

Bias loop corrections to the galaxy bispectrum

Alexander Eggemeier,^{1,2,*} Román Scoccimarro,³ and Robert E. Smith¹

¹*Astronomy Centre, School of Mathematical and Physical Sciences,
University of Sussex, Brighton BN1 9QH, United Kingdom*

²*Institute for Computational Cosmology, Department of Physics, Durham University,
South Road, Durham DH1 3LE, United Kingdom*

³*Center for Cosmology and Particle Physics, Department of Physics, New York University,
New York 10003, New York, USA*



(Received 12 December 2018; published 13 June 2019)

The combination of the power spectrum and bispectrum is a powerful way of breaking degeneracies between galaxy bias and cosmological parameters, enabling us to maximize the constraining power from galaxy surveys. Recent cosmological constraints treat the power spectrum and bispectrum on an uneven footing: they include one-loop bias corrections for the power spectrum but not the bispectrum. To bridge this gap, we develop the galaxy bias description up to fourth order in perturbation theory, conveniently expressed through a basis of Galilean invariants that clearly split contributions that are local and nonlocal in the second derivatives of the linear gravitational potential. In addition, we consider relevant contributions from short-range nonlocality (higher-derivative terms), stress-tensor corrections, and stochasticity. To sidestep the usual renormalization of bias parameters that complicates predictions beyond leading order, we recast the bias expansion in terms of multipoint propagators, which take a simple form in our split basis with loop corrections depending only on bias parameters corresponding to nonlocal operators. We show how to take advantage of Galilean invariance to compute the time evolution of bias and present results for the fourth-order parameters for the first time. We also discuss the possibilities of reducing the bias parameter space by using the evolution of bias and exploiting degeneracies between various bias contributions in the large-scale limit. Our baseline model allows us to verify these simplifications for any application to large-scale structure data sets.

DOI: [10.1103/PhysRevD.99.123514](https://doi.org/10.1103/PhysRevD.99.123514)

I. INTRODUCTION

Any mismodeling of galaxy bias—the relation between galaxies (or any other luminous tracer) and the underlying matter distribution [1–6]—risks an incorrect recovery of cosmological parameters from large-scale structure (LSS) surveys. Fortunately, although the formation of galaxies involves highly nonlinear, small-scale processes, recent developments [7–14] have shown that a perturbative expansion provides a robust treatment of galaxy bias on sufficiently large scales. This comes at the price of a set of unknown bias parameters, which, once marginalized over, degrade the statistical power for constraining cosmological models. It is therefore imperative to combine traditional LSS two-point statistics with higher-order statistics, such as the bispectrum, which allows us to break the degeneracies that exist between cosmological parameters and galaxy bias [15–21]. For this program to succeed we require consistent models of the galaxy power spectrum and bispectrum at leading (“tree-level”) and next-to-leading (“one-loop”) order, which should improve the regime of validity and

robustness of the results [9,22–24]. This is particularly important since at present the analysis of the bispectrum in galaxy surveys [25,26] is done in a manner that is inconsistent with the treatment of the power spectrum, in that bias is treated at one-loop order for the power spectrum but tree level for the bispectrum. The main goal of this study is to present such a unified model that includes all relevant effects in real space, extending the results of [12].

What are the essential elements? First, we need to consider all contributions from the general bias expansion up to fourth order in perturbation theory [6,12,13]. These are generated by the gravitational evolution of the dark matter field and include the common linear and nonlinear bias parameters [1,27,28], as well as the second-order nonlocal (or tidal field) bias [8–10,29]. Second, it has been argued in [30,31] that so-called higher-derivative terms, which have been known for a long time [3] and can be understood as deriving from a dependence of galaxy formation on the spatial distribution of matter [8], can be as important or even dominate over the general bias terms in certain circumstances. Third, on scales of the weakly nonlinear regime stress-tensor corrections to the evolution of dark matter [32–37], it might potentially

*alexander.eggemeier@durham.ac.uk

become relevant. And finally, the impact of very small scales, in the absence of primordial non-Gaussianity largely uncorrelated with the previously mentioned large-scale effects, leads to an additional stochastic bias [38,39].

However, a challenge that afflicts all theoretical predictions of galaxy clustering beyond leading order is a mismatch between the bias parameters from the perturbative expansion, and those an observer would define through the measurement of correlation functions. This was first pointed out in [7], which showed that appropriate renormalizations of the perturbative bias parameters restore agreement with the measurements. These renormalizations must be done on a statistic-by-statistic and parameter-by-parameter basis, and while tractable for the power spectrum [7,8], this becomes increasingly complicated for higher-order statistics [12]. As a way of circumventing this procedure altogether, we advocate the use of multipoint propagators.

The multipoint propagator formalism was originally introduced in the context of renormalized perturbation theory (RPT) and its generalizations [40–45] to obtain an accurate description of the dark matter field in the quasilinear regime. Multipoint propagators have also been extended to include redshift space distortions and galaxy bias in [46–48], but without the clear connection to the renormalization issue that we aim to establish in this paper. In fact, we will demonstrate that the multipoint propagators correspond to the bias parameters that are commonly identified through cross-correlations of galaxy and matter fields [22,49]. Thus, they are observable and have a well-defined physical meaning, based on which we argue that they provide the most natural approach toward galaxy bias. As in RPT they further act as building blocks for the general N -point correlation function, and determination of the first three propagators already fixes the power spectrum and bispectrum at the one-loop level, simplifying the computations significantly.

Modeling higher-order statistics to the same precision as their lower-order counterparts brings about a quickly inflating number of terms from the various bias contributions. In order to reduce the growing parameter space, a further goal of this article is to examine (1) degeneracies between the bias terms, and (2) relations between bias parameters arising from evolution, which are studied here for the first time at fourth order.

The paper is organized as follows. In Sec. II we present the central idea of the multipoint propagator expansion based on the simple and widely known model of local galaxy bias. A complete basis for the general bias and higher-derivative terms is provided in Sec. III. Following [9] this basis makes intuitive use of Galilean invariants of Lagrangian perturbation theory (LPT) potentials and is particularly well suited for the computation of the multipoint propagators that we perform in Sec. IV. We also determine their time evolution and present the relations between initial and evolved bias parameters. Finally, Sec. V

uses these results to compute the power spectrum and bispectrum, and demonstrates how we incorporate stress-tensor corrections and stochasticity. Section VI summarizes the final model components and gives our conclusions. Appendixes A and B give further details and include relations between our basis for galaxy bias and others in the literature. Appendix C shows how to take advantage of Galilean invariance for determining the evolution of galaxy bias, while Appendix D demonstrates that the corresponding results are unaffected by renormalization.

II. BASICS OF THE RENORMALIZED BIAS EXPANSION

We are interested in the statistical properties of the observed galaxy distribution. These are commonly quantified by a hierarchy of correlation functions of the density perturbations, which we write as $\delta_g(\mathbf{x}) = n_g(\mathbf{x})/\bar{n}_g - 1$, with $n_g(\mathbf{x})$ denoting the galaxy number density and $\bar{n}_g$ its average.¹ Model predictions of the correlation functions require a relation between the galaxy perturbations and matter fluctuations $\delta(\mathbf{x})$, which is usually written as some functional $\delta_g[\delta(\mathbf{x})]$ that is then Taylor expanded.

In order to illustrate the main idea pursued in this paper, we start with the simplest and most well-known bias relation, which considers δ_g to be a local function of δ [27,28] and can be written as a Taylor series around $\delta = 0$. Dropping all position arguments we have

$$\delta_g = \sum_n \frac{1}{n!} \left(\frac{\partial^n \delta_g}{\partial \delta^n} \right)_0 \delta^n \equiv \bar{b}_0 + \bar{b}_1 \delta + \frac{\bar{b}_2}{2} \delta^2 + \frac{\bar{b}_3}{3!} \delta^3 + \dots, \quad (1)$$

where the bias parameters are identified as $\bar{b}_n = (\partial^n \delta_g / \partial \delta^n)_0$ ($n > 0$) with $(\)_0$ denoting evaluation at $\delta = 0$ and $\bar{b}_0$ enforces $\langle \delta_g \rangle = 0$. However, the bias parameters so defined are not observables, as we measure correlators (expectation values or ensemble averages of fields), not quantities evaluated at $\delta = 0$. As we will detail below, this is the reason for difficulties in the computation of galaxy correlation functions that ultimately require us to redefine (or “renormalize” [7]) the bias parameters above. We will further show that a more natural bias expansion, in the sense that its coefficients are actually measurable, can be defined in terms of cross-correlations between the galaxy and matter fields. In the language of RPT and its generalizations [40,41] these cross-correlations correspond to the so-called *multipoint propagators*.

¹Note that while the subscript g here stands for galaxies, it could equally well denote any other tracer of the mass field, such as quasars, galaxy clusters, or the Ly- α forest.

A. Galaxy clustering statistics in the standard approach

We define the two- and three-point correlation functions of the galaxy perturbations in Fourier space—the power spectrum and bispectrum—as follows:

$$\langle \delta_g(\mathbf{k}_1) \delta_g(\mathbf{k}_2) \rangle \equiv (2\pi)^3 P_g(k_1) \delta_D(\mathbf{k}_{12}), \quad (2)$$

$$\langle \delta_g(\mathbf{k}_1) \delta_g(\mathbf{k}_2) \delta_g(\mathbf{k}_3) \rangle \equiv (2\pi)^3 B_g(k_1, k_2, k_3) \delta_D(\mathbf{k}_{123}), \quad (3)$$

where $\mathbf{k}_{1\dots n} \equiv \mathbf{k}_1 + \dots + \mathbf{k}_n$ and the appearance of the delta distribution is a manifestation of statistical homogeneity. Statistical isotropy further demands that the power spectrum only depends on the magnitude of the two wave vectors participating in the correlator, while the bispectrum is a function of three wave numbers k_1 , k_2 , and k_3 . Finally, to pass from configuration to Fourier space we have adopted the convention

$$\delta_g(\mathbf{x}) = \int_{\mathbf{k}} \exp(-i\mathbf{k} \cdot \mathbf{x}) \delta_g(\mathbf{k}), \quad (4)$$

using a shorthand notation for k -space integrals, i.e., $\int_{\mathbf{k}_1, \dots, \mathbf{k}_n} \equiv \int d^3k_1 / (2\pi)^3 \dots d^3k_n / (2\pi)^3$.

Analogous definitions hold for the dark matter field, and to begin with let us assume that the bias expansion is done in Lagrangian space, so that the matter fluctuations can be considered linear and Gaussian (we neglect any possible primordial non-Gaussianities for the remainder of this paper). In that case all statistical information is contained in the linear power spectrum, which we denote by $\langle \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) \rangle = (2\pi)^3 P_L(k_1) \delta_D(\mathbf{k}_{12})$, while the bispectrum and all other higher-order statistics vanish. Under this assumption, plugging Eq. (1) into Eq. (2) and using Wick's theorem lead to

$$P_g(k) = \bar{b}_1^2 P_L(k) + [\bar{b}_1 \bar{b}_3 \sigma^2] P_L(k) + \frac{\bar{b}_2^2}{2} \int_q P_L(|\mathbf{k} - \mathbf{q}|) P_L(q) + \dots, \quad (5)$$

where the dots denote contributions of two-loop and higher order, i.e., terms that involve five or more powers of δ . The variance $\sigma^2 \equiv \langle \delta^2(\mathbf{x}) \rangle$, which appears in the first term of the loop contribution in square brackets, is formally infinite or at least highly sensitive to the nonlinear regime, depending on the shape of the linear power spectrum.² This implies that the large-scale galaxy power spectrum would be heavily influenced by scales where our perturbative approach is not expected to hold. However, we notice that this term is also proportional to $P_L(k)$, such that if we

redefine the linear bias parameter to be $b_1 \equiv \bar{b}_1 + \bar{b}_3 \sigma^2 / 2$, we retain the form

$$P_g(k) = b_1^2 P_L(k) + \frac{\bar{b}_2^2}{2} \int_q P_L(|\mathbf{k} - \mathbf{q}|) P_L(q). \quad (6)$$

Clearly, as more bias loops are included, the expression for b_1 keeps changing, but the principle remains the same—the *observed* linear bias is defined as (the square root of) the coefficient in front of P_L .

A similar situation holds for the quadratic bias parameter b_2 . To see this let us now consider the bispectrum up to one-loop order (up to six powers of δ), under the same assumption as above. We then get from Eq. (3)

$$\begin{aligned} B_{g,123} = & \bar{b}_2 \bar{b}_1^2 P_1 P_2 + \left[\bar{b}_1 \bar{b}_2 \bar{b}_3 \sigma^2 + \frac{1}{2} \bar{b}_1^2 \bar{b}_4 \sigma^2 \right] P_1 P_2 \\ & + \frac{1}{2} \bar{b}_1 \bar{b}_2 \bar{b}_3 P_1 \int_q P_L(|\mathbf{k}_2 - \mathbf{q}|) P_L(q) \\ & + \bar{b}_2^3 \int_q P_L(|\mathbf{k}_1 - \mathbf{q}|) P_L(|\mathbf{k}_2 + \mathbf{q}|) P_L(q) + \text{cyc} \\ & + \dots, \end{aligned} \quad (7)$$

where $P_i \equiv P_L(k_i)$ and *cyc.* denotes cyclic permutations of each term over the three wave vectors. Again, the first term in square brackets corresponds to the renormalization of the linear bias seen in the power spectrum, but in addition now there is also a second term that can be absorbed by a renormalization of the *quadratic bias* $b_2 \equiv \bar{b}_2 + \bar{b}_4 \sigma^2 / 2$, so that we can write

$$\begin{aligned} B_{g,123} = & b_2 b_1^2 P_1 P_2 + \frac{1}{2} \bar{b}_1 \bar{b}_2 \bar{b}_3 P_1 \int_q P_L(|\mathbf{k}_2 - \mathbf{q}|) P_L(q) \\ & + \bar{b}_2^3 \int_q P_L(|\mathbf{k}_1 - \mathbf{q}|) P_L(|\mathbf{k}_2 + \mathbf{q}|) P_L(q) + \text{cyc} \end{aligned} \quad (8)$$

A number of questions arise from this procedure. (1) Is the b_2 renormalization that follows from the one-loop bispectrum consistent with the one that follows from the two-loop power spectrum? If so, is that true for all other bias parameters? (2) Is there a way to do all these renormalizations “automatically,” instead of calculating statistic by statistic and order by order? (3) Is there a simple connection between renormalizations of different N -point correlators? As we will see next, the answer to all these questions is “yes” [40,41,43,50].

B. A bias expansion based on observables

First, we notice that the terms that contribute to the renormalization of bias parameters must derive from factorizable loop corrections; i.e., they can be written as products of lower-order terms and new contributions (in the

²Imposing a high- k cutoff for the power spectrum would only mask the problem, as this would lead to all observables being dependent on an arbitrary choice of scale.

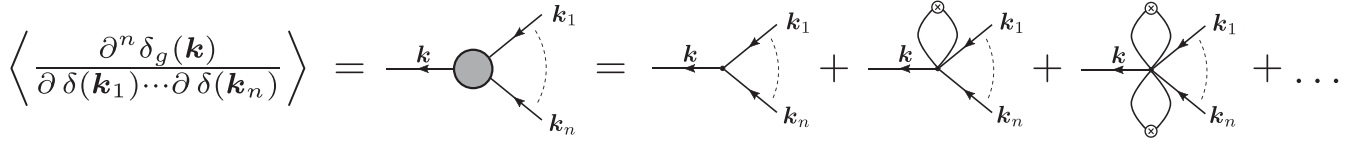


FIG. 1. The n th multipoint propagator is given by the sum of all reducible diagrams with n legs of incoming momentum $\mathbf{k}_1, \dots, \mathbf{k}_n$ (the external lines). The vertices correspond to the kernels associated with terms in the bias expansion; e.g., the vertices of the three diagrams on the right-hand side are given by $\bar{b}_n$, $\bar{b}_{n+2}$, and $\bar{b}_{n+4}$ for the expansion in Eq. (1). If terms nonlocal in the matter density enter the bias relation, these kernels acquire a scale dependence based on all incoming momenta. The crossed circles stand for (linear) power spectra.

example of Sec. II A these new contributions are functions of bias parameters and σ^2). In general, any loop correction can be portrayed as a diagram with a fixed number of external lines and a variable number of internal lines depending on the loop order. Those diagrams that lead to factorizable contributions are so-called *reducible*, meaning they can be decomposed into two or more connected diagrams by cutting one or more internal lines, while all other ones that do not share this property are classified as *irreducible*.

In order to automatically include renormalizations to arbitrary loop order, it is therefore desirable to construct the bias expansion in terms of the sum over all reducible diagrams with a given number of external lines. Such objects are known as multipoint propagators (see Fig. 1) and correspond to the expectation value of functional derivatives of the galaxy field with respect to the matter field: the derivatives produce the external lines, while the expectation value generates the loops [40].

An arbitrarily complicated loop diagram for any statistic can be decomposed easily and uniquely into multipoint propagators: the reducible subdiagrams are absorbed into the multipoint propagators while the irreducible subdiagrams are generated by connecting multipoint propagators among themselves. Since a given multipoint propagator appears in various statistics, this gives us a connection between renormalizations of different N -point correlation functions; e.g., the renormalization of b_2 that follows from the one-loop bispectrum is consistent with the one that follows from the two-loop power spectrum, because both are incorporated into the two-point propagator (i.e., they correspond to the *same* subdiagram in the power spectrum and bispectrum). As a result of this, $\bar{b}_2$ in Eq. (6) can be replaced by b_2 and similarly $\bar{b}_1 \bar{b}_2 \bar{b}_3 \rightarrow b_1 b_2 b_3$ and $\bar{b}_2^3 \rightarrow b_2^3$ in Eq. (8).

The observed bias parameters of order n that appear in correlators thus correspond to the sum over all reducible diagrams with n external legs. For example, when $n = 1$, i.e., linear bias, we simply need to consider the sum of all reducible diagrams with a single external leg (corresponding to P_L at k):

$$b_1 = \left\langle \frac{\partial \delta_g}{\partial \delta} \right\rangle = \bar{b}_1 + \bar{b}_3 \frac{\sigma^2}{2} + \dots = \sum_{n=0}^{\infty} \frac{\bar{b}_{2n+1}}{n!} \left(\frac{\sigma^2}{2} \right)^n, \quad (9)$$

where we use the symbol ∂ to denote a functional derivative. Similarly, for quadratic bias ($n = 2$) we have from Eq. (1)

$$b_2 = \left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle = \bar{b}_2 + \bar{b}_4 \frac{\sigma^2}{2} + \dots = \sum_{n=0}^{\infty} \frac{\bar{b}_{2n+2}}{n!} \left(\frac{\sigma^2}{2} \right)^n, \quad (10)$$

that is, the sum over all reducible diagrams with two external legs [corresponding to P_L at k_1 and k_2 as seen in the leading-order bispectrum, Eq. (8)]. Clearly, the calculations in Eqs. (9) and (10) are significantly easier than performing the renormalization procedure order by order and statistic by statistic leading to Eqs. (6) and (8).

To conclude, we can remove the disconnect between the parameters appearing in the standard bias expansion [Eq. (1)] and those in correlators [Eqs. (6) and (8)], and therefore stop thinking about renormalization altogether, if we construct the bias expansion in terms of reducible diagrams with a given number of external legs. In effect, we trade kernels for multipoint propagators,

$$\left(\frac{\partial^n \delta_g}{\partial \delta^n} \right)_0 \rightarrow \left\langle \frac{\partial^n \delta_g}{\partial \delta^n} \right\rangle, \quad (11)$$

which leads to a new expansion of the form (denoted as Gamma expansion in [41,43])

$$\begin{aligned} \delta_g &= \left\langle \frac{\partial \delta_g}{\partial \delta} \right\rangle \delta + \frac{1}{2!} \left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle [\delta^2 - \langle \delta^2 \rangle] \\ &\quad + \frac{1}{3!} \left\langle \frac{\partial^3 \delta_g}{\partial \delta^3} \right\rangle [\delta^3 - 3\langle \delta^2 \rangle \delta - \langle \delta^3 \rangle] + \dots \quad (12) \\ &= b_1 \delta + \frac{b_2}{2!} [\delta^2 - \langle \delta^2 \rangle] + \frac{b_3}{3!} [\delta^3 - 3\langle \delta^2 \rangle \delta - \langle \delta^3 \rangle] \\ &\quad + \frac{b_4}{4!} [\delta^4 - 4\langle \delta^3 \rangle \delta - 6\langle \delta^2 \rangle \delta^2 + 3\langle \delta^2 \rangle^2 - \langle \delta^4 \rangle_c] \\ &\quad + \dots \quad (13) \end{aligned}$$

The structure of the square brackets is given by δ^n minus all possible actions of $\langle \rangle$ on δ^n with a constant term that respects that the expectation value is zero for non-Gaussian δ , and we note that Eq. (13) automatically satisfies

$\langle \delta_g \rangle = 0$. The second equality assumes local galaxy bias as we have done so far, but we stress that in this expansion the bias parameters b_n are the renormalized ones, that is, replacing the propagators by numbers gives precisely the renormalized local bias expansion.

Furthermore, if the expansion in Eq. (13) is done with respect to a Gaussian δ , for instance, when using the linear fluctuations, terms such as $\langle \delta^3 \rangle$ and $\langle \delta^4 \rangle_c$ will vanish and the expansion of δ_g will be given in terms of Hermite polynomials—a suggestion that was already put forward in [51]. Another crucial property of the multipoint propagators in this case is that they can be shown to be the cross-correlation bias between galaxies and matter fluctuations [50]; e.g., for linear and quadratic bias we have

$$\langle \delta_g \delta \rangle = \left\langle \frac{\partial \delta_g}{\partial \delta} \right\rangle \langle \delta \delta \rangle, \quad (14)$$

$$\langle \delta_g \delta \delta \rangle = \left\langle \frac{\partial^2 \delta_g}{\partial \delta^2} \right\rangle \langle \delta \delta \rangle \langle \delta \delta \rangle. \quad (15)$$

That is, the observables corresponding to the multipoint propagators are no other than the standard cross-correlation Lagrangian bias coefficients routinely measured in N -body simulations, e.g., [52–63].

Using the expansion in Eq. (13) allows us to rederive the renormalization procedure by writing it in the form of Eq. (1) and matching coefficients of δ^n , e.g., $\bar{b}_1 = b_1 - b_3 \sigma^2/2$ and $\bar{b}_2 = b_2 - b_4 \sigma^2/2$, again, without doing any calculations of correlators. This can become useful if such relations are desired for the more general case of nonlocal bias, as we shall discuss below.

Finally, the multipoint propagator expansion has the additional advantage that it simplifies the computation of correlators considerably. For instance, for the power spectrum there is only one irreducible diagram present at each loop order, as opposed to an increasing number of diagrams at each loop order for the standard approach; indeed, assuming local bias for now we simply have

$$P_g(k) = (2\pi)^3 \sum_n \frac{b_n^2}{n!} \int_{q_1, \dots, q_n} [\delta_D]_n P_L(q_1) \dots P_L(q_n), \quad (16)$$

where $[\delta_D]_n \equiv \delta_D(\mathbf{k} - \sum_{i=1}^n \mathbf{q}_i)$. For the correlation function this leads to

$$\xi_g(r) = \sum_n \frac{b_n^2}{n!} [\xi(r)]^n, \quad (17)$$

a result first obtained in [51]. If bias is not local, the functions of wave vectors that replace b_n are just placed inside the integral in Eq. (16).

