \frac{1}{2\pi^2} \int_0^\infty dk_1 k_1^2 \mathcal{M}_L^2(k_1)F_\alpha^{(i)}(k_1)F_\alpha^{(j)}(k_1)$$

$$\sigma_{X\alpha L}^{2(ij)} \equiv \frac{1}{2\pi^2} \int_0^\infty dk_1 \frac{1}{2} \left\{ \tilde{F}_\alpha^{(i)}(k_1) \left[ \frac{2}{3} k_1 \tilde{F}_\alpha^{\prime(j)}(k_1) \right. \right.$$

$$\left. \left. + \frac{1}{3} k_1^2 \tilde{F}_\alpha^{\prime\prime(j)}(k_1) \right] + (i) \leftrightarrow (j) \right\}$$

$$= \frac{1}{12\pi^2} \int_0^\infty dk_1 \{ \tilde{F}_\alpha^{(i)}(k_1)[2k_1 \tilde{F}_\alpha^{\prime(j)}(k_1)$$

$$+ k_1^2 \tilde{F}_\alpha^{\prime\prime(j)}(k_1)] + (i) \leftrightarrow (j) \}. \quad (A6)$$

Note that for all spectral moments, $\sigma^{2(ij)} = \sigma^{2(ji)}$. This allows us to write Eq. (24) in the compact form

$$\langle \delta_L(1)\delta_L^2(2) \rangle = \sum_\alpha \{ 2\xi_{\alpha L}^{(1)}(r)\sigma_{\alpha L}^{2(23)}$$

$$+ \xi_{\nabla^2 \alpha L}^{(1)}(r)\sigma_{X\alpha L}^{2(23)} + 2 \text{ perm} \}, \quad (A7)$$

where the two permutations stand for the cyclic permutations $123 \rightarrow 231, 312$.

We now turn to Eq. (37). This correlator is given by

$$\langle \delta_L(1)y_*(2) \rangle = \frac{1}{2\sigma_s^2} \langle \delta_L(1)\delta_s^2(2) \rangle$$

$$= \frac{1}{2\sigma_s^2} \int \frac{d^3 k}{(2\pi)^3} \mathcal{M}_L(k)e^{i\mathbf{k} \cdot \mathbf{r}} \int \frac{d^3 k_1}{(2\pi)^3} \mathcal{M}_s(|\mathbf{k}_1|)$$

$$\times \mathcal{M}_s(|\mathbf{k}_1 + \mathbf{k}|)B_\phi(|\mathbf{k}|, |\mathbf{k}_1|, |\mathbf{k}_1 + \mathbf{k}|),$$

$$(A8)$$

where $\mathcal{M}_s(k) = \tilde{W}_s(k)\mathcal{M}(k)$. This expression is very similar to Eq. (19), the only difference being that the filter functions under the $k_1$ integral involve $\tilde{W}_s$ rather than $\tilde{W}_L$. Equation (A7) then straightforwardly translates to

$$\langle \delta_L(1)y_*(2) \rangle = \sum_\alpha A_\alpha \left\{ \xi_{\alpha L}^{(1)}(r) \frac{\sigma_{\alpha s}^{2(23)}}{\sigma_s^2} \right.$$

$$\left. + \frac{1}{2} \xi_{\nabla^2 \alpha L}^{(1)}(r) \frac{\sigma_{X\alpha s}^{2(23)}}{\sigma_s^2} + 2 \text{ perm} \right\}, \quad (A9)$$

where

$$\sigma_{\alpha s}^{2(ij)} \equiv \frac{1}{2\pi^2} \int_0^\infty dk_1 k_1^2 \mathcal{M}_s^2(k_1)F_\alpha^{(i)}(k_1)F_\alpha^{(j)}(k_1)$$

$$\sigma_{X\alpha s}^{2(ij)} \equiv \frac{1}{12\pi^2} \int_0^\infty dk_1 \{ \mathcal{M}_s(k_1)F_\alpha^{(i)}(k_1)[2k_1(\mathcal{M}_s F_\alpha^{(j)})'_{k_1}$$

$$+ k_1^2 (\mathcal{M}_s F_\alpha^{(j)})''_{k_1}] + (i) \leftrightarrow (j) \}. \quad (A10)$$

[1] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, Phys. Rep. **402**, 103 (2004).

[2] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday *et al.* (Planck Collaboration), arXiv:1303.5084.

[3] S. Matarrese, F. Lucchin, and S. A. Bonometto, Astrophys. J. Lett. **310**, L21 (1986).

[4] P. Coles, L. Moscardini, F. Lucchin, S. Matarrese, and A. Messina, Mon. Not. R. Astron. Soc. **264**, 749 (1993).

[5] X. Luo and D. N. Schramm, Astrophys. J. **408**, 33 (1993).

[6] L. Verde and A. F. Heavens, Astrophys. J. **553**, 14 (2001).

[7] E. Sefusatti and E. Komatsu, Phys. Rev. D **76**, 083004 (2007).

[8] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, J. Cosmol. Astropart. Phys. 08 (2008) 031.

[9] D. Jeong and E. Komatsu, Astrophys. J. **703**, 1230 (2009).

[10] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, Phys. Rev. D **77**, 123514 (2008).

[11] F. Schmidt and M. Kamionkowski, Phys. Rev. D **82**, 103002 (2010).

[12] S. Matarrese and L. Verde, Astrophys. J. Lett. **677**, L77 (2008).

[13] V. Desjacques, D. Jeong, and F. Schmidt, Phys. Rev. D **84**, 063512 (2011).

[14] R. Scoccimarro, L. Hui, M. Manera, and K. C. Chan, Phys. Rev. D **85**, 083002 (2012).

[15] F. Schmidt, D. Jeong, and V. Desjacques, arXiv:1212.0868.

[16] P. McDonald, Phys. Rev. D **74**, 103512 (2006).

[17] T. Matsubara, Phys. Rev. D **83**, 083518 (2011).

[18] P. McDonald, Phys. Rev. D **78**, 123519 (2008).

[19] T. Giannantonio and C. Porciani, Phys. Rev. D **81**, 063530 (2010).

[20] D. Tseliakhovich, C. Hirata, and A. Slosar, Phys. Rev. D **82**, 043531 (2010).

[21] D. Baumann, S. Ferraro, D. Green, and K. M. Smith, J. Cosmol. Astropart. Phys. 05 (2013) 001.

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

[23] P. McDonald and A. Roy, J. Cosmol. Astropart. Phys. 08 (2009) 020.

[24] K. M. Smith, S. Ferraro, and M. LoVerde, J. Cosmol. Astropart. Phys. 03 (2012) 032.

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