C. Renormalized bias expansion with nonlinear evolution

Let us now consider the renormalized bias expansion in Eulerian space, which requires us to take into account the nonlinear evolution of the matter field. This computation serves two main objectives. First, it highlights the complexity in determining contributions to the renormalized bias parameters from evolved fields, and thus motivates our approach taken in Sec. IV, where we instead evolve the multipoint propagators from their initial conditions. Second, it demonstrates that the bias expansion in Eq. (1) cannot be completely renormalized by the local parameters b_n alone (see also [12]), calling for additional terms in the expansion of δ_g , which we address in Sec. III.

On large scales, where we can take the dark matter velocity field³ $\mathbf{v}(\mathbf{x})$ to be irrotational, the time evolution is governed by the coupled equations of motion for the density perturbations and velocity divergence $\theta(\mathbf{x}) \equiv \nabla \cdot \mathbf{v}(\mathbf{x})$ under the influence of the gravitational potential (see [64] for a detailed review). In standard Eulerian perturbation theory (SPT) these equations are solved as series expansions about the linear density field $\delta^{(1)}$, which we still assume to be Gaussian. At time τ we have (neglecting transients)

$$\delta(\mathbf{x}, \tau) = \sum_{n=1}^{\infty} D^n(\tau) \delta^{(n)}(\mathbf{x}), \quad (18)$$

where, to very good accuracy, all cosmology dependence is encoded in the linear growth factor $D(\tau)$ [19,65,66]. Note that what we denote as δ_L is actually the linear density field extrapolated to some final time (whose argument we will usually drop), i.e., $\delta_L(\mathbf{x}) \equiv D(\tau) \delta^{(1)}(\mathbf{x})$. In Fourier space the n th-order solution is constructed out of n powers of the linear density field, which are coupled via the SPT kernels F_n :

$$\delta^{(n)}(\mathbf{k}) = (2\pi)^3 \int_{\mathbf{k}_1, \dots, \mathbf{k}_n} [\delta_D]_n F_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \delta^{(1)}(\mathbf{k}_1) \dots \delta^{(1)}(\mathbf{k}_n). \quad (19)$$

The velocity divergence can be expanded in a similar manner, and the n th-order solutions are obtained by replacing F_n with G_n in the equation above. Explicit expressions for these kernels can be found in [64,67,68]. We also note that at linear order, we have $\delta = \theta$.

Because of the mode coupling in Eq. (19) the nonlinear (or time-evolved) density field becomes non-Gaussian. In order to retain the usual linear (as opposed to nonlinear)

³Note that we work with the *scaled* velocity field $\mathbf{v}$, which is related to the *peculiar* velocity via $\mathbf{v} \equiv -\mathbf{u}/f\mathcal{H}$, with f the logarithmic growth rate and $\mathcal{H}$ the comoving Hubble rate.

spectra in the loop integrals when computing clustering statistics, we, however, still define the multipoint propagators with respect to the linear Gaussian matter fluctuations; i.e., derivatives are taken with respect to δ_L . However, this implies that the kernels $(\partial^n \delta_g / \partial \delta_L^n)_0$ are not just numbers but acquire a scale dependence, in which case it becomes more convenient to express the multipoint propagators in Fourier space,

$$\left\langle \frac{\partial^n \delta_g}{\partial \delta_L^n} \right\rangle \rightarrow \left\langle \frac{\partial^n \delta_g(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)} \right\rangle, \quad (20)$$

as was already indicated in Fig. 1. While we present the resulting Fourier space equivalent to Eq. (12) in Sec. IV A, in what follows it suffices to notice that the Fourier space multipoint propagators are functions of the n momenta $\mathbf{k}_1, \dots, \mathbf{k}_n$ and include the factor $(2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n})$. Primed angle brackets indicate that we have dropped this common factor.

Returning now to the calculation of the renormalized linear bias parameter, we see that because of second-order SPT (2SPT) we have an effective cubic term inside the quadratic bias contribution, $\bar{b}_2 \delta^2 / 2$, from Eq. (1). Accordingly, the one-point propagator has the following extra term:

$$\begin{aligned} \frac{1}{2} \left\langle \frac{\partial \delta^2(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1)} \right\rangle' &= \left\langle \left[\delta_L * \frac{\partial \delta^{(2)}}{\partial \delta_L} \right](\mathbf{k}_1) \right\rangle' + \dots \\ &= 2 \underbrace{\int_{\mathbf{q}} F_2(\mathbf{k}_1, \mathbf{q}) P_L(q)}_{\bar{F}_2 \sigma^2} + \dots = \frac{34}{21} \sigma^2 + \dots, \end{aligned} \quad (21)$$

where $*$ denotes a convolution and $\bar{F}_2$ is the angular average of the 2SPT kernel (i.e., the contribution from spherical collapse to F_2). This leads to $b_1 = \bar{b}_1 + (34\bar{b}_2/21 + \bar{b}_3/2)\sigma^2 + \dots$ [7]. More broadly, since each term in the bias expansion of Eq. (1) appears at all orders in perturbation theory when δ corresponds to the Eulerian matter field, the renormalization of each bias parameter depends on all $\bar{b}_n$ ($n \geq 2$). Thus, the simplicity of Eqs. (9) and (10) is lost.

Let us consider what that implies for the renormalization of the quadratic bias, where we now have an effective quartic term (again due to 2SPT) inside the $\bar{b}_3 \delta^3 / 3!$ contribution in Eq. (1), extending the two-point propagator by

$$\begin{aligned} \frac{1}{3!} \left\langle \frac{\partial^2 \delta^3}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' &= \left\langle \delta * \frac{\partial \delta}{\partial \delta_1^L} * \frac{\partial \delta}{\partial \delta_2^L} \right\rangle' + \frac{1}{2} \left\langle \delta^2 * \frac{\partial^2 \delta}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' \\ &= [4\bar{F}_2 + F_2(\mathbf{k}_1, \mathbf{k}_2)] \sigma^2 + \dots \end{aligned} \quad (22)$$

with $\delta_i^L \equiv \delta_L(\mathbf{k}_i)$. The second term depending on the full F_2 kernel simply gives the desired contribution to the renormalization of the linear bias parameter by $\bar{b}_3$ as found for the one-point propagator before. That is because

$$\bar{b}_1 \left\langle \frac{\partial^2 \delta}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' = \bar{b}_1 2F_2(\mathbf{k}_1, \mathbf{k}_2) + \dots, \quad (23)$$

such that when calculating $\langle \partial^2 \delta_g / \partial \delta_1^L \partial \delta_2^L \rangle$ the second term in Eq. (22) corresponds to changing $\bar{b}_1 \rightarrow \bar{b}_1 + \bar{b}_3 \sigma^2 / 2$ in Eq. (23). The first term in Eq. (22) on the other hand is the contribution to the renormalization of the quadratic bias parameter from nonlinear evolution, which now reads $b_2 = \bar{b}_2 + (68\bar{b}_3/21 + \bar{b}_4/2)\sigma^2 + \dots$. In addition, quadratic bias contributes to the renormalization of b_2 itself, since

$$\begin{aligned} \frac{1}{2} \left\langle \frac{\partial^2 \delta^2}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' &= \left\langle \frac{\partial \delta}{\partial \delta_1^L} * \frac{\partial \delta}{\partial \delta_2^L} \right\rangle' + \left\langle \delta * \frac{\partial^2 \delta}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' \\ &= 1 + \left\langle \frac{\partial \delta^{(2)}}{\partial \delta_1^L} * \frac{\partial \delta^{(2)}}{\partial \delta_2^L} \right\rangle' + \left\langle \delta_L * \frac{\partial^2 \delta^{(3)}}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' \\ &\quad + \left\langle \frac{\partial \delta^{(3)}}{\partial \delta_1^L} + \frac{\partial \delta^{(3)}}{\partial \delta_2^L} \right\rangle' + \dots \end{aligned} \quad (24)$$

The last term in this expression dresses the external lines to include the one-loop propagator due to nonlinear evolution (i.e., the standard P_{13} contribution from the one-loop matter power spectrum [64]), while the remaining two terms give rise to the renormalization of the quadratic bias we are interested in, and they read

$$\begin{aligned} &\left\langle \frac{\partial \delta^{(2)}}{\partial \delta_1^L} * \frac{\partial \delta^{(2)}}{\partial \delta_2^L} \right\rangle' + \left\langle \delta_L * \frac{\partial^2 \delta^{(3)}}{\partial \delta_1^L \partial \delta_2^L} \right\rangle' \\ &= \int_{\mathbf{q}} [4F_2(-\mathbf{q}, \mathbf{k}_1) F_2(\mathbf{q}, \mathbf{k}_2) + 6F_3(\mathbf{q}, \mathbf{k}_1, \mathbf{k}_2)] P_L(q) \\ &= \frac{68}{21} F_2(\mathbf{k}_1, \mathbf{k}_2) \sigma^2 + \frac{8126}{2205} \sigma^2 + \frac{254}{2205} K(\mathbf{k}_1, \mathbf{k}_2) \sigma^2 \\ &\quad + \int_{\mathbf{q}} \mathcal{F}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}) P_L(q) \end{aligned} \quad (25)$$

with $[\mu \equiv \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2, K(\mathbf{k}_1, \mathbf{k}_2) \equiv \mu^2 - 1]$, and $\mathcal{L}_\ell$ denoting Legendre polynomials]

$$\begin{aligned}
\mathcal{F}(\mathbf{k}_1, \mathbf{k}_2, q) \equiv & \frac{(k_1^2 - q^2)^3}{168k_1^5 q^3} [(q^2 - k_1^2) \mathcal{L}_2(\mu) \\
& - 9k_1 k_2 \mu] \ln \left| \frac{k_1 + q}{k_1 - q} \right| + k_1 \leftrightarrow k_2 \\
& + \left(\frac{k_1^2 + k_2^2}{7k_1 k_2} - \frac{19k_1 k_2}{84q^2} - \frac{3(k_1^4 + k_2^4)q^2}{56k_1^3 k_2^3} \right) \mu \\
& + \left[\frac{73}{630} + \frac{k_1^2 + k_2^2}{84q^2} - \frac{11q^2}{252} \left(\frac{1}{k_1^2} + \frac{1}{k_2^2} \right) \right. \\
& \left. + \frac{q^4}{84} \left(\frac{1}{k_1^4} + \frac{1}{k_2^4} \right) \right] \mathcal{L}_2(\mu). \quad (26)
\end{aligned}$$

While the integral over the function $\mathcal{F}$ in Eq. (26) represents the finite part (going as $1/q^2$ as $q \rightarrow \infty$) of the one-loop two-point propagator due to quadratic bias, the three terms proportional to σ^2 in the second line of Eq. (25) must be absorbed by renormalizations. We recognize that the first of these corresponds to a linear bias renormalization by $\bar{b}_2$, leading to $\bar{b}_1 \rightarrow b_1 = \bar{b}_1 + (34\bar{b}_2/21 + \bar{b}_3/2)\sigma^2$ in Eq. (23), which is consistent with the result found for the one-point propagator in Eq. (21). The second term renormalizes the quadratic bias, giving $b_2 = \bar{b}_2 + (8126\bar{b}_2/2205 + 68\bar{b}_3/21 + \bar{b}_4/2)\sigma^2$ [12]. Finally, the third term has a different scale dependence compared to the previous ones, encoded by $K(\mathbf{k}_1, \mathbf{k}_2)$, such that it cannot be absorbed by a redefinition of any local bias parameter b_n . This illustrates that nonlinear evolution gives rise to additional effects that enter the relation between the galaxy and matter density and thus renders the local-in-matter expansion incomplete. In particular, we will see in Appendix D that the term proportional to $K(\mathbf{k}_1, \mathbf{k}_2)\sigma^2$ corresponds to a renormalization of the nonlocal quadratic bias $\gamma_2 = \bar{\gamma}_2 + 127\bar{b}_2\sigma^2/2205$ [12], which is associated with the gravitational tidal field.

As we alluded to at the beginning of this section, compared to the calculation when the bias relation is written in Lagrangian space, the steps taken here are a lot more complicated (and we just scratched the surface, since we only assumed local galaxy bias so far). On the other hand, if we would like to maintain the simplicity of the Lagrangian expansion, we would instead have to evolve the Lagrangian propagators by nonlinear evolution, which is where the complications may surface again. However, this is not the case for the main reason that the evolution from Lagrangian to Eulerian bias by conserving the number of objects cannot generate unphysical terms going as σ^2 . Therefore, the time-evolved propagators do not require extra renormalizations from nonlinear dynamics, and we obtain the “finite parts” such as Eq. (26) automatically.

It is further worth noting that the evolved propagators connect the Lagrangian bias parameters with the (late time) Eulerian ones; i.e., they give us the time evolution of the *observable* bias parameters, as opposed to the evolution of

the *bare* bias parameters. We discuss this in some detail in Appendix D, as to whether renormalization affects the time evolution of bias parameters.

III. A COMPLETE GALILEAN INVARIANT BASIS FOR GALAXY BIAS

We have already seen how the renormalization of local quadratic bias, when applied in Eulerian space, requires the existence of at least one additional term in the bias expansion, which was not initially included in Eq. (1). That is because gravitational evolution automatically generates a variety of terms nonlocal in δ , such as the tidal field, which in principle can affect the formation of galaxies and should consequently be taken into account in the perturbative description of galaxy bias [8,9,12,29,69].

We will refer to any such terms (both local and nonlocal in δ) as *operators*, and it is our aim in this section to provide a complete basis of operators required by the bispectrum at one-loop order. By basis we mean a set of linearly independent operators at each order of perturbation theory. We will largely follow up on the earlier work of [6,9,12,13], but distinguish between operators local in second derivatives of the linear potential and those which are not. The former will be generated even if nonlinear evolution is entirely local (Zel’dovich approximation), whereas the latter derive from corrections at second-order LPT and beyond. Accordingly, we denote these as “local evolution” (LE) and “nonlocal evolution” (NLE) operators.

Our choice of basis will make this distinction explicit and therefore differs in the type of operators from that given in [6,8,13,70]. The rationale for this choice is as follows: (1) to give the multipoint propagators a particularly simple form, and (2) to provide a hierarchy of approximations based on which one can reduce the total number of bias parameters. While the usual local Lagrangian bias approximation (where only LE operators that are also local in δ are present at the initial conditions) may well be too restrictive for cosmological parameter estimation, it might prove useful in practice to relax this at least to all LE operators. This approximation is more accurate but still reduces the number of free bias parameters compared to the case when we also allow for NLE operators at the initial conditions. Other bases in the literature mix our two sets of operators, so one might not be able to see these subtleties when measuring bias parameters from simulations or data. We discuss their relation to ours in Appendix A 2.

A. Galileons as general basis operators

Let us begin with two physical scales important for the process of galaxy formation: (1) the spatial extent R_* on which this process depends on the precise distribution of matter, and (2) the typical time T_* it takes for this matter distribution to collapse into a bound object. While the latter is a significant fraction of the Hubble time H^{-1} , the scale R_*

usually corresponds to the Lagrangian radius of the galaxy's host halo, which is of the order ~ 1 Mpc. If we are interested in the clustering of galaxies on scales $r \gg R_*$, then we can consider galaxy formation as essentially local *in space*. For now we will take this to be the case, before relaxing this assumption in Sec. III D.

As stated at the beginning of this section, other properties than density of the matter field, such as the tidal field, must enter the bias relation [8,9,29,69]. More generally, we can argue that these should depend on the (scaled) gravitational and velocity potentials, defined by

$$\nabla^2 \Phi(\mathbf{x}, \tau) \equiv \delta(\mathbf{x}, \tau), \quad (27)$$

$$\nabla^2 \Phi_v(\mathbf{x}, \tau) \equiv \theta(\mathbf{x}, \tau), \quad (28)$$

as these drive the time evolution in the regime where the dark matter flow is irrotational. According to the equivalence principle, all leading local gravitational effects must stem from second derivatives, which we write as $\nabla_{ij} \Phi(\mathbf{x}, \tau) \equiv \partial_i \partial_j \Phi(\mathbf{x}, \tau)$. Similarly, if we assume that dark matter and galaxies are comoving (i.e., no velocity bias), Galilean invariance of the equations of motion [71] implies that only second derivatives of the velocity potential are allowed to appear.

Furthermore, as δ_g is a scalar and therefore invariant under spatial coordinate transformations, we can limit ourselves to all scalar invariants of the tensors $\nabla_{ij} \Phi$ and $\nabla_{ij} \Phi_v$. In three dimensions the Cayley-Hamilton theorem [72] guarantees that there can be only three such invariants, which can be expressed by the so-called *Galileons* [9] (repeated indices are summed over):

$$\mathcal{G}_1(\Phi) \equiv \nabla^2 \Phi, \quad (29)$$

$$\mathcal{G}_2(\Phi) \equiv (\nabla_{ij} \Phi)^2 - (\nabla^2 \Phi)^2, \quad (30)$$

$$\mathcal{G}_3(\Phi) \equiv (\nabla^2 \Phi)^3 + 2 \nabla_{ij} \Phi \nabla_{jk} \Phi \nabla_{ki} \Phi - 3 (\nabla_{ij} \Phi)^2 \nabla^2 \Phi, \quad (31)$$

and similarly for Φ_v . We note that their leading SPT expressions are of first, second, and third order, respectively. By inverting the Poisson equation [Eq. (27)] we can derive their Fourier space analogs, and for the latter two we obtain

$$\mathcal{G}_2(\mathbf{k}|\Phi) = \int_{\mathbf{k}_1, \mathbf{k}_2} [\delta_D]_2 K(\mathbf{k}_1, \mathbf{k}_2) \delta(\mathbf{k}_1) \delta(\mathbf{k}_2), \quad (32)$$

$$\mathcal{G}_3(\mathbf{k}|\Phi) = \int_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} [\delta_D]_3 L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3), \quad (33)$$

where we have defined the following two kernel functions:

$$K(\mathbf{k}_1, \mathbf{k}_2) \equiv \mu_{12}^2 - 1, \quad (34)$$

$$L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \equiv 2\mu_{12}\mu_{23}\mu_{31} - \mu_{12}^2 - \mu_{23}^2 - \mu_{31}^2 + 1, \quad (35)$$

with $\mu_{ij} \equiv \mathbf{k}_i \cdot \mathbf{k}_j / k_i k_j$. Similar expressions hold for Φ_v by replacing δ with θ in Eqs. (32) and (33). Note that these kernels vanish as k^2 when $k = |\sum_i \mathbf{k}_i| \rightarrow 0$, and therefore $\langle \mathcal{G}_2 \rangle = \langle \mathcal{G}_3 \rangle = 0$ (see, e.g., [9,12]). In addition $\langle \partial \mathcal{G}_3 / \partial \delta \rangle = 0$ since $L(\mathbf{k}, \mathbf{q}, -\mathbf{q}) = 0$, which means that $\mathcal{G}_3$ cannot contribute to the one-loop galaxy propagator, as we shall discuss in Sec. IV B 1 below. All these properties make using Galileons as basis functions better suited for calculations compared to other choices that do not obey these relations, e.g., [8].

B. Local evolution operators

Let us consider the bias relation on some initial time slice in the far past. In that case we are only dealing with linear fluctuations and the single degree of freedom is Φ_L , as at linear order we have $\Phi = \Phi_v$. If we assume that objects can be identified by a local procedure on $\nabla_{ij} \Phi_L$, the only terms that can appear in the bias relation for objects at the initial conditions are $\mathcal{G}_m(\Phi_L)$ ($m \leq 3$), such that there will be n basis operators at n th order in the expansion ($n \leq 5$), i.e.,

$$\begin{aligned} \text{first: } & \mathcal{G}_1, \\ \text{second: } & \mathcal{G}_1^2, \mathcal{G}_2, \\ \text{third: } & \mathcal{G}_1^3, \mathcal{G}_1 \mathcal{G}_2, \mathcal{G}_3, \\ \text{fourth: } & \mathcal{G}_1^4, \mathcal{G}_1^2 \mathcal{G}_2, \mathcal{G}_1 \mathcal{G}_3, \mathcal{G}_2^2, \end{aligned} \quad (36)$$

and we assign a free bias parameter to each of these operators. We stress that the corresponding tracer density at the initial time will thus be a local function of $\nabla_{ij} \Phi_L$, which is similar in spirit with more phenomenological approaches, such as the peak and excursion set bias models [3,4,73,74]. In reality though, tracers are not identified at the initial time, but in the late-time, nonlinear field, which means that the basis in Eq. (36) will be insufficient if gravitational instability produces terms nonlocal in $\nabla_{ij} \Phi_L$. This is the case starting with second-order corrections to the Zel'dovich approximation [75], as we will see in Sec. III C. Therefore, even when traced back to the initial conditions such terms can in principle not be ignored. If we do so, the nonlocal terms are still produced during the evolution process, but with their bias coefficients fixed in terms of the parameters associated with the operators in Eq. (36) [9]. For that reason, this provides us with a very useful approximation (with fewer free parameters) that can be tested against numerical simulations [76].

C. Nonlocal evolution operators

Because of nonlinear evolution, the gravitational and velocity potentials will no longer be identical. Following

our previous arguments, it thus seems obvious to simply double the number of operators at each order by including a set of Galileons for both, the evolved Φ and Φ_v , and also allow for their combinations, in order to complete the basis in Eq. (36). Unfortunately, this produces a lot of redundancy as many of these operators are degenerate, so our task will be to identify those, which are linearly independent. We follow the strategy first developed in [9].

At first order in SPT we have already established that $\mathcal{G}_1(\Phi) = \mathcal{G}_1(\Phi_v)$, and we choose the former, i.e., the matter fluctuation itself, as our first basis operator. Likewise, the two second-order Galileons are degenerate at second order in SPT, and furthermore, we have

$$\mathcal{G}_1^{(2)}(\Phi) - \mathcal{G}_1^{(2)}(\Phi_v) = \delta^{(2)}(\mathbf{x}) - \theta^{(2)}(\mathbf{x}) = -\frac{2}{7}\mathcal{G}_2(\Phi_L), \quad (37)$$

proving that the basis in Eq. (36) is complete up to that order. The need for an additional operator occurs for the first time at third order. Using the notation $\Delta_n \mathcal{G}_m \equiv \mathcal{G}_m^{(n)}(\Phi) - \mathcal{G}_m^{(n)}(\Phi_v)$ for the difference between the m th Galileons evaluated at the n th order in SPT, we see that in addition to the ones already written in Eq. (36) there are the following four combinations:

$$\Delta_3 \mathcal{G}_1, \quad \Delta_3 \mathcal{G}_2, \quad \delta \Delta_2 \mathcal{G}_1, \quad \Delta_3(\mathcal{G}_1^2). \quad (38)$$

From Eq. (37) follows that the latter two are degenerate with $\delta \mathcal{G}_2(\Phi_v)$, while $\Delta_3 \mathcal{G}_1$ contains a contribution that cannot be written in terms of second derivatives of Φ_v and is thus not Galilean invariant. This only leaves the second combination, which gives

$$\Delta_3 \mathcal{G}_2 = -\frac{4}{7}[\nabla_{ij}\Phi_L \nabla_{ij}\nabla^{-2}\mathcal{G}_2(\Phi_L) - \delta \mathcal{G}_2(\Phi_L)], \quad (39)$$

demonstrating, as claimed above, that the additional basis operators induced by gravity can no longer be expressed as local functions of the linear gravitational potential, i.e., $\nabla_{ij}\Phi_L$. As these effects are precisely captured by second-order LPT and beyond, instead of explicitly calculating the differences between Galileons of Φ and Φ_v , we can follow an alternative strategy (which builds on [9]) that will prove particularly useful for extending the basis beyond third order. Let us consider LPT, which summarizes all of the dynamics in its Lagrangian displacement field

$$\Psi(\mathbf{q}, \tau) = D_1(\tau)\Psi^{(1)}(\mathbf{q}) + D_2(\tau)\Psi^{(2)}(\mathbf{q}) + \dots, \quad (40)$$

that moves particles from their initial positions $\mathbf{q}$ to their final destinations $\mathbf{x} = \mathbf{q} + \Psi(\mathbf{q}, \tau)$ at conformal time τ . The functions $\Psi^{(n)}(\mathbf{q})$ are the n th-order contributions and $D_n(\tau)$ are the corresponding growth factors ($D_1 \equiv D$ is the linear growth factor). At any order both the gravitational and velocity potentials can be expressed in terms of these

$\Psi^{(n)}$, which are in turn given by the LPT potentials φ_n [66,77], e.g.,

$$\nabla \cdot \Psi^{(1)} = \nabla^2 \varphi_1 = -\delta, \quad (41)$$

$$\nabla \cdot \Psi^{(2)} = \nabla^2 \varphi_2 = -\mathcal{G}_2(\varphi_1). \quad (42)$$

Any set of linearly independent operators induced by gravity must therefore be connected to combinations of the LPT potentials. In order to guarantee that these are still Galilean invariant, we can generalize the definition of the Galileons to [9]

$$\mathcal{G}_2(A, B) \equiv \nabla_{ij}A \nabla_{ij}B - \nabla^2 A \nabla^2 B, \quad (43)$$

and similarly for $\mathcal{G}_3(A, B, C)$. From Eq. (41) we have $\varphi_1 = -\Phi_L$, so that the first new combination appears at the third order of perturbation theory: $\mathcal{G}_2(\varphi_2, \varphi_1)$. Evaluating this Galileon using Eq. (42) shows that it is precisely related to the only gravity induced operator that we previously identified at third order, i.e., [9]

$$\Delta_3 \mathcal{G}_2 = -\frac{4}{7}\mathcal{G}_2(\varphi_2, \varphi_1). \quad (44)$$

Following this line of argument we can now easily determine the additional operators at fourth order: apart from the combination $\delta \mathcal{G}_2(\varphi_2, \varphi_1)$, we can construct the following three invariants out of the LPT potentials:

$$\mathcal{G}_2(\varphi_2, \varphi_2), \quad \mathcal{G}_2(\varphi_3, \varphi_1), \quad \mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1). \quad (45)$$

However, beyond second order the LPT solutions are no longer purely potential anymore and at third order in particular it consists of two scalar and a transverse vector potentials, all with different time dependencies,

$$\begin{aligned} \Psi(\mathbf{q}, \tau)|_{3\text{rd}} &= D_3^{(a)}(\tau)\Psi^{(3,a)}(\mathbf{q}) + D_3^{(b)}(\tau)\Psi^{(3,b)}(\mathbf{q}) \\ &+ D_3^{(c)}(\tau)\Psi^{(3,c)}(\mathbf{q}), \end{aligned} \quad (46)$$

where in the Einstein–de Sitter (EdS) approximation the growth factors are given by $D_3^{(a)}(\tau) = 1/18D(\tau)^3$, $D_3^{(b)}(\tau) = 5/42D(\tau)^3$, and $D_3^{(c)}(\tau) = 1/14D(\tau)^3$. The functions $\Psi^{(3,a/b/c)}$ and their corresponding potentials satisfy the following Poisson equations [78]:

$$\nabla \cdot \Psi^{(3,a)} = \nabla^2 \varphi_3^{(a)} = -\mathcal{G}_3(\varphi_1), \quad (47)$$

$$\nabla \cdot \Psi^{(3,b)} = \nabla^2 \varphi_3^{(b)} = -\mathcal{G}_2(\varphi_2, \varphi_1), \quad (48)$$

$$-\nabla \times \Psi^{(3,c)} = \nabla^2 \mathbf{A}_3 = -\hat{e}_i \epsilon_{ijk} (\nabla_{jl} \varphi_1) (\nabla_{kl} \varphi_2), \quad (49)$$

where ϵ_{ijk} denotes the fully antisymmetric Levi-Civita symbol and $\hat{e}_i$ the unit vector in direction i .

TABLE I. Overview of basis operators for galaxy bias, along with their associated bias parameters. Each single column presents a different order in SPT, and we have categorized operators into groups, which are local and nonlocal in second derivatives of the linear gravitational potential (left and middle columns, respectively), and which contain higher than second derivatives of the potentials. Note that to simplify notation, we have relabeled some bias parameters compared to [9]—what was formerly $\gamma_3^\times$ is now $\gamma_2^\times$ and γ_3^- has become $-7/4\gamma_{21}$. A superscript $\times$ denotes that a Galileon field has been multiplied by δ . Operators related to noise terms are not listed here; see Sec. V C.

Local evolution operators				Nonlocal evolution operators		Higher derivative	
First	Second	Third	Fourth	Third	Fourth	Third	Fourth
$b_1\delta$	$b_2\delta^2/2$ $\gamma_2\mathcal{G}_2(\Phi_v)$	$b_3\delta^3/3!$ $\gamma_2^\times\delta\mathcal{G}_2(\Phi_v)$ $\gamma_3\mathcal{G}_3(\Phi_v)$	$b_4\delta^4/4!$ $\gamma_2^{\times\times}\delta^2\mathcal{G}_2(\Phi_v)$ $\gamma_3^\times\delta\mathcal{G}_3(\Phi_v)$ $\gamma_2^{\text{sq}}\mathcal{G}_2(\Phi_v)^2$	$\gamma_{21}\mathcal{G}_2(\varphi_2, \varphi_1)$	$\gamma_{21}^\times\delta\mathcal{G}_2(\varphi_2, \varphi_1)$ $\gamma_{211}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1)$ $\gamma_{22}\mathcal{G}_2(\varphi_2, \varphi_2)$ $\gamma_{31}\mathcal{G}_2(\varphi_3, \varphi_1)$	$\beta_1\nabla^2\delta$	$\beta_{2,1}\nabla^2\delta^2$ $\beta_{2,2}(\nabla\delta)^2$ $\beta_{2,3}\nabla^2\mathcal{G}_2(\Phi_v)$ $\beta_{2,4}\mathcal{G}_2(\nabla\Phi_v)$

The combination of the third and first-order LPT potentials is thus made up of three pieces, and factoring out $D(\tau)^3$ from the EdS solutions, we define

$$\begin{aligned} \mathcal{G}_2(\varphi_3, \varphi_1) \equiv & \frac{1}{18}\mathcal{G}_2(\varphi_3^{(a)}, \varphi_1) + \frac{5}{42}\mathcal{G}_2(\varphi_3^{(b)}, \varphi_1) \\ & + \frac{1}{14}\nabla_i(\nabla \times \mathbf{A}_3)_j \nabla_{ij}\varphi_1. \end{aligned} \quad (50)$$

Because of the different time dependencies we should in principle allow these three pieces to enter the bias basis individually, but in practice the departures from the EdS approximation to growth factors are below the percent level, so for all purposes of this paper we are safe to ignore this complication.

To conclude, our choice of a complete Galilean invariant basis for the evolved galaxy perturbations is given by a set of 15 operators up to fourth order, which are summarized in the first two columns of Table I. We have separated what we denote as local evolution operators (first column), which are local in $\nabla_{ij}\Phi_L$, from those that correspond to nonlinear corrections to the gravitational and velocity potentials that are nonlocal in $\nabla_{ij}\Phi_L$ (middle column). With respect to the number of basis operators we are thus in agreement with [6], the relation between our and their basis up to fourth order is given in Appendix A 2.

D. Higher-derivative galaxy bias

Although some of the basis operators we derived in Secs. III B and III C are nonlocal in the matter fluctuations and gravitational potential, we made the central assumption that the formation of galaxies is *spatially local*, meaning it is determined by the value of these operators at a single point in space. On small scales this approximation must break down because galaxies form due to matter collapsing from a finite region of size R_* , which is of the order of the Lagrangian radius of the host dark matter halo. As shown in [6,8,70] this gives rise to additional operators in the bias

expansion that contain higher than second derivatives of the gravitational and velocity potentials.⁴

As these operators must still be scalars, the simplest one involves four derivatives of Φ and is given by $\nabla^2\delta$. Its effect can be interpreted as an emerging scale dependence of the linear bias parameter, because upon Fourier transformation and grouping all terms linear in δ , we get

$$b_1(k) = b_1 - \beta_1 k^2 + \mathcal{O}(k^4), \quad (51)$$

where β_1 is the bias parameter associated with $\nabla^2\delta$ and $\mathcal{O}(k^4)$ stands for contributions involving even higher derivatives. Note that β_1 has units of length squared, and as we can expect it to scale with R_* , we see that on scales much larger than the Lagrangian radius, i.e., $kR_* \ll 1$, the higher-derivative contributions are suppressed by powers of $(kR_*)^2$ compared to the spatially local linear term. This is similar to the nonlinear bias terms, where an increase in nonlinear order is suppressed by powers of $\delta \simeq (k/k_{\text{nl}})^{(n_{\text{eff}}+3)/2}$ at large scales, with $k_{\text{nl}}^3 P(k_{\text{nl}})/(2\pi^2) \equiv 1$ and n_{eff} the effective spectral index at the nonlinear scale. However, comparing the nonlinear and higher-derivative contributions against each other *a priori* is difficult, since this depends on the relative size of R_* (which in turn depends on the biased tracer in question) and k_{nl} and n_{eff} (which depend on the linear spectrum and redshift), in addition to the size of the bias parameters. We therefore take the following approach: we consider *a priori* each higher-derivative factor in correlators as a nonlinear bias loop, counting each additional derivative acting on the gravitational or velocity potentials as an increase of the SPT order by one. The operator $\nabla^2\delta$ would thus be considered as third order. After performing a measurement of bias parameters from clustering data, one can reassess *a posteriori* whether bias loops of derivatives are more important

⁴These have been known for a very long time, e.g., [3] since they naturally appear in peak models of biased tracers.

for a particular biased tracer. In [31], the authors conclude that higher-derivative biases are more important than loops for tracers of a very wide range of masses. We revisit this issue in [76], but Figs. 4 and 5 below already suggest that loop corrections are as important for the bispectrum as for the power spectrum and that they matter on scales commonly used in the analysis of galaxy surveys.

Based on our counting, the fourth-order higher-derivative terms will simply be given by acting with two derivatives on the spatially local quadratic bias operators. This allows only for the following four independent combinations:

$$\nabla^2 \delta^2, \quad (\nabla \delta)^2, \quad \nabla^2 \mathcal{G}_2(\Phi_v), \quad \mathcal{G}_2(\nabla_i \Phi_v, \nabla_i \Phi_v), \quad (52)$$

where $\delta \nabla^2 \delta$ as well as $\mathcal{G}_2(\Phi_v, \nabla^2 \Phi_v)$ do not enter individually, as they can be expressed through combinations of the operators above. In total we thus obtain a set of five higher-derivative operators, which are summarized in Table I according to their classification in terms of SPT order as discussed above. Note that Ref. [6] gives a total of ten operators up to the same order. We find that the set presented here is equivalent with the first five operators in their Eq. (2.74), while the last five can be expressed in terms of the other ones, making them superfluous.

IV. MULTIPOINT PROPAGATORS IN LAGRANGIAN AND EULERIAN SPACE

We now return to the main idea presented in Sec. II B: in order to guarantee that the bias parameters corresponding to all of the Galilean basis operators are observable quantities, we should construct the galaxy density field out of multipoint propagators,

$$\left\langle \frac{\partial^n \delta_g(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)} \right\rangle, \quad (53)$$

as opposed to the usual kernel functions that we obtain from $[\partial^n \delta_g(\mathbf{k}) / \partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)]_0$. Our goal is to compute the complete multipoint propagators at initial and final times, while paying particular attention to the evolution of the bias parameters. Before that we illustrate their relation to the Wiener-Hermite functionals, which formalizes the multipoint propagator expansion in Fourier space and will prove useful for determining their time evolution.

A. Relating multipoint propagators and the Wiener-Hermite expansion

For a linear Gaussian dark matter density field, the probability density function (PDF) for a mode in Fourier space is given by

$$\mathcal{P}[\delta_L] = N \exp \left[-\frac{1}{2} \int_q \frac{|\delta_L(\mathbf{q})|^2}{P_L(q)} \right], \quad (54)$$

with normalization factor N . The n th generalized Wiener-Hermite functional $\mathcal{H}_n$ is then defined by taking n functional derivatives of the PDF [79]:

$$\frac{\mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n)}{\prod_{i=1}^n P_L(\mathbf{k}_i)} \equiv \frac{(-1)^n}{\mathcal{P}[\delta_L]} \frac{\partial^n \mathcal{P}[\delta_L]}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)}. \quad (55)$$

Using this definition, we obtain the following first three functionals (suppressing the momentum arguments):

$$\begin{aligned} \mathcal{H}_1 &= \delta_L^*(\mathbf{k}), \\ \mathcal{H}_2 &= \delta_L^*(\mathbf{k}_1) \delta_L^*(\mathbf{k}_2) - \langle \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) \rangle, \\ \mathcal{H}_3 &= \delta_L^*(\mathbf{k}_1) \delta_L^*(\mathbf{k}_2) \delta_L^*(\mathbf{k}_3) \\ &\quad - [\langle \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) \rangle \delta_L^*(\mathbf{k}_3) + \text{cyc}], \end{aligned} \quad (56)$$

where a superscript $*$ denotes complex conjugation. As the standard Hermite polynomials they satisfy an orthogonality relation, which can be shown to be (see [79] and Appendix B 1)

$$\begin{aligned} &\langle \mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \mathcal{H}_m^*(\mathbf{q}_1, \dots, \mathbf{q}_n) \rangle \\ &= (2\pi)^{3n} \delta_{nm}^K [\delta_D(\mathbf{k}_1 - \mathbf{q}_1) \dots \delta_D(\mathbf{k}_n - \mathbf{q}_n) \\ &\quad + \text{sym}] \prod_{i=1}^n P_L(\mathbf{k}_i), \end{aligned} \quad (57)$$

where δ_{nm}^K is the Kronecker delta, and “sym.” stands for the remaining $(n! - 1)$ combinations of the arguments. Expanding the Fourier space galaxy overdensity in terms of Hermite functionals gives a convolution over $\mathcal{H}_n$ at each order, such that

$$\begin{aligned} \delta_g(\mathbf{k}) &= \sum_n \frac{(2\pi)^3}{n!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_n} [\delta_D]_n \Gamma_g^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) \\ &\quad \times \mathcal{H}_n^*(\mathbf{k}_1, \dots, \mathbf{k}_n), \end{aligned} \quad (58)$$

where the coefficients $\Gamma_g^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n)$ are scale-dependent functions (kernels) that can be interpreted as the corresponding bias parameters. Multiplying both sides of Eq. (58) with $\mathcal{H}_m$ and using the orthogonality relation [Eq. (57)], we see that

$$\begin{aligned} \frac{\langle \mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \delta_g(\mathbf{k}) \rangle}{\prod_{i=1}^n P_L(\mathbf{k}_i)} &= (2\pi)^3 \Gamma_g^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) \\ &\quad \times \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n}), \end{aligned} \quad (59)$$

which generalizes Eqs. (14) and (15). Cross-correlating the galaxy density field with Hermite polynomials to measure local bias parameters has been proposed in a similar way in [80]. Moreover, this is also how propagators for the mapping from linear to nonlinear fluctuations have been

measured in [41,43,50]. Note that in general for nonlocal bias this cross-correlation has to be further decomposed in terms of the structures inside $\Gamma_g^{(n)}$ to end up with parameter estimates (i.e., one has to separate b_2 from γ_2 contributions, etc.), e.g., [23,60,61,81,82], but that procedure is basis dependent.

We can derive a different relation by plugging Eq. (55) into Eq. (59) and replacing the ensemble average by its definition—the functional integral of all modes δ_L over their joint PDF,

$$\begin{aligned} & \frac{\langle \mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \delta_g(\mathbf{k}) \rangle}{\prod_{i=1}^n P_L(\mathbf{k}_i)} \\ &= (-1)^n \int \mathcal{D}[\delta_L] \left[\frac{\partial^n \mathcal{P}[\delta_L]}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)} \right] \delta_g(\mathbf{k}) \\ &= \int \mathcal{D}[\delta_L] \mathcal{P}[\delta_L] \frac{\partial^n \delta_g(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)} \\ &= \left\langle \frac{\partial^n \delta_g(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)} \right\rangle, \end{aligned} \quad (60)$$

where we have integrated by parts n times in going from the second to the third line and ignored any surface terms as $\mathcal{P}[\delta_L] \rightarrow 0$ (same as its derivatives) for $\delta_L \rightarrow \pm\infty$. Thus, it follows that the kernels of the Wiener-Hermite expansion are indeed the multipoint propagators:

$$\begin{aligned} & \left\langle \frac{\partial^n \delta_g(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \dots \partial \delta_L(\mathbf{k}_n)} \right\rangle \\ &= (2\pi)^3 \Gamma_g^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n}). \end{aligned} \quad (61)$$

This equation is equivalent to the definition of multipoint propagators in the previous work of [41,43], which considered the nonlinear matter fluctuations in place of the galaxy fluctuations. As they showed that the multipoint propagators function as basic building blocks for constructing arbitrary N -point spectra, the same holds for the $\Gamma_g^{(n)}$, and we will use this fact to compute the galaxy power spectrum and bispectrum in Sec. V.

B. Initial conditions

We now determine the first three multipoint propagators at an initial time where nonlinearities in the dark matter field can be ignored, which implies we can set all SPT kernels F_n and G_n for $n \geq 2$ to zero. We include all of the basis operators summarized in Table I and work up to fourth order in SPT, as required by the one-loop bispectrum. For reasons explained below, the general structure of the multipoint propagators in our basis is very simple: they are given by a tree-level contribution consisting of all basis operators that correspond to the order of the propagator itself, in addition to loop corrections involving only NLE operators.

1. The one-point propagator

To compute the one-point propagator, we need to take a single derivative of δ_g and take the expectation value, which implies due to Gaussianity of δ_L that only odd orders in the bias expansion enter. More generally, as each derivative cancels exactly one factor of δ_L , we see that odd (even) numbered propagators can only contain terms stemming from odd (even) orders of the bias expansion.

The first term in the bias expansion gives just linear bias; thus to compute $\Gamma_g^{(1)}$ to one loop we need the derivative of a generic third-order term $\mathcal{O}_B^{(3)}$, which can be written as the convolution

$$\begin{aligned} \mathcal{O}_B^{(3)}(\mathbf{k}) &= (2\pi)^3 \int_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} [\delta_D]_3 \mathcal{K}_B^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \\ &\quad \times \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) \delta_L(\mathbf{k}_3), \end{aligned} \quad (62)$$

with $B \in \{\delta^3, \delta\mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_2(\varphi_2, \varphi_1)\}$ and the kernels $\mathcal{K}_B^{(3)}$ are given in Eqs. (A5)–(A8). Making use of the symmetry of $\mathcal{K}_B^{(3)}$, we obtain

$$\left\langle \frac{\partial \mathcal{O}_B^{(3)}(\mathbf{k})}{\partial \delta_L(\mathbf{k}')} \right\rangle' = 3 \int_{\mathbf{q}} \mathcal{K}_B^{(3)}(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_L(\mathbf{q}), \quad (63)$$

using the same notation for ensemble averages as in Sec. II C. We note that Eq. (63) is analogous to the result obtained for the one-loop contribution to the matter propagator, which gives rise to $P_{13}/2P_L$ [40].

As pointed out already, most of the possibilities for third-order kernel contributions are trivial or zero in our choice of basis functions for bias. In fact, no basis function local in $\nabla_{ij}\Phi$ (denoted as LE operators in Table I) can give a nontrivial contribution to the loop corrections of $\Gamma_g^{(1)}$ (with trivial contributions meaning renormalizations of the linear bias parameter, as discussed in Sec. II B). The reason is that in propagator loop integrals no wave vector angles can appear inside power spectra (due to being reducible diagrams); thus one can always perform angular integrations over momenta of the resulting kernels, which for operators local in $\nabla_{ij}\Phi$ simply give rise to numerical prefactors (many of them zero in our choice of basis). Therefore such loops are only functions of σ^2 , e.g., for $B = \delta\mathcal{G}_2$,

$$\left\langle \frac{\partial \mathcal{O}_{\delta\mathcal{G}_2}^{(3)}(\mathbf{k})}{\partial \delta_L(\mathbf{k}')} \right\rangle' = 2 \int_{\mathbf{q}} K(\mathbf{k}, \mathbf{q}) P_L(\mathbf{q}) = -\frac{4}{3} \sigma^2, \quad (64)$$

similar to the case $B = \delta^3$, while $\langle \partial \mathcal{G}_3 / \partial \delta_L \rangle = 0$ since $L(\mathbf{k}, \mathbf{q}, -\mathbf{q}) = 0$. On the other hand, for the only NLE operator (see Table I) at third order $B = \mathcal{G}_2(\varphi_2, \varphi_1)$, we get

$$\left\langle \frac{\partial \mathcal{O}_{\mathcal{G}_2(\varphi_2, \varphi_1)}^{(3)}(\mathbf{k})}{\partial \delta_L(\mathbf{k}')} \right\rangle' = 2 \int_q K(\mathbf{k} - \mathbf{q}, \mathbf{q}) K(\mathbf{k}, \mathbf{q}) P_L(q). \quad (65)$$

Collecting all these results, the first galaxy propagator at initial time is given by

$$\Gamma_g^{(1)}(\mathbf{k}) = b_1 - \beta_1 k^2 + 2\gamma_{21} \int_q K(\mathbf{k} - \mathbf{q}, \mathbf{q}) K(\mathbf{k}, \mathbf{q}) P_L(q), \quad (66)$$

where

$$b_1 \equiv \bar{b}_1 + \left[\frac{1}{2} \bar{b}_3 - \frac{4}{3} \bar{\gamma}_2^\times \right] \sigma^2, \quad \gamma_{21} \equiv \bar{\gamma}_{21}, \quad (67)$$

and we have added the higher-derivative term. As noted in Sec. II B, Eq. (67) can also be derived from the renormalized expansion in Eq. (13) without explicitly computing the loops for the terms local in $\nabla_{ij}\Phi$.

In the limit $k \rightarrow 0$, the γ_{21} integral scales as

$$\int_q K(\mathbf{k} - \mathbf{q}, \mathbf{q}) K(\mathbf{k}, \mathbf{q}) P_L(q) = \frac{8}{15} k^2 \int_q \frac{P_L(q)}{q^2} + \mathcal{O}(k^4), \quad (68)$$

which ensures that $\Gamma_g^{(1)}$ indeed corresponds to the linear bias parameter on large scales. Moreover, we note that if terms of order k^4 and beyond are negligible, the γ_{21} contribution is entirely degenerate with the higher-derivative term [see also [8,24,83]]. As long as we restrict ourselves to sufficiently large scales, this implies it is not necessary to include the higher-derivative parameter in excess of γ_{21} . However, *a priori* it is difficult to tell what “sufficient” means for a given data set, and therefore requires careful testing using mock data.

2. The two-point propagator

In a similar manner we can now derive all remaining multipoint propagators. The two-point propagator gets contributions from second and fourth-order bias operators, which, analogous to Eq. (62), we write as the symmetric kernels $\mathcal{K}_B^{(2)}$ and $\mathcal{K}_B^{(4)}$. The corresponding expressions are given in Eqs. (A2) and (A3), and Eqs. (A12)–(A19), respectively. Differentiating a generic second or fourth-order contribution twice results in

$$\left\langle \frac{\partial^2 \mathcal{O}_B^{(2)}(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \partial \delta_L(\mathbf{k}_2)} \right\rangle' = 2\mathcal{K}_B^{(2)}(\mathbf{k}_1, \mathbf{k}_2) \quad (69)$$

and

$$\left\langle \frac{\partial^2 \mathcal{O}_B^{(4)}(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \partial \delta_L(\mathbf{k}_2)} \right\rangle' = 12 \int_q \mathcal{K}_B^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, -\mathbf{q}) P_L(q). \quad (70)$$

As explained above, the nontrivial loop corrections to the two-point galaxy propagator can only consist of the NLE operators, and so we get

$$\begin{aligned} \Gamma_g^{(2)}(\mathbf{k}_1, \mathbf{k}_2) &= b_2 + 2\gamma_2 K(\mathbf{k}_1, \mathbf{k}_2) \\ &+ 12 \int_q \left[\gamma_{21}^\times \mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_1)}^{(4,F)} + \gamma_{211} \mathcal{K}_{\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1)}^{(4)} \right. \\ &+ \gamma_{22} \mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_2)}^{(4)} + \gamma_{31} \mathcal{K}_{\mathcal{G}_2(\varphi_3, \varphi_1)}^{(4)} \left. \right] P_L(q) \\ &- \beta_{2,1} k_{12}^2 - \beta_{2,2} (\mathbf{k}_1 \cdot \mathbf{k}_2) \\ &- [\beta_{2,3} k_{12}^2 + \beta_{2,4} (\mathbf{k}_1 \cdot \mathbf{k}_2)] K(\mathbf{k}_1, \mathbf{k}_2), \end{aligned} \quad (71)$$

where $k_{12}^2 \equiv |\mathbf{k}_1 + \mathbf{k}_2|^2$ and the bracket inside the loop integral is evaluated at $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, -\mathbf{q})$. Furthermore, the kernel $\mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_1)}^{(4,F)}$ denotes the finite part of $\mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_1)}^{(4)}$ (i.e., removing contributions proportional to σ^2), and we have added the higher-derivative terms in the last two lines. For the same reasons mentioned above, all of the initial operators again only contribute to renormalizations of

$$b_2 \equiv \bar{b}_2 + \left[\frac{1}{2} \bar{b}_4 - \frac{16}{3} \bar{\gamma}_2^{\times\times} + \frac{32}{15} \bar{\gamma}_{21}^\times + \frac{64}{15} \bar{\gamma}_2^{\text{sq}} \right] \sigma^2, \quad (72)$$

$$\gamma_2 \equiv \bar{\gamma}_2 + \left[\bar{\gamma}_2^{\times\times} + \frac{2}{5} \bar{\gamma}_{21}^\times - \frac{1}{2} \bar{\gamma}_3^\times + \frac{8}{15} \bar{\gamma}_2^{\text{sq}} \right] \sigma^2. \quad (73)$$

These are the observable quadratic and tidal tensor bias parameters, which match those in Eqs. (3.16) and (3.17) in [12] up to the signs of the $\gamma_{21}^\times$ and $\gamma_3^\times$ terms in the expression for γ_2 (note that in their notation $b_{\Gamma_{3\delta}} = -4/7\gamma_{21}^\times$ and $b_{\mathcal{G}_3\delta} = \gamma_3^\times$). A more direct comparison with [12] including renormalization due to nonlinear time evolution is discussed in Appendix D. Again, the relations in Eqs. (72) and (73) can be obtained directly from Eq. (13) without computing the loops explicitly.

The large-scale behavior of the nonlocal evolution terms that we found for the one-point propagator in the previous section also applies to the two-point propagator. More precisely, it can be shown that the four-loop integrals in Eq. (71) are fully expressible in terms of combinations of the four higher-derivative operators (and vice versa), and we present the corresponding relations in Appendix A 3. As noted above, this means that the higher-derivative operators become degenerate with fourth-order contributions from the general bias expansion, and their impact is automatically accounted for by the latter, as long as terms of order k_1^4 , k_2^4 , etc., are negligible.

3. The three-point propagator

Finally, for the three-point propagator, we compute three derivatives of $\mathcal{O}_B^{(3)}$, giving

$$\left\langle \frac{\partial^3 \mathcal{O}_B^{(3)}(\mathbf{k})}{\partial \delta_L(\mathbf{k}_1) \partial \delta_L(\mathbf{k}_2) \partial \delta_L(\mathbf{k}_3)} \right\rangle' = 6\mathcal{K}_B^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \quad (74)$$

At the order of SPT we are working in, we only need the tree-level expression for $\Gamma_g^{(3)}$, and we simply get

$$\begin{aligned} \Gamma_g^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = & b_3 + 2\gamma_2^\times [K(\mathbf{k}_1, \mathbf{k}_2) + \text{cyc}] \\ & + 2\gamma_{21} [K(\mathbf{k}_1, \mathbf{k}_2)K(\mathbf{k}_2, \mathbf{k}_3) + \text{cyc}] \\ & + 6\gamma_3 L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \end{aligned} \quad (75)$$

C. Time evolution

As we already discussed in Sec. II C for the case of local galaxy bias, until the time when galaxies are observed, all of the basis operators will have evolved and thus developed nonlinear corrections that are of higher SPT orders. Consequently, if we were to compute the galaxy multipoint propagators at late times, we would have to account for these corrections, which lead to numerous new renormalizations [see discussion related to Eqs. (25) and (26)]. Especially when pushed to fourth order in SPT this approach becomes very cumbersome.

For that reason this section presents an alternative—we start from the multipoint propagator expansion at initial time using the propagators derived in Sec. IV B and then evolve this expansion instead of resorting to the usual SPT solutions. Because the mapping from the initial time to the final time conserves the number of tracers, no new divergences going like σ^2 can arise, and this bypasses the need to deal with these complexities. Therefore, the only bias renormalizations done in our approach are trivial ones, done at initial time, and, in fact, dealt simply by using the multipoint propagators: in this sense, in our approach one does not have to think about renormalization at all.

Since time evolution from the initial conditions also gives rise to an evolution of the bias parameters, we first consider them separately from the propagators. This allows us to illustrate precisely how various assumptions made about the initial bias relation affect the late-time galaxy overdensity, and thus generalizes previous results in the literature up to fourth order. In order to distinguish initial from evolved quantities, we will make the following changes to notation:

$$b, \quad \gamma \rightarrow b_{\mathcal{L}}, \gamma_{\mathcal{L}} \quad \text{and} \quad \Gamma^{(n)} \rightarrow \Gamma_{\mathcal{L}}^{(n)}, \quad (76)$$

for the initial bias parameters and propagators, respectively.

1. Bias parameters

One of the key principles underpinning the bias relation is that it must be Galilean invariant if there is no velocity bias. We can exploit this symmetry to determine the evolution of the bias parameters in a significantly simpler way by restricting all quantities that appear in intermediate steps of the calculation to a Galilean invariant basis of operators, i.e., the basis we presented in Sec. III. To illustrate this novel technique we start from the evolution equations for conserved tracers, which in the absence of velocity bias can be directly integrated to give [see Eq. (48) in [9]]

$$1 + \delta_g(\mathbf{x}) = \frac{1 + \delta_g^{\text{IC}}(\mathbf{q})}{1 + \delta^{\text{IC}}(\mathbf{q})} [1 + \delta(\mathbf{x})], \quad (77)$$

where Lagrangian fields (with argument $\mathbf{q}$) are evaluated at the initial time when $\delta = \delta_L$. Since we are not interested in decaying modes, we can neglect $\delta^{\text{IC}}(\mathbf{q})$ in the denominator because it is suppressed by one growth factor compared to the Eulerian fields, leading to the well-known expression for the evolution of bias [84]. Similarly, we can ignore decaying modes and write the initial bias relation in terms of the extrapolated (to final time) linear density fluctuations and Lagrangian bias parameters (identical to the parameters in Sec. IV B),

$$\begin{aligned} \delta_g^{\text{IC}} = & b_{1,\mathcal{L}} \delta_L + \frac{b_{2,\mathcal{L}}}{2} \delta_L^2 + \gamma_{2,\mathcal{L}} \mathcal{G}_2 + \frac{b_{3,\mathcal{L}}}{6} \delta_L^3 + \gamma_{3,\mathcal{L}} \mathcal{G}_3 \\ & + \gamma_{2,\mathcal{L}}^\times \delta_L \mathcal{G}_2 + \gamma_{21,\mathcal{L}} \mathcal{G}_2(\varphi_2, \varphi_1) + \dots, \end{aligned} \quad (78)$$

where all right-hand side (RHS) fields are linear Gaussian evaluated at the final time. To proceed we would then have to use $\mathbf{q} = \mathbf{x} - \Psi$ to relate initial to final positions and expand all quantities around $\mathbf{x}$. However, this can be done trivially by realizing that in order for the left-hand side (LHS) of Eq. (77) to be Galilean invariant, so must the RHS, which implies that the dipole terms that arise from the $\mathbf{q}$ to $\mathbf{x}$ mapping must cancel against those in the nonlinear matter density. All we need to do is thus to consider the Galilean invariant contributions to the nonlinear density and velocity expressed in terms of our bias basis, which is done in Appendix C 1. Indicating this by the superscript “GI,” it follows that

$$\delta_g(\mathbf{x}) = \delta_g^{\text{IC}}(\mathbf{x}) + \delta^{\text{GI}}(\mathbf{x}) + \delta_g^{\text{IC}}(\mathbf{x}) \delta^{\text{GI}}(\mathbf{x}) \quad (79)$$

and by using our bias basis at each order we can then find the coefficients (bias parameters) for the Eulerian expansion in terms of those in the initial expansion. For instance, at linear order we have $\delta_g^{\text{IC}}(\mathbf{x}) = b_{1,\mathcal{L}} \delta_L(\mathbf{x})$ and plugging this into Eq. (79) gives

$$b_1 = 1 + b_{1,\mathcal{L}}, \quad (80)$$

which agrees with the result from full evolution [69] when neglecting transients. For higher-order bias evolution we

proceed as in [9]; i.e., to find bias parameters at a given order we subtract the contributions expected from the lower-order bias parameters. The full details of this computation are given in Appendix C 2, and at second order we get the following well known relations:

$$b_2 = b_{2,\mathcal{L}}, \quad \gamma_2 = -\frac{2}{7}b_{1,\mathcal{L}} + \gamma_{2,\mathcal{L}}, \quad (81)$$

while we obtain

$$\begin{aligned} b_3 &= b_{3,\mathcal{L}} - 3b_2, \quad \gamma_3 = -\frac{1}{9}b_{1,\mathcal{L}} - \gamma_2 + \gamma_{3,\mathcal{L}}, \\ \gamma_2^\times &= -\frac{2}{7}b_2 + \gamma_{2,\mathcal{L}}^\times, \quad \gamma_{21} = \frac{2}{21}b_{1,\mathcal{L}} + \frac{6}{7}\gamma_2 + \gamma_{21,\mathcal{L}} \end{aligned} \quad (82)$$

for cubic bias [note that there was a typo in Eq. (116) of [9], which should have had a minus sign for its second term], and finally

$$\begin{aligned} b_4 &= b_{4,\mathcal{L}} - 12b_2 - 8b_3, \\ \gamma_2^{\times\times} &= -\frac{3}{7}b_2 - \frac{1}{7}b_3 - \gamma_2^\times + \gamma_{2,\mathcal{L}}^{\times\times}, \\ \gamma_3^\times &= -\frac{b_2}{9} - \gamma_2^\times + \gamma_{3,\mathcal{L}}^\times, \end{aligned} \quad (83)$$

$$\begin{aligned} \gamma_2^{\text{sq}} &= -\frac{2}{49}b_2 - \frac{2}{7}\gamma_2^\times + \gamma_{2,\mathcal{L}}^{\text{sq}}, \\ \gamma_{21}^\times &= \frac{2}{21}b_2 + \frac{6}{7}\gamma_2^\times + \gamma_{21,\mathcal{L}}^\times, \\ \gamma_{31} &= -\frac{4}{11}b_{1,\mathcal{L}} - 6\gamma_2 + \gamma_{31,\mathcal{L}}, \end{aligned} \quad (84)$$

$$\begin{aligned} \gamma_{22} &= -\frac{6}{539}b_{1,\mathcal{L}} - \frac{9}{49}\gamma_2 + \gamma_{22,\mathcal{L}}, \\ \gamma_{211} &= \frac{5}{77}b_{1,\mathcal{L}} + \frac{15}{14}\gamma_2 + \gamma_{21} - \frac{9}{7}\gamma_3 + \gamma_{211,\mathcal{L}}, \end{aligned} \quad (85)$$

for quartic bias. Appendix C 3 shows how to recover the familiar spherically symmetric results from these expressions, which serves as a robust consistency check. Although the relations above have been derived for the *bare* bias parameters, they are equally valid for the renormalized ones. This is demonstrated explicitly in Appendix D, which considers what happens if one applies Eq. (79) to the renormalized bias expansion. However, as we mentioned above, in this paper we advocate a different, simpler route that evolves the initial multipoint propagators, as we discuss next.

2. Propagators

As a first step we write a combined evolution equation for matter, its velocity divergence, and galaxies, and for that purpose we define the three-component vector

$$\Psi(\mathbf{k}, \tau) \equiv (\delta(\mathbf{k}, \tau), \theta(\mathbf{k}, \tau)/f\mathcal{H}, \delta_g(\mathbf{k}, \tau)). \quad (86)$$

In terms of Ψ and by using the logarithm of the linear growth rate as our new time variable, i.e., $\eta \equiv \ln D(\tau)$, the evolution equations can be recast as (see [9])

$$\begin{aligned} \frac{\partial \Psi_a(\mathbf{k}, \eta)}{\partial \eta} + \Omega_{ab} \Psi_b(\mathbf{k}, \eta) \\ = (2\pi)^3 \int_{\mathbf{k}_1, \mathbf{k}_2} [\delta_D]_2 \gamma_{abc}(\mathbf{k}_1, \mathbf{k}_2) \\ \times \Psi_b(\mathbf{k}_1, \eta) \Psi_c(\mathbf{k}_2, \eta), \end{aligned} \quad (87)$$

where we follow the convention that repeated indices are summed over, and we assume galaxies move with the dark matter, i.e., no velocity bias. The matrix

$$\Omega_{ab} \equiv \frac{1}{2} \begin{bmatrix} 0 & -2 & 0 \\ -3 & 1 & 0 \\ 0 & -2 & 0 \end{bmatrix} \quad (88)$$

describes the coupling between densities and velocities, while γ_{abc} encodes the nonlinear interactions between different Fourier modes. Its only nonzero components are given by

$$\gamma_{121} = \gamma_{323} = \alpha(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{\mathbf{k}_{12} \cdot \mathbf{k}_1}{k_1^2}, \quad (89)$$

$$\gamma_{222} = \beta(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{k_{12}^2(\mathbf{k}_1 \cdot \mathbf{k}_2)}{2k_1^2 k_2^2}, \quad (90)$$

and $\gamma_{112}(\mathbf{k}_1, \mathbf{k}_2) = \gamma_{121}(\mathbf{k}_2, \mathbf{k}_1)$. Given some arbitrary initial conditions $\Psi_a(\mathbf{k}, \eta = 0) \equiv \phi_a(\mathbf{k})$, Eq. (87) has the integral solution [85]

$$\begin{aligned} \Psi_a(\mathbf{k}, \eta) &= g_{ab}(\eta) \phi_b(\mathbf{k}) + (2\pi)^3 \int_0^\eta d\eta' g_{ab}(\eta - \eta') \\ &\times \int_{\mathbf{k}_1, \mathbf{k}_2} [\delta_D]_2 \gamma_{bcd}(\mathbf{k}_1, \mathbf{k}_2) \Psi_c(\mathbf{k}_1, \eta') \Psi_d(\mathbf{k}_2, \eta'). \end{aligned} \quad (91)$$

This expression depends on the linear propagator $g_{ab}(\eta)$, which solves the linearized equations of motion [i.e., setting the right-hand side of Eq. (87) to zero] and presents a mixture of growing and decaying, as well as time independent modes [9,85]:

$$g_{ab}(\eta) = \frac{e^\eta}{5} \begin{bmatrix} 3 & 2 & 0 \\ 3 & 2 & 0 \\ 3 & 2 & 0 \end{bmatrix} - \frac{e^{-3\eta/2}}{5} \begin{bmatrix} -2 & 2 & 0 \\ 3 & -3 & 0 \\ -2 & 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}. \quad (92)$$

As we did for the galaxy overdensity in Sec. IV A, we now expand $\Psi(\mathbf{k}, \eta)$ in terms of generalized Wiener-Hermite functionals,

$$\Psi_a(\mathbf{k}, \eta) = \sum_n \frac{(2\pi)^3}{n!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_n} [\delta_D]_n \Gamma_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta) \times \mathcal{H}_n^*(\mathbf{k}_1, \dots, \mathbf{k}_n), \quad (93)$$

where $\Gamma_3^{(n)} \equiv \Gamma_g^{(n)}$. An equivalent expansion holds at initial time $\eta = 0$, and we denote the corresponding propagators by the symbol $\Gamma_{\mathcal{L}}$. Since the dark matter and velocity fields are linear at that time, we have $\Gamma_{1,\mathcal{L}}^{(1)} = 1 = \Gamma_{2,\mathcal{L}}^{(1)}$, while all higher-order propagators must vanish, i.e.,

$$\Gamma_{a,\mathcal{L}}^{(n)} = \delta_{n,1}^K \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \Gamma_{g,\mathcal{L}}^{(n)} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (94)$$

The initial galaxy propagators, on the other hand, are given by the expressions from Sec. IV B. As mentioned in the introduction to Sec. III, it is here where one might take advantage of some simplifying assumptions, for instance, removing fourth-order NLE operators from the initial conditions. This sets

$$\gamma_{21,\mathcal{L}}^\times = \gamma_{31,\mathcal{L}} = \gamma_{22,\mathcal{L}} = \gamma_{211,\mathcal{L}} = 0 \quad (95)$$

in Eqs. (84) and (85), which does not imply that we ignore such NLE operators completely because, as demonstrated by these equations, time evolution will generate amplitudes $\gamma_{21}^\times, \gamma_{31}, \gamma_{22}, \gamma_{211}$ fixed through lower-order bias parameters (note this approximation is more general than the local Lagrangian approximation, which sets *all* nonlocal in δ Lagrangian bias parameters to zero).

By multiplying both sides of Eq. (91) with $\mathcal{H}_n$ and taking the ensemble average, we can derive a recursion relation for the time evolved multipoint propagators. Following the steps detailed in Appendix B, we arrive at the expression

$$\begin{aligned} \Gamma_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta) &= g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) \\ &+ \sum_{r=0}^{n+r-\delta_{r,0}^K} \sum_{m=\max\{r,1\}} \frac{1}{r!} \int_0^\eta d\eta' g_{ab}(\eta-\eta') \Gamma_b^{(n,m,r)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta'), \end{aligned} \quad (96)$$

where the quantity $\Gamma_a^{(n,m,r)}$ represents r loop integrals over propagators of orders m and $n-m+2r$,

$$\begin{aligned} \Gamma_a^{(n,m,r)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta') &\equiv \int_{\mathbf{q}_1, \dots, \mathbf{q}_r} \left[\gamma_{abc}(\mathbf{k}_{1\dots m-r} + \mathbf{q}_{1\dots r}, \mathbf{k}_{m-r+1\dots n} - \mathbf{q}_{1\dots r}) \right. \\ &\times \Gamma_b^{(m)}(\mathbf{k}_1, \dots, \mathbf{k}_{m-r}, \mathbf{q}_1, \dots, \mathbf{q}_r; \eta') \\ &\times \Gamma_c^{(n-m+2r)}(\mathbf{k}_{m-r+1}, \dots, \mathbf{k}_n, -\mathbf{q}_1, \dots, -\mathbf{q}_r; \eta') \\ &\times \prod_{i=1}^r P_L(\mathbf{q}_i) + \text{sym}(\mathbf{k}_i) \Big], \end{aligned} \quad (97)$$

with $\text{sym}(\mathbf{k}_i)$ standing for the symmetrization over all possibilities of building a subset of $m-r$ $\mathbf{k}$ modes from a total group of n , i.e., $\binom{n}{m-r}$ terms.

Furthermore, Eq. (97) illustrates the point we made at the beginning of this section—the time evolved multipoint propagators are free of potentially divergent contributions proportional to σ^2 , which means we do not have to perform any additional renormalization steps. This is a consequence of the scale dependence of the mode coupling kernels γ_{abc} , which contribute to $\Gamma_g^{(n)}$ only via the symmetrized combination $\alpha^s(\mathbf{k}_1, \mathbf{k}_2) = 1/2[\alpha(\mathbf{k}_1, \mathbf{k}_2) + \alpha(\mathbf{k}_2, \mathbf{k}_1)]$ coming from the continuity equation (conservation of tracers).

Plugging in the wave vectors from Eq. (97) and expanding in inverse powers of $q_{1\dots r}$, we find that to leading order

$$\alpha^s(\mathbf{K} - \mathbf{q}_{1\dots r}, \mathbf{k}_{1\dots n} - \mathbf{K} - \mathbf{q}_{1\dots r}) \sim \left(\frac{k_{1\dots n}}{q_{1\dots r}} \right)^2 + \dots, \quad (98)$$

where $\mathbf{K} = \mathbf{k}_{1\dots m-r}$. In this paper we are interested in corrections only up to the one-loop level, so we are led to consider expressions of the form

$$\begin{aligned} &\sim k_{1\dots n}^2 \int_{\mathbf{q}} \frac{P_L(\mathbf{q})}{q^2} \Gamma^{(m)}(\mathbf{k}_1, \dots, \mathbf{k}_{m-1}, \mathbf{q}) \\ &\times \Gamma^{(n-m+2)}(\mathbf{k}_m, \dots, \mathbf{k}_n, -\mathbf{q}), \end{aligned} \quad (99)$$

where both propagators are evaluated at tree level. That means the propagators remain finite when the loop momentum q becomes large, which in turn implies that the overall integrand scales, at most, as $1/q^3$ in this limit [using that

$P_L(q) \sim 1/q^3$ for $q \rightarrow \infty$]. This guarantees that any integral of the above type is quickly convergent.

For higher than one-loop corrections this argument no longer holds, as each loop adds an additional power spectrum to the integrand, while the scaling with the loop momenta remains the same. However, the scale dependence on the external momenta, i.e., $\sim k_{1\dots n}^2$, suggests that these terms can be absorbed by redefinitions of the higher-derivative bias parameters, which display the exact same scale dependence, as shown in Sec. III D. Using higher-derivative terms to absorb sensitivities to the highly nonlinear regime was already presented in a slightly different manner in [86,87]. We note that this behavior of the loop integrals is well known in the context of matter perturbations [43,88–91], where the small scale sensitivity can be understood in terms of nonzero stress-tensor corrections. Indeed, we will show in Sec. V B that these are completely degenerate with the contributions from higher-derivative bias.

Using Eq. (96) it is straightforward to compute the evolved multipoint propagators by starting from the lowest order at tree level and constructing all higher-order solutions recursively. One part of the solution is due to the evolution of the bias parameters and can be simply obtained by replacing the Lagrangian parameters in the initial multipoint propagators by their Eulerian analogs (according to the relations given in Sec. IV C 1). Here we focus on the remaining corrections from nonlinear evolution, which we write as $\Delta\Gamma_g^{(n)}$, and neglecting all but the fastest growing mode and expressing the result back in terms of the SPT kernels, we find that these corrections for the one-point propagator at one-loop ($r = 1$) are given by

$$\Delta\Gamma_g^{(1)}|_{1\text{-loop}} = 3b_1 \int_{\mathbf{q}} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_L(q) + 4\gamma_2 \int_{\mathbf{q}} K(\mathbf{k} - \mathbf{q}, \mathbf{q}) G_2(\mathbf{k}, -\mathbf{q}) P_L(q). \quad (100)$$

The two- and three-point propagators are already affected at tree level by nonlinear evolution, and the resulting corrections are

$$\Delta\Gamma_g^{(2)}|_{\text{tree}} = 2b_1 F_2(\mathbf{k}_1, \mathbf{k}_2), \quad (101)$$

$$\Delta\Gamma_g^{(3)}|_{\text{tree}} = 6b_1 F_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + 2[b_2 F_2(\mathbf{k}_1, \mathbf{k}_2) + 2\gamma_2 K(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3) G_2(\mathbf{k}_1, \mathbf{k}_2) + \text{cyc}]. \quad (102)$$

For the one-loop bispectrum we also require nonlinear corrections of $\Gamma_g^{(2)}$ at one-loop order, which are, however, not easily expressed in terms of SPT kernels, so we give the full result in terms of the initial multipoint propagators in Appendix B 3.

V. POWER SPECTRUM AND BISPECTRUM

This section details the final step of this paper: the computation of the galaxy power spectrum and bispectrum in terms of the multipoint propagators. In addition, we focus on residual sensitivities of the loop corrections on the highly nonlinear regime, where our perturbative approach breaks down. These affect the statistics not only on small scales, but also on asymptotically large scales, and we discuss how they can be regularized by the addition of physically motivated terms.

A. Reconstructing correlators from multipoint propagators

The multipoint propagators serve as the basic building blocks for computing N -point spectra. This follows easily from the orthogonality relations of the generalized Wiener-Hermite functionals (see Appendix B 1), and was already shown for the dark matter density in [41], whose results we can directly apply to the present case of galaxy clustering.

In particular, by evaluating $\langle \Psi_a(\mathbf{k}) \Psi_b(\mathbf{k}') \rangle$ and using Eq. (57) one finds that the power spectrum is given by a series of two contracted multipoint propagators of the same order. Diagrammatically this can be represented by gluing together two of the objects shown in Fig. 1, where each combination of the incoming lines gives rise to a (linear) power spectrum. This is demonstrated in Fig. 2, and as the shaded circles include all vertex loop corrections (i.e., vertex renormalizations) we only need to consider one distinct diagram for the power spectrum at one-loop level, compared to the usual two in the standard treatment. The galaxy power spectrum is thus given by [41]

$$P_g(k) = [\Gamma_g^{(1)}(k)]^2 P_L(k) + \frac{1}{2} \int_{\mathbf{q}} [\Gamma_g^{(2)}(\mathbf{k} - \mathbf{q}, \mathbf{q})]^2 \times P_L(|\mathbf{k} - \mathbf{q}|) P_L(q), \quad (103)$$

where $\Gamma_g^{(1)}$ is evaluated up to one-loop order, while tree-level terms are sufficient for $\Gamma_g^{(2)}$.

We proceed in a similar manner for the bispectrum, which is obtained from $\langle \Psi_a(\mathbf{k}_1) \Psi_b(\mathbf{k}_2) \Psi_c(\mathbf{k}_3) \rangle$

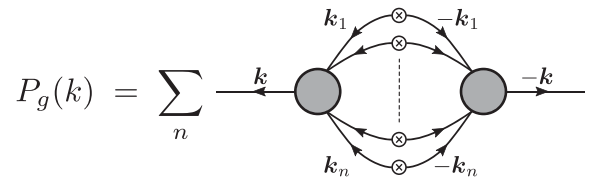


FIG. 2. Galaxy power spectrum, reconstructed from multipoint propagators, which are represented by the shaded circles with incoming and outgoing momentum $\mathbf{k}$. The sum runs over the number of connected internal lines, each of which produces a linear power spectrum depicted by a crossed circle.

and application of Eq. (B12). The complete solution can be found in [41], but its diagrammatic depiction in Fig. 3 is straightforward—a combination of three multipoint propagators with a varying number of connecting lines between

each pair (maximally one such pair is allowed to be disconnected to avoid an overall disconnected graph). Taking care of the appropriate symmetry factors that arise in these various combinations, it follows that [41]

$$B_g(k_1, k_2, k_3) = \Gamma_g^{(2)}(\mathbf{k}_1, \mathbf{k}_2) \Gamma_g^{(1)}(k_1) \Gamma_g^{(1)}(k_2) P_L(k_1) P_L(k_2) + \text{cyc} \\ + \left[\int_q \Gamma_g^{(2)}(\mathbf{k}_1 - \mathbf{q}, \mathbf{q}) \Gamma_g^{(2)}(\mathbf{k}_2 + \mathbf{q}, -\mathbf{q}) \Gamma_g^{(2)}(\mathbf{k}_1 - \mathbf{q}, \mathbf{k}_2 + \mathbf{q}) P_L(|\mathbf{k}_1 - \mathbf{q}|) P_L(|\mathbf{k}_2 + \mathbf{q}|) P_L(q) \right. \\ \left. + \frac{1}{2} \int_q \Gamma_g^{(3)}(\mathbf{k}_3, \mathbf{k}_2 - \mathbf{q}, \mathbf{q}) \Gamma_g^{(2)}(\mathbf{k}_2 - \mathbf{q}, \mathbf{q}) \Gamma_g^{(1)}(k_3) P_L(|\mathbf{k}_2 - \mathbf{q}|) P_L(q) P_L(k_3) + \text{cyc} \right]. \quad (104)$$

Therefore, for the one-loop galaxy bispectrum we require both, $\Gamma_g^{(1)}$ and $\Gamma_g^{(2)}$, up to one-loop order in the first term of Eq. (104), but tree-level expressions for them and $\Gamma_g^{(3)}$ are enough in the loop integrals, i.e., in the second and third lines. Note that there are only two one-loop diagrams instead of the four in SPT; this is because there are two diagrams in SPT that are reducible, and thus incorporated into the one-loop $\Gamma_g^{(1)}$ (giving the 321-II diagram in the notation of [92]) and the one-loop $\Gamma_g^{(2)}$ diagram (giving the 411 diagram).

From comparing Eqs. (103) and (104) we note that the two-point propagator contributes to the leading-order bispectrum, while showing up as a loop correction for the power spectrum. This structure extends to consecutively higher orders, for instance, the three-point propagator, which appears as a one-loop expression in the bispectrum, will enter at tree level for the trispectrum. That suggests that constraints on bias parameters required to fit the small-scale behavior of a given correlator (and thus cosmological parameters, too) will already highly benefit from the large-scale information of the next order correlator.

When splitting each expression in Eq. (104) into its individual contributions, one finds that in total there are 40 additional terms in the one-loop galaxy bispectrum compared to tree level. Each of these terms is multiplied by a

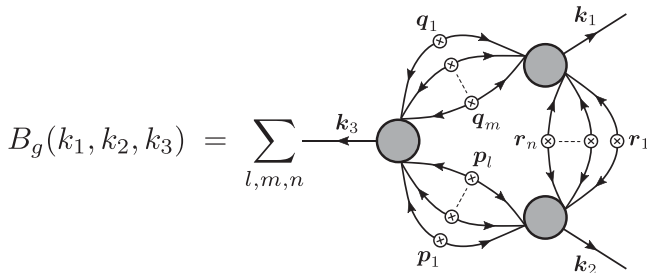


FIG. 3. Galaxy bispectrum, expressed through multipoint propagators. The sum runs over the number of connected internal lines of each pair of propagators (shaded circles). At most one of the three indices can be zero, so that the overall diagram remains a connected graph.

combination of various bias parameters, making a comparison of the individual terms not particularly meaningful. However, given a set of bias parameters, it is interesting to consider whether the bispectrum loop contributions become relevant on similar scales as for the power spectrum. To that end, in Fig. 4 we show a comparison of the one-loop galaxy power spectrum evaluated from Eq. (103) and its tree-level prediction $P_g^{\text{tree}} = b_1^2 P_L$, where different colors correspond to a different choice of the linear bias parameter. In order to adopt representative values for the higher-order bias parameters, we fix them in terms of b_1 by making use of the local Lagrangian approximation [obtained by setting all parameters with subscript “ $\mathcal{L}$ ” in Eqs. (81)–(85) to zero], as well as the peak-background split relations for b_2 and b_3 .⁵ These were calibrated using separate universe simulations in [93], yielding

$$b_2(b_1) = 0.412 - 2.143b_1 + 0.929b_1^2 + 0.008b_1^3, \quad (105)$$

$$b_3(b_1) = -1.028 + 7.646b_1 - 6.227b_1^2 + 0.912b_1^3. \quad (106)$$

As is demonstrated by the lower panel of Fig. 4, the one-loop corrections to the galaxy power spectrum become relevant on scales $k \gtrsim 0.1 h/\text{Mpc}$. A similar trend can be observed for the galaxy bispectrum, whose relative difference between the tree-level and one-loop models for $b_1 = 1.8$ is plotted in Fig. 5. To show its configuration dependence we evaluate the bispectrum as a function of $x_2 = k_2/k_1$ and $x_3 = k_3/k_1$, averaged over a thin shell centered at a given value of k_1 , i.e.,

$$\bar{B}_g(k_1, x_2, x_3) = \frac{1}{\Delta k} \int_{k_1 - \Delta k/2}^{k_1 + \Delta k/2} dq B_g(q, x_2 q, x_3 q), \quad (107)$$

⁵Note that the power spectrum and bispectrum models in Figs. 4 and 5 include the stochastic corrections in the low- k limit to be discussed in Sec. V C, but with all noise parameters set to zero.

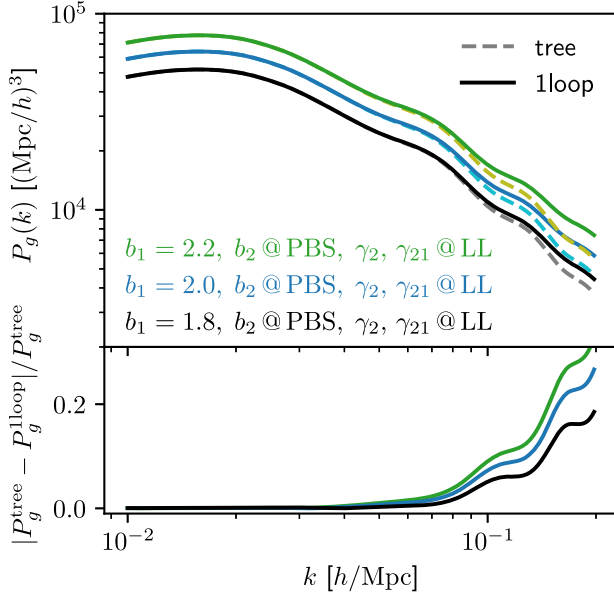


FIG. 4. Top panel: comparison of the tree-level and one-loop galaxy power spectrum for three different values of the linear bias parameter. Second- and third-order bias parameters are fixed in terms of b_1 by means of the peak-background split (PBS) and local Lagrangian (LL) predictions. Bottom panel: relative difference between the tree-level and one-loop models for the same three cases.

where we have chosen $\Delta k = 0.02 h/\text{Mpc}$ and three different values for k_1 (see Fig. 5). We notice that for all triangle configurations the relative difference surpasses 10% for $k_1 \gtrsim 0.1 h/\text{Mpc}$, with the strongest deviations to be found for equilateral triangles. Of course, the precise numerical values obtained in this demonstration depend sensitively on the set of bias parameters, but we do not expect any big impact on the overall conclusion: bias loop corrections affect the power spectrum and bispectrum starting from comparable scales.

B. Stress-tensor corrections and their degeneracy with higher-derivative bias

While the introduction of multipoint propagators has automatically removed potentially divergent contributions proportional to σ^2 [i.e., $(k/q)^0$ with k the external momentum and q the loop momentum], the loop integrals remain sensitive to the nonlinear regime through terms scaling as powers of $(k/q)^2$ [and $(k/q)^4$, etc.]. This is true for loop corrections of individual multipoint propagators, as discussed in Sec. IV B and at the end of Sec. IV C, but also for the loop integrals over tree-level propagators appearing in Eqs. (103) and (104). Sensitivity of loop integrals to the nonlinear regime is obviously a problem since perturbation theory does not hold at small scales.

At such small scales the dark matter field can no longer be treated as a pressureless perfect fluid, as is assumed in SPT. Initially, or on large scales, this is a good approximation as dark matter particles tend to move within single coherent flows, which implies a vanishing stress tensor σ_{ij} . At later times, multistreaming induces nonzero stresses, which can have an impact on quasilinear scales, with the same $(k/q)^2$ scaling as one-loop SPT and one-loop bias terms [32].

These stress-tensor corrections have been computed in the framework of the effective field theory (EFT) [34–37] and directly from the Vlasov equation [94] following [32,95]. This leads to additional terms in the power spectrum and bispectrum that at lowest-order scale with powers of k^2 , identically to one-loop corrections sensitive to the nonlinear regime. In the particularly economic parametrization of [94] (which is otherwise equivalent to the EFT calculations mentioned above) they read

$$P_\sigma(k) = -2\beta_P k^2 P_L(k), \quad (108)$$

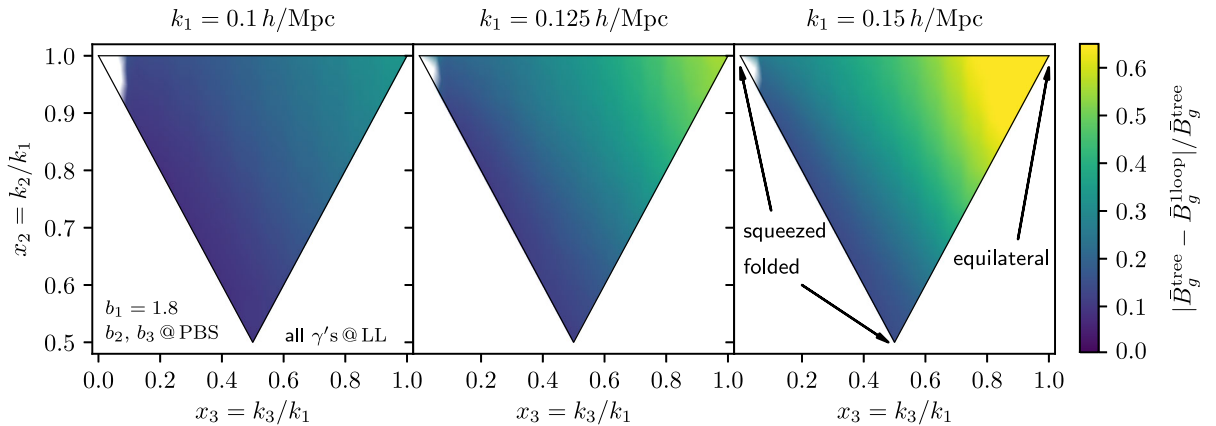


FIG. 5. Relative difference between the tree-level and one-loop galaxy bispectrum as a function of $x_2 = k_2/k_1$ and $x_3 = k_3/k_1$, integrated over a thin shell centered on three different values of k_1 . The linear bias parameter is given by $b_1 = 1.8$, while b_2 and b_3 are determined from the PBS relations, and all other parameters are fixed using the LL approximation. Note that Eq. (106) was evaluated using a grid of bispectra with fixed binning for k_1 , k_2 , and k_3 , leading to an absence of very squeezed configurations.

$$\begin{aligned}
B_{\sigma,123} = & -[(\beta_{B,a}(k_1^2 + k_2^2) + \beta_{B,b}k_3^2)F_2(\mathbf{k}_1, \mathbf{k}_2) \\
& + (\beta_{B,c}(k_1^2 + k_2^2) + \beta_{B,d}k_3^2)K(\mathbf{k}_1, \mathbf{k}_2)] \\
& \times P_L(k_1)P_L(k_2) + \text{cyc}, \quad (109)
\end{aligned}$$

where β_P and $\beta_{B,i}$ are numbers that result from integrating over stress-tensor components weighted by functions of time (growth factors). Since the time dependence of the stress tensor is not calculable in a generic Λ CDM universe, these numbers are free parameters. Crucially though, they can absorb and regularize all contributions from SPT one-loop integrals that are sensitive to a range of scales where our perturbative approach breaks down and have the same k^2 scaling.

As we have seen in Sec. III D, the higher-derivative bias terms display the same momentum scaling and thus give rise to very similar terms. Ignoring all precoefficients, evaluation of Eqs. (103) and (104) shows that the higher-derivative contribution to the power spectrum is exactly degenerate with Eq. (108), while for the bispectrum we get five different terms:

$$\left. \begin{aligned}
& \text{(i) } k_3^2 \\
& \text{(ii) } \mathbf{k}_1 \cdot \mathbf{k}_2 \\
& \text{(iii) } k_3^2 F_2(\mathbf{k}_1, \mathbf{k}_2) \\
& \text{(iv) } k_3^2 K(\mathbf{k}_1, \mathbf{k}_2) \\
& \text{(v) } (\mathbf{k}_1 \cdot \mathbf{k}_2) K(\mathbf{k}_1, \mathbf{k}_2)
\end{aligned} \right\} \times P_L(k_1)P_L(k_2) + \text{cyc} \quad (110)$$

The contributions (iii) and (iv) are clearly degenerate with terms in Eq. (109), and using that $\mathbf{k}_1 \cdot \mathbf{k}_2 = 1/2(k_3^2 - k_1^2 - k_2^2)$ we see that (v) can be written as a combination where $\beta_{B,d} = 1/2 = -\beta_{B,c}$. Furthermore, one can show that

$$\begin{aligned}
k_3^2 = & -[k_1^2 + k_2^2 - k_3^2]F_2(\mathbf{k}_1, \mathbf{k}_2) \\
& - \left[\frac{5}{7}(k_1^2 + k_2^2) + \frac{2}{7}k_3^2 \right] K(\mathbf{k}_1, \mathbf{k}_2), \quad (111)
\end{aligned}$$

such that a combination of all four terms in Eq. (109) with $\beta_{B,a} = 1$, $\beta_{B,b} = -1$, $\beta_{B,c} = 5/7$, and $\beta_{B,d} = 2/7$ can also accommodate for (i). Only (ii) cannot be expressed through the stress-tensor terms and must consequently enter the galaxy bispectrum as an independent contribution. In total, this demonstrates that the stress-tensor corrections to the bispectrum are also completely degenerate with those from higher-derivative bias. For that reason we consider them collectively, using the following basis:

$$\begin{aligned}
B_{\sigma+\nabla,123} = & -\{[\beta_{B,a}(k_1^2 + k_2^2) + \beta_{B,b}k_3^2]F_2(\mathbf{k}_1, \mathbf{k}_2) \\
& + [\beta_{B,c}(k_1^2 + k_2^2) + \beta_{B,d}k_3^2]K(\mathbf{k}_1, \mathbf{k}_2) \\
& + \beta_{B,e}\mathbf{k}_1 \cdot \mathbf{k}_2\}P_L(k_1)P_L(k_2) + \text{cyc}, \quad (112)
\end{aligned}$$

which reduces the number of free parameters to five, in addition to another one for the power spectrum. Note that if we had considered only higher-derivative bias, then we would have only five parameters in total for the power spectrum and bispectrum, since in that case $2\beta_P = \beta_{B,a} + \beta_{B,b}$. However, that is not the case for stress-tensor contributions since β_P and β_B 's result from integrating the stress tensor weighted with different powers of the growth factor [94]. In addition, it is worth noting that some of the contributions of velocity bias are also degenerate with higher-derivative bias (see [96]). Finally, note that the contributions in Eq. (112) are degenerate with one-loop galaxy bias contributions in the low- k limit; see Appendix A 3 for an explicit discussion of this.

C. Stochastic contributions

The effects discussed in the last section become relevant toward smaller scales, but the bias loop corrections can also have an impact on the large-scale power spectrum and bispectrum. As already pointed out in [7,97], that is because a subset of the bias loop integrals do not vanish in the large-scale limit, and thus come to dominate for values of k^{-1} above a certain scale. Taking the limit $k \rightarrow 0$ of Eq. (103) shows that the galaxy power spectrum approaches a constant,

$$\begin{aligned}
\lim_{k \rightarrow 0} P_g(k) &= \frac{1}{2} \int_q [\Gamma_g^{(2)}(-\mathbf{q}, \mathbf{q})]^2 P_L(q)^2 \\
&= \frac{b_2^2}{2} \int_q P_L(q)^2, \quad (113)
\end{aligned}$$

where we have assumed that the linear power spectrum falls off to zero on large scales and used that $K(-\mathbf{q}, \mathbf{q}) = 0$. Crucially, corrections from successively higher orders of SPT add in comparable measures to this large-scale limit, meaning that this low- k limit is not controlled by any particular order in perturbation theory, and thus one must introduce a new parameter that describes its size.

The same extends to the loop corrections of the galaxy bispectrum, and in the limit of all three triangle sides approaching zero, Eq. (104) equally becomes constant,

$$\begin{aligned}
\lim_{k_1, k_2 \rightarrow 0} B_g(k_1, k_2, k_3) &= \int_q [\Gamma_g^{(2)}(-\mathbf{q}, \mathbf{q})]^3 P_L(q)^3 \\
&= b_2^3 \int_q P_L(q)^3. \quad (114)
\end{aligned}$$

As for the power spectrum the value of this constant receives corrections from higher-order SPT terms, which are shown schematically up to the three-loop level by the diagrams in Fig. 6. Although diagrams that contain a one-point propagator at one of their external legs do not contribute to this limit, they are affected by large-scale loop corrections in a similar manner. The corresponding terms can be identified as follows:

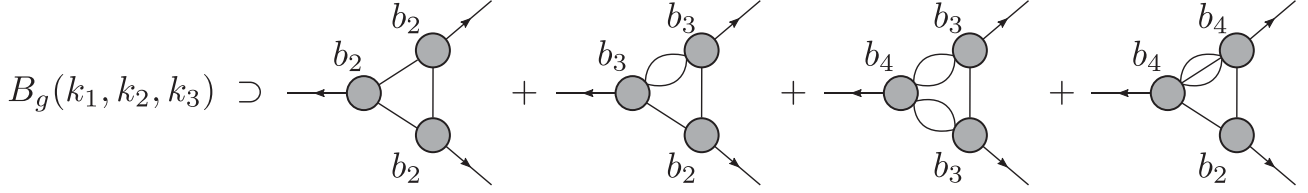


FIG. 6. Subset of diagrams that contribute to a nonzero large-scale limit ($k_1, k_2, k_3 \rightarrow 0$) of the galaxy bispectrum. The first diagram appears at one-loop level, the second at two-loop, and the last two at three-loop. The bias constants b_i indicate the value of the multipoint propagators (shaded circles), i.e., constants that do not vanish in the large-scale limit.

$$\lim_{k_1, k_2 \rightarrow 0} \frac{B_g(k_1, k_2, k_3)|_{\propto P_L(k_1)}}{P_L(k_1)} = b_1 b_2 \left[\frac{115}{42} b_2 + b_3 - \frac{8}{3} \gamma_2^\times \right] \int_q P_L(q)^2, \quad (115)$$

where $B_g(k_1, k_2, k_3)|_{\propto P_L(k_1)}$ denotes all terms in the galaxy bispectrum that are proportional to $P_L(k_1)$. After cyclic permutations their total contribution to the one-loop bispectrum is thus given by

$$b_1 b_2 \left(\frac{115}{42} b_2 + b_3 - \frac{8}{3} \gamma_2^\times \right) \times [P_L(k_1) + P_L(k_2) + P_L(k_3)] \int_q P_L(q)^2. \quad (116)$$

These effects can be interpreted as the impact of small-scale perturbations on the formation of galaxies that cannot be captured by any perturbative bias model, and were originally introduced as a “stochastic” bias in the literature [38,39]. Its defining property is that it must be mostly uncorrelated on large scales (assuming Gaussian initial conditions as we do), meaning it will manifest in the same way as shot noise. Inspecting Eq. (113) and Eqs. (114) and (116), we notice that this is indeed the case for the large-scale limit of the one-loop power spectrum and bispectrum. We could have chosen to incorporate these effects by including stochastic fields in the expansion for δ_g (see, for instance, [6,98,99]), instead we now introduce *a posteriori* three effective shot noise parameters [7,8] for the power spectrum and bispectrum, such that

$$P_{\text{shot}}(k) = N_0, \quad (117)$$

$$B_{\text{shot}}(k_1, k_2, k_3) = \epsilon_0 + \eta_0 [P_L(k_1) + P_L(k_2) + P_L(k_3)]. \quad (118)$$

Similar to the stress-tensor parameters, these parameters are able to absorb any residual large-scale contributions stemming from the bias loops. We stress that the values of N_0 , ϵ_0 , and η_0 are typically not given by their Poissonian shot noise predictions, i.e., $N_0 = 1/\bar{n}$, $\epsilon_0 = 1/\bar{n}^2$, and $\eta_0 = 1/\bar{n}$

[100] for an average number density of galaxies $\bar{n}$, but must be determined from the data itself.

Finally, in order to take into account a slight correlation of the stochastic bias on large scales, we can think of Taylor expanding its contributions in powers of k^2 , where in general the expansion coefficients must also be considered as free parameters. This is motivated by requiring that not only the leading-order terms in the large-scale limits above can be absorbed but also their next-to-leading order (NLO) contributions. These terms exclusively scale as powers of k^2 (or k_1^2 , k_2^2 , and k_3^2 in case of the bispectrum), and it is straightforward to show that the allowed terms can be summarized as follows:

$$P_{\text{shot,NLO}}(k) = N_2 k^2, \quad (119)$$

$$B_{\text{shot,NLO}}(k_1, k_2, k_3) = \epsilon_2 (k_1^2 + k_2^2 + k_3^2) + [\eta_{2,1} k_1^2 + \eta_{2,2} (k_2^2 + k_3^2)] P(k_1) + \text{cyc}, \quad (120)$$

which leads to an additional four parameters.

VI. DISCUSSION AND CONCLUSIONS

In this paper, we presented the complete set of one-loop (next-to-leading order) galaxy bias corrections for the real-space galaxy bispectrum. These corrections serve to increase the range of scales where the bispectrum can be robustly used to extract information from the clustering of galaxies (or any other tracer), and brings the bispectrum to the same state-of-the-art as the galaxy power spectrum, meaning joint analyses can be performed consistently. We carry this out in detail and compare against numerical simulations of biased tracers in a follow-up paper [76].

Our perturbative bias model systematically combines a variety of effects that become relevant on scales of the weakly nonlinear regime. These effects include contributions from the following: (1) the general bias expansion, generated by gravitational evolution of the matter field; (2) higher-derivative terms, due to the spatial nonlocality of galaxy formation, and a nonzero stress tensor; and (3) stochasticity, resulting from the impact of nonperturbative physics on the large-scale galaxy distribution.

TABLE II. Number of bias parameters required for the galaxy power spectrum and bispectrum up to one-loop order in perturbation theory, organized according to the various contributions to the bias expansion. Of the 11 general bias parameters for the $P_g + B_g$ description, 4 are enough to describe P_g alone ($b_1, b_2, \gamma_2, \gamma_{21}$), while adding B_g requires 7 more parameters ($\gamma_2^\times, \gamma_{21}^\times, b_3, \gamma_3, \gamma_{22}, \gamma_{31}, \gamma_{211}$).

	General bias expansion	Stochastic		Stress tensor and higher derivative
		$\mathcal{O}(k^0)$	$\mathcal{O}(k^2)$	
P_g	11	1	1	1
B_g		2	3	5

In all generality, we found that a joint one-loop power spectrum and bispectrum analysis in real-space requires 24 free parameters. In agreement with [12], we showed that 11 parameters are due to the general bias expansion, which are shared among the power spectrum and bispectrum. They include the usual linear bias parameter at first order (b_1), and the local and nonlocal quadratic parameters at second (b_2, γ_2). To account for higher-derivative effects, which were not considered by [12], we demonstrated that we need to introduce an additional five parameters for the bispectrum and one for the power spectrum, five less than given in [6]. We also found that the stress-tensor corrections are entirely degenerate with those from higher derivatives and therefore do not consider them separately. Finally, at leading order stochasticity contributes with two terms (one term) to the bispectrum (power spectrum), whereas at next-to-leading order we identified a further three (one). An overview of these numbers is given in Table II.

Obviously, it would be bad news if a joint analysis of power spectrum and bispectrum demands that each of the 24 parameters enters the model with a freely adjustable amplitude. In that case the gain from including smaller scales in the analysis could easily be canceled out by the loss in constraining power from having to marginalize over so many nuisance parameters. From that point of view the work presented in this paper is best regarded as the theory on which we can analyze various ways of reducing the parameter space, as will be discussed in detail in a follow-up paper [76]. For example, it is known in the case of the one-loop galaxy power spectrum that higher derivatives are degenerate with bias loops [8,24,101,102], and as now shown here (see Appendix A 3) the structure of one-loop bias at low- k is also degenerate with higher-derivative bias for the bispectrum. However, it remains to be seen how the higher-derivative terms compare to those from the general bias expansion. While [30,31] suggest that the former dominate for tracers residing in very massive halos, this is not necessarily correct for certain types of tracers that are targeted by upcoming galaxy surveys such as DESI. This includes, for instance, emission line galaxies, which are usually found in less massive halos.

Apart from testing the significance of the higher-derivative contributions, another clear possibility of reducing the parameter space is to check which bias operators determine the galaxy perturbations at the initial time. Our bias basis is designed so that we separate operators that are local in second derivatives of the linear potential (what we call local evolution operators) and those that are nonlocal (induced by nonlocal evolution operators) that involve the nonlinear Lagrangian potentials. Other bases in the literature mix these properties; see Appendix A 2.

The most drastic reduction in parameter space is achieved by initially putting all NLE operators to zero and all LE operators nonlocal in δ to zero, this is the local Lagrangian approximation: it reduces the 11 parameters in Table II to just 3 free parameters (b_1, b_2, b_3), and the other eight parameters get determined from them by the time evolution arguments in Sec. IV C 1 (putting all of the eight Lagrangian values to zero). A hierarchy of more accurate approximations (but with more free parameters) can be made where one takes, e.g., only the NLE operators to be zero in the far past (as in peak or excursion set models of bias), or just the fourth-order ones that only enter through the loop correction of a single diagram in the bispectrum. The validity of such assumptions must be tested numerically and compared to the full model, which we will present in [76]. For practical applications to galaxy surveys, given finite error bars, one should be able to reduce the parameter space sequentially going from most general to least general checking the final results are not impacted by the assumptions.

On the more technical side, the central concept explored in this work is the use of multipoint propagators, which correspond to the sum over all reducible diagrams, as a way to writing the perturbative expansion in terms of explicit observables (that corresponds for Gaussian fluctuations to the commonly measured cross-correlation bias between galaxies and matter). As a result of this reformulation we were able to circumvent the renormalization procedure—a tedious redefinition of the standard bias parameters required to absorb diverging contributions in the computation of correlation functions at one-loop and beyond. To achieve this we showed that it is best to proceed in two steps: first, compute the multipoint propagators at initial time where nonlinearities in the matter field can be ignored, and then evolve these initial propagators to the time of observation. Evolving the propagators conserving the number of tracers (the so-called “coevolution” first given by [69]) cannot generate any diverging terms proportional to $\sigma^2 = \langle \delta(x)^2 \rangle$, and therefore the late-time propagators are already renormalized (and any k^2 or higher-order renormalizations are simple to handle by adding the corresponding terms). Once the evolved multipoint propagators have been computed, they serve as simple building blocks for the general N -point correlation function.

We also discussed how our choice of basis, written in terms of Galileons and with a clear split between local and nonlocal functions of second derivatives of the linear potential, simplifies the calculation of the multipoint propagators: they are given by the tree-level expressions plus loop integrals over the NLE operator contributions only. Other bases that mix our operators (see Appendix A 2) complicate the calculation, as one has to separate terms that get absorbed by the renormalization procedure.

A significant part of the evolution of the propagators is absorbed into the evolution of the bias parameters in them. Our calculation of this time evolution was carried out using a new approach that takes full advantage of Galilean invariance (see Appendix C), and we present results for the fourth-order parameters for the first time (see Sec. IV C 1), generalizing the third-order results originally given in [9]. Up to quadratic bias we also show explicitly that renormalization by fourth-order operators does not change the time evolution of bias parameters obtained at tree level (see Appendix D), as expected.

Finally, to summarize the steps necessary for applying the results of this paper in practice, one would proceed as follows: first, one computes the Lagrangian multipoint propagators given in Eqs. (66), (71), and (75). These serve as initial conditions for the evolution equations yielding the Eulerian multipoint propagators, i.e., Eqs. (B23)–(B26), which in turn determine the power spectrum and bispectrum at the time of observation via Eqs. (103) and (104). The final expressions are obtained after adding the combination of stress-tensor and higher-derivative corrections from Eqs. (108) and (112), as well as the stochastic contributions up to next-to-leading order [Eqs. (117)–(120)].

This will be demonstrated explicitly in our follow-up paper [76]. In particular, we carry out a detailed likelihood analysis of the power spectrum and bispectrum in numerical simulations of biased tracers to ascertain the performance of the results derived here, the importance of the different terms in the bias expansion, the accuracy of theoretical assumptions that reduce the number of free parameters, and the improvement of our results compared to current models in the literature.

ACKNOWLEDGMENTS

We thank M. Crocce, A. Sanchez, and R. Sheth for useful discussions. A. E. acknowledges support from the UK Science and Technology Facilities Council (STFC) via Research Training Grant (No. ST/M503836/1), as well as from the European Research Council (Grant No. ERC-StG-716532-PUNCA). A. E. also thanks the Center for Cosmology and Particle Physics of New York University for hospitality while part of this work was undertaken, and the STFC for supporting this research stay. R. E. S. acknowledges support from the STFC (Grant No. ST/P000525/1).

APPENDIX A: FURTHER NOTES ON GALILEAN BASIS FOR GALAXY BIAS

1. Basis operators in Fourier space

Here we briefly summarize Fourier space expressions for our basis operators given in Table I, which are being used in the computation of the multipoint propagators. In general, we write any n th-order operator $\mathcal{O}^{(n)}$ as an integral over n linear matter perturbations,

$$\mathcal{O}_B^{(n)}(\mathbf{k}) = (2\pi)^3 \int_{\mathbf{k}_1, \dots, \mathbf{k}_n} [\delta_D]_n \mathcal{K}_B^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) \prod_{i=1}^n \delta_L(\mathbf{k}_i), \quad (\text{A1})$$

where B stands for any of the basis operators at that order and we use the notation $[\delta_D]_n \equiv \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n})$. For $n = 2$ we have from Eq. (32)

$$\mathcal{K}_{\mathcal{G}^2}^{(2)}(\mathbf{k}_1, \mathbf{k}_2) = 1, \quad (\text{A2})$$

$$\mathcal{K}_{\mathcal{G}_2}^{(2)}(\mathbf{k}_1, \mathbf{k}_2) = K(\mathbf{k}_1, \mathbf{k}_2), \quad (\text{A3})$$

and the kernel $K(\mathbf{k}_1, \mathbf{k}_2)$ was already defined in Eq. (34). At third order the only nontrivial operator is $\mathcal{G}_2(\varphi_2, \varphi_1)$, which becomes upon Fourier transformation

$$\begin{aligned} \mathcal{G}_2(\varphi_2, \varphi_1 | \mathbf{k}) &= (2\pi)^3 \int_{\mathbf{k}_1, \mathbf{k}_2} [\delta_D]_2 K(\mathbf{k}_1, \mathbf{k}_2) \delta_L(\mathbf{k}_1) \mathcal{G}_2(\Phi_L | \mathbf{k}_2) \\ &= (2\pi)^3 \int_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} [\delta_D]_3 K(\mathbf{k}_1, \mathbf{k}_{23}) K(\mathbf{k}_2, \mathbf{k}_3) \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) \delta_L(\mathbf{k}_3), \end{aligned} \quad (\text{A4})$$

where we have made use of Eq. (32) and made the redefinition $\mathbf{k}_2 \rightarrow \mathbf{k}_{23}$ in the second step. After symmetrization we then obtain

$$\mathcal{K}_{\mathcal{G}^3}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 1, \quad (\text{A5})$$

$$\mathcal{K}_{\mathcal{G}_2}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{3} [K(\mathbf{k}_1, \mathbf{k}_2) + \text{cyc}], \quad (\text{A6})$$

$$\mathcal{K}_{\mathcal{G}_3}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \quad (\text{A7})$$

$$\mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_1)}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{3} [K(\mathbf{k}_1, \mathbf{k}_{23}) K(\mathbf{k}_2, \mathbf{k}_3) + \text{cyc}]. \quad (\text{A8})$$

Next, let us consider the most complicated combination that appears at fourth order, $\mathcal{G}_2(\varphi_3, \varphi_1)$, and all other operators will follow in a very similar manner. Starting from the definition in Eq. (50) and using the relations (47) to (49) for the LPT potentials, we have

$$\begin{aligned} \mathcal{G}_2(\varphi_3, \varphi_1 | \mathbf{k}) = (2\pi)^3 \int_{\mathbf{k}_1, \mathbf{k}_2} [\delta_D]_2 \left[\frac{1}{18} K(\mathbf{k}_1, \mathbf{k}_2) \left(\mathcal{G}_3(\varphi_1 | \mathbf{k}_2) + \frac{15}{7} \mathcal{G}_2(\varphi_2, \varphi_1 | \mathbf{k}_2) \right) \right. \\ \left. + \frac{1}{14} \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2) k_{1,j} k_{2,l}}{k_1^2 k_2^2} [\nabla_{lm} \varphi_1 \nabla_{jm} \varphi_2 - \nabla_{lm} \varphi_2 \nabla_{jm} \varphi_1](\mathbf{k}_2) \right] \delta_L(\mathbf{k}_1). \end{aligned} \quad (\text{A9})$$

Plugging in the Fourier expressions for the remaining potentials and Galileons [using Eq. (A4)], and replacing $\mathbf{k}_2 \rightarrow \mathbf{k}_{234}$, we get

$$\begin{aligned} \mathcal{G}_2(\varphi_3, \varphi_1 | \mathbf{k}) = (2\pi)^3 \int_{\mathbf{k}_1, \dots, \mathbf{k}_4} [\delta_D]_4 \left[\frac{1}{18} K(\mathbf{k}_1, \mathbf{k}_{234}) \left(\frac{15}{7} K(\mathbf{k}_{23}, \mathbf{k}_4) K(\mathbf{k}_2, \mathbf{k}_3) - L(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \right) \right. \\ \left. + \frac{1}{14} (M(\mathbf{k}_1, \mathbf{k}_{23}, \mathbf{k}_4, \mathbf{k}_{234}) - M(\mathbf{k}_1, \mathbf{k}_{234}, \mathbf{k}_{23}, \mathbf{k}_4)) K(\mathbf{k}_2, \mathbf{k}_3) \right] \prod_{i=1}^4 \delta_L(\mathbf{k}_i), \end{aligned} \quad (\text{A10})$$

where we have introduced the new kernel

$$M(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)(\mathbf{k}_2 \cdot \mathbf{k}_3)(\mathbf{k}_3 \cdot \mathbf{k}_4)(\mathbf{k}_4 \cdot \mathbf{k}_1)}{(k_1 k_2 k_3 k_4)^2}, \quad (\text{A11})$$

which is symmetric under cyclic permutations of its four momenta. The fully symmetric kernels for the basis operators at fourth order are thus given by

$$\mathcal{K}_{\delta^4}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv 1, \quad (\text{A12})$$

$$\mathcal{K}_{\delta^2 \mathcal{G}_2}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{6} [K(\mathbf{k}_1, \mathbf{k}_2) + \text{sym}(6)], \quad (\text{A13})$$

$$\mathcal{K}_{\delta \mathcal{G}_3}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{4} [L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \text{sym}(4)], \quad (\text{A14})$$

$$\mathcal{K}_{\mathcal{G}_2^2}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{3} [K(\mathbf{k}_1, \mathbf{k}_2) K(\mathbf{k}_3, \mathbf{k}_4) + \text{sym}(3)], \quad (\text{A15})$$

$$\mathcal{K}_{\delta \mathcal{G}_2(\varphi_2, \varphi_1)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{12} [K(\mathbf{k}_1, \mathbf{k}_{23}) K(\mathbf{k}_2, \mathbf{k}_3) + \text{sym}(12)], \quad (\text{A16})$$

$$\mathcal{K}_{\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{6} [L(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_{34}) K(\mathbf{k}_3, \mathbf{k}_4) + \text{sym}(6)], \quad (\text{A17})$$

$$\mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_2)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{3} [K(\mathbf{k}_{12}, \mathbf{k}_{34}) K(\mathbf{k}_1, \mathbf{k}_2) K(\mathbf{k}_3, \mathbf{k}_4) + \text{sym}(3)], \quad (\text{A18})$$

$$\begin{aligned} \mathcal{K}_{\mathcal{G}_2(\varphi_3, \varphi_1)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{12} \left[\frac{1}{18} K(\mathbf{k}_1, \mathbf{k}_{234}) \left(\frac{15}{7} K(\mathbf{k}_{23}, \mathbf{k}_4) K(\mathbf{k}_2, \mathbf{k}_3) - L(\mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \right) \right. \\ \left. + \frac{1}{14} (M(\mathbf{k}_1, \mathbf{k}_{23}, \mathbf{k}_4, \mathbf{k}_{234}) - M(\mathbf{k}_1, \mathbf{k}_{234}, \mathbf{k}_{23}, \mathbf{k}_4)) K(\mathbf{k}_2, \mathbf{k}_3) + \text{sym}(12) \right], \end{aligned} \quad (\text{A19})$$

where $\text{sym}(n)$ denotes the total number of terms the expressions have to be symmetrized over with respect to the four wave vectors.

2. Relation to other bias bases in the literature

In [6], the authors extend the bias basis from [13]. Let us establish the connection between their basis of operators

and ours up to fourth order. They use operators that correspond to tracing objects denoted as $\Pi^{[n]}$ defined from convective derivatives. The calculation of the connection between both bases is conceptually simple but in practice long and tedious beyond second order, so we only provide the final results. To linear order, we simply have $\text{Tr}[\Pi^{[1]}] = \delta$. The only nontrivial (i.e., beyond δ^2) at

quadratic order is $\text{Tr}[(\Pi^{[1]})^2] = \delta^2 + \mathcal{G}_2$. At cubic order we have for the nontrivial ones

$$\begin{aligned}\text{Tr}[(\Pi^{[1]})^3] &= \delta^3 + \frac{3}{2}\mathcal{G}_2\delta + \frac{1}{2}\mathcal{G}_3, \\ \text{Tr}[\Pi^{[1]}\Pi^{[2]}] &= \delta^3 + \frac{11}{14}\mathcal{G}_2\delta + \frac{1}{2}\mathcal{G}_3 - \frac{5}{7}\mathcal{G}_2(\varphi_2, \varphi_1),\end{aligned}\quad (\text{A20})$$

whereas at fourth order we have

$$\text{Tr}[(\Pi^{[1]})^4] = \delta^4 + 2\mathcal{G}_2\delta^2 + \frac{2}{3}\mathcal{G}_3\delta + \frac{1}{2}\mathcal{G}_2^2, \quad (\text{A21})$$

$$\begin{aligned}\text{Tr}[\Pi^{[1]}\Pi^{[1]}\Pi^{[2]}] &= \delta^4 + \frac{9}{7}\mathcal{G}_2\delta^2 + \frac{2}{3}\mathcal{G}_3\delta + \frac{1}{7}\mathcal{G}_2^2 \\ &\quad - \frac{5}{7}\mathcal{G}_2(\varphi_2, \varphi_1)\delta - \frac{5}{14}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1),\end{aligned}\quad (\text{A22})$$

$$\begin{aligned}\text{Tr}[\Pi^{[2]}\Pi^{[2]}] &= \delta^4 + \frac{4}{7}\mathcal{G}_2\delta^2 + \frac{2}{3}\mathcal{G}_3\delta + \frac{29}{98}\mathcal{G}_2^2 \\ &\quad - \frac{10}{7}\mathcal{G}_2(\varphi_2, \varphi_1)\delta - \frac{5}{7}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1) \\ &\quad + \frac{25}{49}\mathcal{G}_2(\varphi_2, \varphi_2),\end{aligned}\quad (\text{A23})$$

$$\begin{aligned}\text{Tr}[\Pi^{[1]}\Pi^{[3]}] &= 2\delta^4 + \frac{15}{7}\mathcal{G}_2\delta^2 - \frac{7}{9}\mathcal{G}_3\delta + \frac{1}{14}\mathcal{G}_2^2 \\ &\quad - \frac{4}{21}\mathcal{G}_2(\varphi_2, \varphi_1)\delta + \frac{13}{14}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1) \\ &\quad + 14\mathcal{G}_2(\varphi_3, \varphi_1).\end{aligned}\quad (\text{A24})$$

Note that these relations show that this basis mixes operators that are local and nonlocal in second derivatives of the linear potential.

Another set of papers use a yet different basis [30,31,70,103]. The comparison to these cases is more complicated since they also include velocity bias operators mixed together with standard ones. The simplest to compare with is [31], which when constrained to the no-velocity bias case shows one more operator (five in total) at third order than in our basis. In fact, this set is precisely the same set used in [8]; however, as pointed out in [9] not all of these five operators are independent (only four of them are). In terms of our basis, apart from δ^3 they have at cubic order the following operators:

$$\begin{aligned}s^3 &= \frac{1}{2}\mathcal{G}_3 + \frac{1}{2}\delta\mathcal{G}_2 + \frac{2}{9}\delta^3, & \delta s^2 &= \delta\mathcal{G}_2 + \frac{2}{3}\delta^3, \\ st &= \frac{2}{7}\mathcal{G}_2(\varphi_2, \varphi_1) + \frac{4}{21}\delta\mathcal{G}_2, \\ \psi &= \frac{5}{7}\mathcal{G}_3 + \delta\mathcal{G}_2 - \frac{30}{49}\mathcal{G}_2(\varphi_2, \varphi_1),\end{aligned}\quad (\text{A25})$$

which means that we can express δs^2 , s^3 , st , ψ each in terms of the four other basis elements in their basis, e.g.,

$$\psi = -\frac{15}{7}st + \frac{10}{7}s^3 + \frac{34}{49}\delta s^2 - \frac{344}{441}\delta^3. \quad (\text{A26})$$

Again, it is worth noting that Eq. (A25) shows that this basis mixes operators that are local and nonlocal in second derivatives of the linear potential.

The papers [30,31,103] do not consider fourth-order operators, but they do include higher-derivative operators. For example, the bispectrum [31] includes terms corresponding to our β_1 , $\beta_{2,1}$, $\beta_{2,2}$ (but they do not include $\beta_{2,3}$, $\beta_{2,4}$) and for the noise terms they include $\eta_{2,1}$, $\eta_{2,2}$ and seem to be missing ϵ_2 (though they do include such terms in the power spectrum, which for reasons that are unclear also includes a k^4 noise term); in addition, they have the usual ϵ_0 , η_0 shot noise terms. See our Table I and Eqs. (118) and (120) for reference of the operators associated with these coefficients.

Finally, the approach described in [47,48] uses a number of basis of operators obtained from different phenomenological models of biasing (e.g., halos, peaks). Note that the “renormalized bias functions” they define correspond to our initial multipoint propagators, whereas what they refer to as the multipoint propagators correspond to our evolved multipoint propagators. The LPT formalism of [46,47] was later extended by [104], and in the context of the two-point correlation function the authors of [105] introduced second-order bias operators equivalent to those given in [8], as well as a higher-derivative term. However, they neglect third-order bias parameters and thus do not cover the complete bias model at one-loop order.

3. Large-scale limit of fourth-order NLE operators

As shown in Sec. IV B, loop corrections of the multipoint propagators are exclusively given in terms of the NLE operators. In the large-scale limit these loop integrals can be Taylor expanded in powers of the participating wave vectors. For the fourth-order contributions to the two-point propagator we thus obtain

$$\int_q \mathcal{K}_{\delta\mathcal{G}_2(\varphi_2, \varphi_1)}^{(4,F)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, -\mathbf{q})P_L(q) = \frac{1}{1890} \{8k_{12}^2 - 16\mathbf{k}_1 \cdot \mathbf{k}_2 - [9k_{12}^2 - 18\mathbf{k}_1 \cdot \mathbf{k}_2]K(\mathbf{k}_1, \mathbf{k}_2)\} \int_q \frac{P_L(q)}{q^2} + \mathcal{O}(k_i^4), \quad (\text{A27})$$

$$\int_q \mathcal{K}_{\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, -\mathbf{q}) P_L(q) = \frac{1}{135} [k_{12}^2 - 2\mathbf{k}_1 \cdot \mathbf{k}_2] K(\mathbf{k}_1, \mathbf{k}_2) \int_q \frac{P_L(q)}{q^2} + \mathcal{O}(k_i^4), \quad (\text{A28})$$

$$\int_q \mathcal{K}_{\mathcal{G}_2(\varphi_2, \varphi_2)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, -\mathbf{q}) P_L(q) = -\frac{8}{945} \{3k_{12}^2 + [k_{12}^2 - \mathbf{k}_1 \cdot \mathbf{k}_2] K(\mathbf{k}_1, \mathbf{k}_2)\} \int_q \frac{P_L(q)}{q^2} + \mathcal{O}(k_i^4), \quad (\text{A29})$$

$$\int_q \mathcal{K}_{\mathcal{G}_2(\varphi_3, \varphi_1)}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, -\mathbf{q}) P_L(q) = -\frac{1}{39690} \{69k_{12}^2 - [12k_{12}^2 - 26\mathbf{k}_1 \cdot \mathbf{k}_2] K(\mathbf{k}_1, \mathbf{k}_2)\} \int_q \frac{P_L(q)}{q^2} + \mathcal{O}(k_i^4), \quad (\text{A30})$$

where $\mathcal{O}(k_i^4)$ denotes terms that are of order k_1^4 , k_2^4 , or k_{12}^4 and higher. These expressions serve to illustrate that the nonlocal evolution terms in the large-scale limit are given by combinations of the higher-derivative operators. The converse is equally true as can easily be verified by inverting the relations above.

APPENDIX B: TIME EVOLUTION OF MULTIPOINT PROPAGATORS

In this Appendix we give a detailed derivation for the time evolution of the multipoint propagators. We proceed in two steps: first, we evaluate expectation values of products of two or three Wiener-Hermite functionals, and second, by using these results we directly show how to obtain the recursion relations reported in Eq. (97).

1. Orthogonality relations for generalized Wiener-Hermite functionals

Let us consider the PDF of δ_L , shifted by a generic source term $\alpha(\mathbf{k})$, which we take to be an arbitrary function of wave number $\mathbf{k}$. A Taylor expansion around $\alpha = 0$ yields

$$\mathcal{P}[\alpha - \delta_L] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_n} \frac{\partial^n \mathcal{P}[\alpha - \delta_L]}{\partial \alpha_1 \dots \partial \alpha_n} \Big|_{\alpha=0} \alpha_1 \dots \alpha_n, \quad (\text{B1})$$

where $\alpha_i \equiv \alpha(\mathbf{k}_i)$. Swapping the derivatives from α to δ_L and using the definition of the Wiener-Hermite functionals from Eq. (55) we get

$$\begin{aligned} \frac{\partial^n \mathcal{P}[\alpha - \delta_L]}{\partial \alpha_1 \dots \partial \alpha_n} \Big|_{\alpha=0} &= (-1)^n \frac{\partial^n \mathcal{P}[\delta_L]}{\partial \delta_{L,1} \dots \partial \delta_{L,n}} \Big|_{\alpha=0} \\ &= \frac{\mathcal{P}[\delta_L] \mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n)}{P_L(k_1) \dots P_L(k_n)}, \end{aligned} \quad (\text{B2})$$

and thus

$$\frac{\mathcal{P}[\alpha - \delta_L]}{\mathcal{P}[\delta_L]} = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_n} \frac{\mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n)}{P_L(k_1) \dots P_L(k_n)} \alpha_1 \dots \alpha_n. \quad (\text{B3})$$

To derive the orthogonality relation between two Wiener-Hermite functionals of orders m and n , we first compute the following integral:

$$\int \mathcal{D}[\delta_L] \mathcal{P}[\delta_L] \frac{\mathcal{P}[\alpha - \delta_L] \mathcal{P}[\beta - \delta_L]}{\mathcal{P}[\delta_L] \mathcal{P}[\delta_L]} = \exp \left[\int_q \frac{\alpha(\mathbf{q}) \beta(-\mathbf{q})}{P_L(q)} \right], \quad (\text{B4})$$

where we have plugged in Eq. (54). However, using Eq. (B3) to replace the PDFs, we must also have

$$\sum_{m,n} \frac{1}{m!n!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_m} \int_{\mathbf{q}_1, \dots, \mathbf{q}_n} \frac{\langle \mathcal{H}_m(\mathbf{k}_1, \dots, \mathbf{k}_m) \mathcal{H}_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \rangle}{P_L(k_1) \dots P_L(k_m) P_L(q_1) \dots P_L(q_n)} \alpha_1 \dots \alpha_m \beta_1 \dots \beta_n = \exp \left[\int_q \frac{\alpha(\mathbf{q}) \beta(-\mathbf{q})}{P_L(q)} \right], \quad (\text{B5})$$

and by Taylor expanding the right-hand side of the expression above, we see that we need to match up all $\mathbf{k}$ and $\mathbf{q}$ modes, which is possible only if $m = n$. From that observation it immediately follows that

$$\langle \mathcal{H}_m(\mathbf{k}_1, \dots, \mathbf{k}_m) \mathcal{H}_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \rangle = (2\pi)^{3m} \delta_{mn}^K \delta_D(\mathbf{k}_{\{1,m\}}, \mathbf{q}_{\{1,m\}}) \prod_{i=1}^m P_L(k_i), \quad (\text{B6})$$

where we have used the shorthand notation

$$\delta_D(\mathbf{k}_{\{1,m\}}, \mathbf{q}_{\{1,m\}}) \equiv \delta_D(\mathbf{k}_1 + \mathbf{q}_1) \dots \delta_D(\mathbf{k}_m + \mathbf{q}_m) + \text{sym} \quad (\text{B7})$$

for writing all possible ways ($m!$ in total) of matching up the two sets of modes.

Let us now compute the expectation value of three Wiener-Hermite functionals, i.e., $\langle \mathcal{H}_m \mathcal{H}_n \mathcal{H}_l \rangle$. Similar to the above procedure, we first evaluate an integral over PDFs, now with the three different sources α , β , and γ :

$$\int \mathcal{D}[\delta_L] \mathcal{P}[\delta_L] \frac{\mathcal{P}[\alpha - \delta_L]}{\mathcal{P}[\delta_L]} \frac{\mathcal{P}[\beta - \delta_L]}{\mathcal{P}[\delta_L]} \frac{\mathcal{P}[\gamma - \delta_L]}{\mathcal{P}[\delta_L]} = \exp \left[\int_{\mathbf{q}} \frac{\alpha(\mathbf{q})\beta(-\mathbf{q}) + \alpha(\mathbf{q})\gamma(-\mathbf{q}) + \beta(\mathbf{q})\gamma(-\mathbf{q})}{P_L(\mathbf{q})} \right]. \quad (\text{B8})$$

Expanding both sides of Eq. (B8), we get

$$\begin{aligned} (\text{B8, LHS}) &= \sum_{m,n,l} \frac{1}{m!n!l!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_m} \int_{\mathbf{q}_1, \dots, \mathbf{q}_n} \int_{\mathbf{p}_1, \dots, \mathbf{p}_l} \langle \mathcal{H}_m(\mathbf{k}_1, \dots, \mathbf{k}_m) \mathcal{H}_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \mathcal{H}_l(\mathbf{p}_1, \dots, \mathbf{p}_l) \rangle \\ &\quad \times \left[\prod_{i=1}^m \frac{\alpha(\mathbf{k}_i)}{P_L(\mathbf{k}_i)} \right] \left[\prod_{i=1}^n \frac{\beta(\mathbf{q}_i)}{P_L(\mathbf{q}_i)} \right] \left[\prod_{i=1}^l \frac{\gamma(\mathbf{p}_i)}{P_L(\mathbf{p}_i)} \right] \end{aligned} \quad (\text{B9})$$

and

$$(\text{B8, RHS}) = \sum_{a,b,c} \frac{1}{a!b!c!} \int_{\mathbf{k}_1, \dots, \mathbf{k}_a} \int_{\mathbf{q}_1, \dots, \mathbf{q}_b} \int_{\mathbf{p}_1, \dots, \mathbf{p}_c} \left[\prod_{i=1}^a \frac{\alpha(\mathbf{k}_i)\beta(-\mathbf{k}_i)}{P_L(\mathbf{k}_i)} \right] \left[\prod_{i=1}^b \frac{\alpha(\mathbf{q}_i)\gamma(-\mathbf{q}_i)}{P_L(\mathbf{q}_i)} \right] \left[\prod_{i=1}^c \frac{\beta(\mathbf{p}_i)\gamma(-\mathbf{p}_i)}{P_L(\mathbf{p}_i)} \right]. \quad (\text{B10})$$

In order for Eqs. (B9) and (B10) to be equal, they need to contain the same number of source terms, which requires that $l + m + n = 2(a + b + c)$ and thus, $l + m + n \in 2\mathbb{N}$. Moreover, the indices must satisfy the conditions

$$\left. \begin{aligned} a + b &= m \\ a + c &= n \\ b + c &= l \end{aligned} \right\} \Leftrightarrow \left\{ \begin{aligned} a &= \frac{m+n-l}{2} \\ b &= \frac{l+m-n}{2} \\ c &= \frac{n+l-m}{2} \end{aligned} \right., \quad (\text{B11})$$

from which follows that $m + n \geq l$, and cyclic permutations thereof. According to Eq. (B11), we can divide all $\mathbf{k}$, $\mathbf{q}$, and $\mathbf{p}$ modes into two subsets each, of size a , b , or c . For instance, the $\mathbf{k}$ modes will be split into a group containing a modes, and another containing b , such that $a + b = m$. Each mode in the former group can then be assigned a mode from an equally sized group of $\mathbf{q}$'s, i.e., $\mathbf{q}_1 = -\mathbf{k}_1, \dots, \mathbf{q}_a = -\mathbf{k}_a$, while the latter are matched in a similar manner with a set of $\mathbf{p}$ modes, which guarantees that we obtain Eq. (B10). Altogether, this means we need to require that the expectation value of the three Wiener-Hermite functionals is given by

$$\begin{aligned} \langle \mathcal{H}_m(\mathbf{k}_1, \dots, \mathbf{k}_m) \mathcal{H}_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \mathcal{H}_l(\mathbf{p}_1, \dots, \mathbf{p}_l) \rangle &= (2\pi)^{\frac{3}{2}(m+n+l)} \left[\prod_{i=1}^b P_L(\mathbf{k}_{a+i}) \right] \left[\prod_{i=1}^a P_L(\mathbf{q}_i) \right] \left[\prod_{i=1}^c P_L(\mathbf{p}_{b+i}) \right] \\ &\quad \times [\delta_D(\mathbf{k}_{\{1,a\}}, \mathbf{q}_{\{1,a\}}) \delta_D(\mathbf{k}_{\{a+1,m\}}, \mathbf{p}_{\{1,b\}}) \delta_D(\mathbf{q}_{\{a+1,n\}}, \mathbf{p}_{\{b+1,l\}}) + \text{sym}], \end{aligned} \quad (\text{B12})$$

and it must vanish if the conditions above are not satisfied. The number of terms that have to be added in order to symmetrize Eq. (B12) is given by the number of possibilities of selecting subsets of $\mathbf{k}$, $\mathbf{q}$, and $\mathbf{p}$ modes that are of size b , a , and c , respectively. The total number of terms in the square brackets is thus

$$\binom{m}{b} \times \binom{n}{a} \times \binom{l}{c} \times a!b!c! = \frac{m!n!l!}{a!b!c!}. \quad (\text{B13})$$

2. The Γ -recursion relation

We are interested in the time evolution of the n th-order multipoint propagator. Exploiting the orthogonality of Wiener-Hermite functionals, we can single out its contribution to the series expansion in Eq. (93) by multiplying both sides with $\mathcal{H}_n$ and taking the ensemble average:

$$\langle \mathcal{H}_n(\delta_L | \mathbf{k}_1, \dots, \mathbf{k}_n) \Psi_a(\mathbf{k}, \eta) \rangle = (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n}) \Gamma_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta) \prod_{i=1}^n P_L(k_i). \quad (\text{B14})$$

Applying the same procedure to the integral solution for Ψ_a from Eq. (91) we get

$$\begin{aligned}
\Gamma_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta) &= g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) + \sum_{m,l=1} \frac{1}{m!l!} \int_0^\eta d\eta' g_{ab}(\eta - \eta') \\
&\times \int_{\mathbf{q}_1, \dots, \mathbf{q}_m} \int_{\mathbf{p}_1, \dots, \mathbf{p}_l} \gamma_{bcd}(\mathbf{q}_{1\dots m}, \mathbf{p}_{1\dots l}) \Gamma_c^{(m)}(\mathbf{q}_1, \dots, \mathbf{q}_m; \eta') \Gamma_d^{(l)}(\mathbf{p}_1, \dots, \mathbf{p}_l; \eta') \\
&\times \langle \mathcal{H}_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \mathcal{H}_m^*(\mathbf{q}_1, \dots, \mathbf{q}_m) \mathcal{H}_l^*(\mathbf{p}_1, \dots, \mathbf{p}_l) \rangle \prod_{i=1}^n \frac{1}{P_L(k_i)}, \tag{B15}
\end{aligned}$$

where $\Gamma_{a,\mathcal{L}}$ denotes multipoint propagators at initial time $\eta = 0$. Upon inserting Eq. (B12) we perform the first $a = (n + m - l)/2$ integrations over $\mathbf{q}$ modes as well as over all of the $\mathbf{p}$ modes, such that the second term above reduces to

$$\begin{aligned}
&\sum_{\substack{m,l=1 \\ n+m+l \in 2\mathbb{N} \\ n+m \geq l, \text{cyc}}} \frac{1}{m!l!} \kappa \int_0^\eta d\eta' g_{ab}(\eta - \eta') \int_{\mathbf{q}_{a+1}, \dots, \mathbf{q}_m} [\gamma_{bcd}(\mathbf{k}_{1\dots a} + \mathbf{q}_{a+1\dots m}, \mathbf{k}_{a+1\dots n} - \mathbf{q}_{a+1\dots m}) \\
&\times \Gamma_c^{(m)}(\mathbf{k}_1, \dots, \mathbf{k}_a, \mathbf{q}_{a+1}, \dots, \mathbf{q}_m; \eta') \Gamma_d^{(l)}(\mathbf{k}_{a+1}, \dots, \mathbf{k}_n, -\mathbf{q}_{a+1}, \dots, -\mathbf{q}_m; \eta') + \text{sym}] \prod_{i=a+1}^m P_L(q_i). \tag{B16}
\end{aligned}$$

The symmetrization is carried out by summing over all $\binom{n}{a}$ subsets of $\mathbf{k}$ modes, which implies that the combinatorial factor κ is given by

$$\kappa = \frac{m!n!l!}{a!b!c!} \binom{n}{a}^{-1} = \frac{m!l!}{(m-a)!}. \tag{B17}$$

Next, let us change the summation index from l to $r = (l + m - n)/2 = m - a$ and relabel the remaining integrations over $\mathbf{q}$'s into $\mathbf{q}_1, \dots, \mathbf{q}_r$. The conditions $n + m \geq l$ (and cyclic permutations) then transform into $r_* \leq r \leq m$, where

$$r_* = \begin{cases} 0, & n > m \\ 1, & n = m \\ m - n, & n < m \end{cases}. \tag{B18}$$

Defining the quantity

$$\begin{aligned}
\Gamma_a^{(n,m,r)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta) &\equiv \int_{\mathbf{q}_1, \dots, \mathbf{q}_r} \left[\gamma_{abc}(\mathbf{k}_{1\dots m-r} + \mathbf{q}_{1\dots r}, \mathbf{k}_{m-r+1\dots n} - \mathbf{q}_{1\dots r}) \Gamma_b^{(m)}(\mathbf{k}_1, \dots, \mathbf{k}_{m-r}, \mathbf{q}_1, \dots, \mathbf{q}_r; \eta) \right. \\
&\times \Gamma_c^{(2r+n-m)}(\mathbf{k}_{m-r+1}, \dots, \mathbf{k}_n, -\mathbf{q}_1, \dots, -\mathbf{q}_r; \eta) \prod_{i=1}^r P_L(q_i) + \text{sym}(\mathbf{k}_i) \Big], \tag{B19}
\end{aligned}$$

we finally obtain the desired recursion relation for multipoint propagators,

$$\Gamma_a^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta) = g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) + \sum_{r=0}^{r+n-\delta_{r,0}^K} \sum_{m=\max\{r,1\}} \frac{1}{r!} \int_0^\eta d\eta' g_{ab}(\eta - \eta') \Gamma_b^{(n,m,r)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \eta'), \tag{B20}$$

where we have arranged the two summations in the second contribution such that the outer sum runs over the number of loop integrals, r , and used Eq. (B18) to determine the allowed range of the index m .

3. Computation of time-evolved propagators up to one-loop order

We now use the recursion relation from Sec. B 2 to derive explicit expressions for the building blocks of the one-loop bispectrum, which requires us to evaluate the first two multipoint propagators at one-loop order and the three-point propagator at tree level. In order to make the connection to the initial conditions evident, where convenient, we write them completely in terms of the Lagrangian multipoint propagators, which we computed in Sec. IV B.

To simplify the resulting expressions it is useful to define the following two quantities that arise through contractions of the linear propagator g_{ab} with the vertices γ_{abc} :

$$g_{abc}^{(2)}(\mathbf{k}_1, \mathbf{k}_2; \eta) \equiv \int_0^\eta d\eta' g_{ad}(\eta - \eta') \gamma_{def}(\mathbf{k}_1, \mathbf{k}_2) g_{eb}(\eta') g_{fc}(\eta'), \quad (\text{B21})$$

$$g_{abcd}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \eta) \equiv \int_0^\eta d\eta' g_{ae}(\eta - \eta') \gamma_{efg}(\mathbf{k}_1, \mathbf{k}_{23}) g_{fb}(\eta') g_{gc}^{(2)}(\mathbf{k}_2, \mathbf{k}_3; \eta'), \quad (\text{B22})$$

where we let the superscript indicate the number of external lines that are attached to these objects, i.e., two for $g_{abc}^{(2)}$, etc. The tree-level expression ($r = 0$) for $\Gamma_a^{(1)}$ is trivial since the second piece in Eq. (B20) makes no contribution. For that reason we start with the two-point propagator, and by using the kernel from Eq. (B21) we obtain

$$\Gamma_a^{(2)}(\mathbf{k}_1, \mathbf{k}_2; \eta)|_{\text{tree}} = g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(2)}(\mathbf{k}_1, \mathbf{k}_2) + 2g_{abc}^{(2)}(\mathbf{k}_1, \mathbf{k}_2; \eta) \Gamma_{b,\mathcal{L}}^{(1)}(\mathbf{k}_1) \Gamma_{c,\mathcal{L}}^{(1)}(\mathbf{k}_2), \quad (\text{B23})$$

with all initial multipoint propagators being evaluated at tree level. The tree-level three-point propagator receives two contributions from the second term in Eq. (B20), such that

$$\begin{aligned} \Gamma_a^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \eta)|_{\text{tree}} &= g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + [2g_{abc}^{(2)}(\mathbf{k}_1, \mathbf{k}_{23}; \eta) \Gamma_{b,\mathcal{L}}^{(1)}(\mathbf{k}_1) \Gamma_{c,\mathcal{L}}^{(2)}(\mathbf{k}_2, \mathbf{k}_3) \\ &\quad + 4g_{abcd}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \eta) \Gamma_{b,\mathcal{L}}^{(1)}(\mathbf{k}_1) \Gamma_{c,\mathcal{L}}^{(1)}(\mathbf{k}_2) \Gamma_{d,\mathcal{L}}^{(1)}(\mathbf{k}_3) + \text{cyc}], \end{aligned} \quad (\text{B24})$$

where we have plugged in the solution for $\Gamma_a^{(2)}$ from above. Similarly, Eq. (B23) can be used to determine the nonlinear correction to the one-point propagator at one-loop order (or $r = 1$), giving

$$\begin{aligned} \Gamma_a^{(1)}(\mathbf{k}; \eta)|_{\text{loop}} &= g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(1)}(\mathbf{k})|_{\text{loop}} + 2 \int_{\mathbf{q}} g_{abc}^{(2)}(\mathbf{q}, \mathbf{k} - \mathbf{q}; \eta) \Gamma_{b,\mathcal{L}}^{(1)}(\mathbf{q}) \Gamma_{c,\mathcal{L}}^{(2)}(\mathbf{k}, -\mathbf{q}) P_L(q) \\ &\quad + 4 \int_{\mathbf{q}} g_{abcd}^{(3)}(\mathbf{q}, \mathbf{k}, -\mathbf{q}; \eta) \Gamma_{b,\mathcal{L}}^{(1)}(\mathbf{q}) \Gamma_{c,\mathcal{L}}^{(1)}(\mathbf{k}) \Gamma_{d,\mathcal{L}}^{(1)}(-\mathbf{q}) P_L(q). \end{aligned} \quad (\text{B25})$$

Equations (B23)–(B25) are the full results including all transients and the evolution of the Lagrangian bias parameters. For comparison to observational data only the fastest growing mode is relevant, and we can further absorb all terms due to bias evolution into the final Eulerian bias parameters (see Sec. IV C 1), allowing us to greatly simplify these expressions. Moreover, if we separate the contributions that get newly generated by nonlinear evolution (i.e., they are not due to bias evolution of terms already present in the initial multipoint propagators), and

expressing them in terms of the SPT kernels, we arrive at Eqs. (100)–(102).

While Eqs. (100)–(102) can more easily be obtained from the nonlinear evolution of the matter and velocity fields directly, starting from $\Gamma^{(2)}$ at one-loop order the evolution of the multipoint propagators becomes useful, since we do not have to deal with cumbersome renormalizations as shown in Sec. II C. Because the result is still complicated, we have chosen to present it in terms of the previously derived expressions for the evolved propagators,

$$\begin{aligned} \Gamma_a^{(2)}(\mathbf{k}_1, \mathbf{k}_2; \eta)|_{\text{loop}} &= g_{ab}(\eta) \Gamma_{b,\mathcal{L}}^{(2)}(\mathbf{k}_1, \mathbf{k}_2)|_{\text{loop}} + 2 \int_0^\eta d\eta' g_{ab}(\eta - \eta') \gamma_{bcd}(\mathbf{k}_1, \mathbf{k}_2) [\Gamma_c^{(1)}(\mathbf{k}_1; \eta')|_{\text{loop}} \Gamma_d^{(1)}(\mathbf{k}_2; \eta')|_{\text{tree}} \\ &\quad + \Gamma_c^{(1)}(\mathbf{k}_1; \eta')|_{\text{tree}} \Gamma_d^{(1)}(\mathbf{k}_2; \eta')|_{\text{loop}}] + 2 \int_{\mathbf{q}} P_L(q) \int_0^\eta d\eta' g_{ab}(\eta - \eta') [\gamma_{bcd}(\mathbf{q}, \mathbf{k}_{12} - \mathbf{q}) \Gamma_c^{(1)} \\ &\quad \times (\mathbf{q}; \eta')|_{\text{tree}} \Gamma_d^{(3)}(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{q}; \eta')|_{\text{tree}} + \gamma_{bcd}(\mathbf{k}_1 + \mathbf{q}, \mathbf{k}_2 - \mathbf{q}) \Gamma_c^{(2)}(\mathbf{k}_1, \mathbf{q}; \eta')|_{\text{tree}} \Gamma_d^{(2)}(\mathbf{k}_2, -\mathbf{q}; \eta')|_{\text{tree}}], \end{aligned} \quad (\text{B26})$$

where we have indicated at which order the propagators are evaluated.

APPENDIX C: USING GALILEAN INVARIANCE TO COMPUTE TIME EVOLUTION OF BIAS

1. SPT in Galilean basis for galaxy bias

Let us consider the Galilean invariant contributions to SPT at each order in perturbation theory (PT). This will allow us to

calculate in a more straightforward way the time evolution of bias parameters up to any order in PT (we will compute up to fourth order here). Since our bias basis contains the LPT potentials, we start from the expression for density fluctuations in terms of the Lagrangian displacement field,

$$1 + \delta(\mathbf{x}) = \frac{1}{\det[\delta_{ij}^K + \Psi_{i,j}(\mathbf{q})]} = \frac{1}{1 + G_1(\Psi) - \frac{1}{2}G_2(\Psi) + \frac{1}{6}G_3(\Psi)} \Big|_{\mathbf{q}}, \quad (\text{C1})$$

where $\Psi_{i,j} \equiv \partial \Psi_i / \partial q_j$ (δ_{ij}^K denotes the Kronecker delta) and the G_i 's are the Galileons built out of the deformation tensor; working up to fourth order in PT they read

$$G_1(\Psi) \equiv \Psi_{i,i} = -\delta + \frac{3}{14}\mathcal{G}_2 + \frac{1}{18}\mathcal{G}_3 - \frac{5}{42}\mathcal{G}_2(\varphi_2, \varphi_1) + \frac{7}{11}\mathcal{G}_2(\varphi_3, \varphi_1) + \frac{51}{4312}\mathcal{G}_2(\varphi_2, \varphi_2) - \frac{13}{308}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1), \quad (\text{C2})$$

$$G_2(\Psi) \equiv (\Psi_{i,j})^2 - (\Psi_{i,i})^2 = \mathcal{G}_2 - \frac{3}{7}\mathcal{G}_2(\varphi_2, \varphi_1) + 2\mathcal{G}_2(\varphi_3, \varphi_1) + \frac{9}{196}\mathcal{G}_2(\varphi_2, \varphi_2), \quad (\text{C3})$$

$$G_3(\Psi) \equiv (\Psi_{i,i})^3 + 2\Psi_{i,j}\Psi_{j,k}\Psi_{k,i} - 3(\Psi_{i,j})^2(\Psi_{i,i}) = -\mathcal{G}_3 + \frac{9}{14}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1). \quad (\text{C4})$$

By using these results in Eq. (C1), and using that $\mathbf{x} = \mathbf{q} + \Psi$, one can derive the standard Eulerian PT expansion. However, we are interested in finding out the bias relation in the absence of velocity bias; therefore any dipole terms coming from displacement from $\mathbf{q}$ to $\mathbf{x}$ will not appear in any of the final expressions, and our basis being Galilean invariant does not include any such operators. Therefore, we can simply work with the Galilean invariant SPT expression at each order, since dipole terms will cancel at the end of the calculation. Thus we have

$$1 + \delta_{\text{GI}}(\mathbf{x}) = \frac{1}{1 + G_1(\Psi) - \frac{1}{2}G_2(\Psi) + \frac{1}{6}G_3(\Psi)} \Big|_{\mathbf{x}}. \quad (\text{C5})$$

Using Eqs. (C2)–(C4) in Eq. (C5) we obtain the Galilean invariant piece of the SPT expansion at each order in our basis of bias operators,

$$\begin{aligned} \delta_{\text{GI}}^{(1)} &= \delta, & \delta_{\text{GI}}^{(2)} &= \delta^2 + \frac{2}{7}\mathcal{G}_2, \\ \delta_{\text{GI}}^{(3)} &= \delta^3 + \frac{4}{7}\delta\mathcal{G}_2 - \frac{2}{21}\mathcal{G}_2(\varphi_2, \varphi_1) + \frac{1}{9}\mathcal{G}_3, \end{aligned} \quad (\text{C6})$$

$$\begin{aligned} \delta_{\text{GI}}^{(4)} &= \delta^4 + \frac{6}{7}\delta^2\mathcal{G}_2 + \frac{4}{49}\mathcal{G}_2^2 - \frac{4}{21}\delta\mathcal{G}_2(\varphi_2, \varphi_1) \\ &+ \frac{2}{9}\delta\mathcal{G}_3 + \frac{6}{539}\mathcal{G}_2(\varphi_2, \varphi_2) + \frac{4}{11}\mathcal{G}_2(\varphi_3, \varphi_1) \\ &- \frac{5}{77}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1). \end{aligned} \quad (\text{C7})$$

A simple check on these expressions is provided by the spherical collapse dynamics, which is also Galilean

invariant. To do this all we need is the spherical average of our basis operators, which is straightforward,

$$\begin{aligned} \overline{\mathcal{G}_2} &= -\frac{2}{3}\delta^2, & \overline{\mathcal{G}_2(\varphi_2, \varphi_1)} &= 2\overline{\mathcal{G}_3} = \frac{4}{9}\delta^3, \\ \overline{\mathcal{G}_2(\varphi_2, \varphi_2)} &= 2\overline{\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1)} = -\frac{8}{27}\delta^4, \\ \overline{\mathcal{G}_2(\varphi_3, \varphi_1)} &= -\frac{46}{1701}\delta^4. \end{aligned} \quad (\text{C8})$$

Using these results in Eqs. (C6) and (C7) we recover the well-known vertices ν_n from the spherical collapse dynamics [106], i.e.,

$$\begin{aligned} \delta_{\text{SC}} &= \sum_n \frac{\nu_n}{n!} \delta^n, & \nu_1 &= 1, & \nu_2 &= \frac{34}{21}, \\ \nu_3 &= \frac{682}{189}, & \nu_4 &= \frac{446440}{43659}. \end{aligned} \quad (\text{C9})$$

Now we proceed to derive the analogous results for the velocity field, which is what is needed for the nonlinear evolution of Φ_v or the normalized divergence $\theta = \nabla^2 \Phi_v$. By definition, the Eulerian velocity field agrees with the velocity of particles calculated from the time derivative of the displacement field at the same location. Taking the normalized divergence $-\nabla/\mathcal{H}f$ we then have that

$$\theta(\mathbf{x}) = -\nabla_{\mathbf{x}} \cdot \Psi'(\mathbf{q}), \quad (\text{C10})$$

where $' \equiv \partial/\partial \ln D$, with D the linear growth factor. Again, we are only interested in the Galilean invariant piece of this

exact relation. In order to obtain that, all we need is $\Psi'(\mathbf{q})$ as a function of $\mathbf{x}$, with *at most* one displacement free from derivatives acting on it, because other contributions cannot be rendered Galilean invariant by taking a single derivative when calculating $\theta_{\text{GI}}(\mathbf{x})$. A straightforward Taylor expansion gives then

$$\begin{aligned} \Psi'(\mathbf{q}) &= \Psi' - \Psi_i(\nabla_i \Psi') + \Psi_j(\nabla_j \Psi_i)(\nabla_i \Psi') \\ &\quad - \Psi_k(\nabla_k \Psi_j)(\nabla_j \Psi_i)(\nabla_i \Psi') + \dots, \end{aligned} \quad (\text{C11})$$

where all quantities on the RHS of this equation are now evaluated at $\mathbf{x}$. The result for θ_{GI} follows by acting with the divergence in Eq. (C10) on the derivative-free displacement field in each term of Eq. (C11) giving us the simple result

$$\theta_{\text{GI}}(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n} [\text{Tr}(\Psi_{i,j})^n]'. \quad (\text{C12})$$

The traces here can easily be rewritten in terms of the deformation tensor Galileons in Eq. (C2), and the derivative with respect to $\ln D$ just multiplies each contribution by its PT order. Therefore we obtain the Galilean invariant piece of the normalized divergence at each order,

$$\begin{aligned} \theta_{\text{GI}}^{(1)} &= \delta, & \theta_{\text{GI}}^{(2)} &= \delta^2 + \frac{4}{7} \mathcal{G}_2, \\ \theta_{\text{GI}}^{(3)} &= \delta^3 + \frac{6}{7} \delta \mathcal{G}_2 - \frac{2}{7} \mathcal{G}_2(\varphi_2, \varphi_1) + \frac{1}{3} \mathcal{G}_3, \end{aligned} \quad (\text{C13})$$

$$\begin{aligned} \theta_{\text{GI}}^{(4)} &= \delta^4 + \frac{8}{7} \delta^2 \mathcal{G}_2 + \frac{8}{49} \mathcal{G}_2^2 - \frac{8}{21} \delta \mathcal{G}_2(\varphi_2, \varphi_1) \\ &\quad + \frac{4}{9} \delta \mathcal{G}_3 + \frac{24}{539} \mathcal{G}_2(\varphi_2, \varphi_2) + \frac{16}{11} \mathcal{G}_2(\varphi_3, \varphi_1) \\ &\quad - \frac{20}{77} \mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1). \end{aligned} \quad (\text{C14})$$

Again, we can double-check these results comparing them to the spherical collapse dynamics by performing the spherical average using Eq. (C8) in Eqs. (C13) and (C14) to compute the vertices μ_n for the normalized velocity divergence,

$$\begin{aligned} \theta_{\text{SC}} &= \sum_n \frac{\mu_n}{n!} \delta^n, & \mu_1 &= 1, & \mu_2 &= \frac{26}{21}, \\ \mu_3 &= \frac{142}{63}, & \mu_4 &= \frac{236872}{43659}, \end{aligned} \quad (\text{C15})$$

which agree with the standard spherical collapse dynamics values [106].

2. Time evolution of bias

As already explained in Sec. IV C 1, in order to determine the time evolution of bias, we can use a simple trick that exploits the Galilean invariance of the galaxy

overdensity in the absence of velocity bias. That means all dipole terms in the matter density must cancel, such that

$$\delta_g(\mathbf{x}) = \delta_g^{\text{IC}}(\mathbf{x}) + \delta^{\text{GI}}(\mathbf{x}) + \delta_g^{\text{IC}}(\mathbf{x}) \delta^{\text{GI}}(\mathbf{x}), \quad (\text{C16})$$

where δ^{GI} is the Galilean invariant part of the (nonlinear) matter density as determined in the last section, and δ_g^{IC} is the galaxy density at initial time (where all fluctuations are linear and Gaussian). At linear order it is easy to check that Eq. (C16) simply gives $b_1 = 1 + b_{1,\mathcal{L}}$, which agrees with the full evolution [69] when neglecting transients. To find the evolution of higher-order bias parameters, we must first subtract all contributions due to lower-order operators (see [9]). For instance, by computing (suppressing $\mathbf{x}$ arguments)

$$\delta_g - b_1 \delta^{\text{GI}} = \delta_g^{\text{IC}} - b_{1,\mathcal{L}} \delta^{\text{GI}} + \delta_g^{\text{IC}} \delta^{\text{GI}}, \quad (\text{C17})$$

and writing the RHS of this equation in our bias basis using Eqs. (78) and (C6) at second order, we are able to identify the coefficients of the δ^2 and $\mathcal{G}_2$ operators, which leads to

$$b_2 = b_{2,\mathcal{L}}, \quad \gamma_2 = -\frac{2}{7} b_{1,\mathcal{L}} + \gamma_{2,\mathcal{L}}. \quad (\text{C18})$$

At cubic order we proceed in the same way and compute

$$\delta_g^{(3)} - b_1 \delta_{\text{GI}}^{(3)} - \frac{b_2}{2} [\delta_{\text{GI}}^2]^{(3)} - \gamma_2 [\mathcal{G}_2]^{(3)}, \quad (\text{C19})$$

and we use Eqs. (78) and (C6) plus that

$$[\mathcal{G}_2]^{(3)} = \mathcal{G}_3 + \delta \mathcal{G}_2 - \frac{6}{7} \mathcal{G}_2(\varphi_2, \varphi_1), \quad (\text{C20})$$

which follows from Eq. (93) in [9] when constrained to Galilean invariant operators [for which we need Eqs. (C6) and (C13)]. We thus identify the coefficients (i.e., bias parameters) in front of the operators in our bias basis at third order,

$$\begin{aligned} b_3 &= b_{3,\mathcal{L}} - 3b_2, & \gamma_3 &= -\frac{1}{9} b_{1,\mathcal{L}} - \gamma_2 + \gamma_{3,\mathcal{L}}, \\ \gamma_2^\times &= -\frac{2}{7} b_2 + \gamma_{2,\mathcal{L}}^\times, & \gamma_{21} &= \frac{2}{21} b_{1,\mathcal{L}} + \frac{6}{7} \gamma_2 + \gamma_{21,\mathcal{L}}. \end{aligned} \quad (\text{C21})$$

Equations (C18)–(C21) agree with those derived in [9] up to decaying modes (which we neglect here) and a typo in their Eq. (116), which should have a minus sign for its second term.

Finally, at quartic order we need to subtract all fourth-order contributions from up to cubic bias,

$$\delta_g^{(4)} - b_1 \delta_{\text{GI}}^{(4)} - \frac{b_2}{2} [\delta_{\text{GI}}^2]^{(4)} - \gamma_2 [\mathcal{G}_2]^{(4)} - \frac{b_3}{6} [\delta_{\text{GI}}^3]^{(4)} - \gamma_3 [\mathcal{G}_3]^{(4)} - \gamma_2^\times [\delta_{\text{GI}} \mathcal{G}_2]^{(4)} - \gamma_{21} [\mathcal{G}_2(\varphi_2, \varphi_1)]^{(4)}, \quad (\text{C22})$$

and then, using that (again, constrained to Galilean invariant operators)

$$[\mathcal{G}_2]^{(4)} = \delta^2 \mathcal{G}_2 + \frac{2}{7} \mathcal{G}_2^2 - \frac{6}{7} \delta \mathcal{G}_2(\varphi_2, \varphi_1) + \frac{9}{49} \mathcal{G}_2(\varphi_2, \varphi_2) + 6 \mathcal{G}_2(\varphi_3, \varphi_1) + \delta \mathcal{G}_3 - \frac{15}{14} \mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1), \quad (\text{C23})$$

$$[\mathcal{G}_3]^{(4)} = \frac{9}{7} \mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1) + \delta \mathcal{G}_3,$$

$$[\mathcal{G}_2(\varphi_2, \varphi_1)]^{(4)} = -\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1) + \delta \mathcal{G}_2(\varphi_2, \varphi_1), \quad (\text{C24})$$

we can identify the evolved bias parameters from the coefficients of each of the fourth-order operators in our bias basis. We thus obtain

$$\begin{aligned} b_4 &= b_{4,\mathcal{L}} - 12b_2 - 8b_3, \\ \gamma_2^{\times\times} &= -\frac{3}{7}b_2 - \frac{1}{7}b_3 - \gamma_2^\times + \gamma_{2,\mathcal{L}}^{\times\times}, \\ \gamma_3^\times &= -\frac{b_2}{9} - \gamma_2^\times + \gamma_{3,\mathcal{L}}^\times, \end{aligned} \quad (\text{C25})$$

$$\begin{aligned} \gamma_2^{\text{sq}} &= -\frac{2}{49}b_2 - \frac{2}{7}\gamma_2^\times + \gamma_{2,\mathcal{L}}^{\text{sq}}, \\ \gamma_{21}^\times &= \frac{2}{21}b_2 + \frac{6}{7}\gamma_2^\times + \gamma_{21,\mathcal{L}}^\times, \\ \gamma_{31} &= -\frac{4}{11}b_{1,\mathcal{L}} - 6\gamma_2 + \gamma_{31,\mathcal{L}}, \end{aligned} \quad (\text{C26})$$

$$\begin{aligned} \gamma_{22} &= -\frac{6}{539}b_{1,\mathcal{L}} - \frac{9}{49}\gamma_2 + \gamma_{22,\mathcal{L}}, \\ \gamma_{211} &= \frac{5}{77}b_{1,\mathcal{L}} + \frac{15}{14}\gamma_2 + \gamma_{21} - \frac{9}{7}\gamma_3 + \gamma_{211,\mathcal{L}}. \end{aligned} \quad (\text{C27})$$

3. Spherical average limit

Let us now briefly consider the spherical-average limit of these results, which is somewhat subtle due to the fact that operators beyond local bias do not vanish under spherical average. As a result of this, the spherically averaged bias parameters b_n^{sph} are linear combinations of all the bias parameters (local and nonlocal) in our basis.

Consider quadratic bias, where the impact of spherical averaging first appears (linear bias is invariant under spherical average). From Eq. (C8) one can take the spherical average of the quadratic bias relation $(b_2/2)\delta^2 + \gamma_2 \mathcal{G}_2$ and write the spherically averaged quadratic bias parameter in terms of our bias parameters as

$$b_2^{\text{sph}} = b_2 - \frac{4}{3}\gamma_2 = b_{2,\mathcal{L}} - \frac{4}{3}\gamma_{2,\mathcal{L}} + \frac{8}{21}b_{1,\mathcal{L}} \equiv b_{2,\mathcal{L}}^{\text{sph}} + \frac{8}{21}b_{1,\mathcal{L}}^{\text{sph}}, \quad (\text{C28})$$

where in the second equality we have used our time evolution result from Eq. (C18) and then rewritten it in terms of the Lagrangian spherical averaged quadratic bias parameter, giving the well-known spherical collapse model bias evolution at quadratic order [73].

At cubic order, we see for the first time something subtler. The reason is that the spherical average of Eq. (C19) does not correspond to

$$\delta_g^{(3)} - b_1 \delta_{\text{SC}}^{(3)} - \frac{b_2^{\text{sph}}}{2} [\delta_{\text{SC}}^2]^{(3)} \quad (\text{C29})$$

due to the fact that at third order the spherically averaged Galileon operator obeys from Eqs. (C8) and (C20)

$$\overline{[\mathcal{G}_2]^{(3)}} = -\frac{52}{63}\delta^3 \neq -\frac{2}{3}\overline{[\delta_{\text{GI}}^2]^{(3)}} = -\frac{2}{3}\nu_2\delta^3 = -\frac{68}{63}\delta^3. \quad (\text{C30})$$

Therefore, at cubic order in our bias basis part of what would be considered b_3^{sph} is encoded by the deviation of the spherical average of $[\mathcal{G}_2]^{(3)}$ from that of $(-2/3)[\delta^2]^{(3)}$, the precise relation is then

$$b_3^{\text{sph}} = b_3 - 4\gamma_2^\times + \frac{4}{3}\gamma_3 + \frac{8}{3}\gamma_{21} + \frac{32}{21}\gamma_2. \quad (\text{C31})$$

Note this apparent dependence of b_3^{sph} on a quadratic bias parameter such as γ_2 only happens at the final conditions, since at the initial conditions matter fluctuations are linear and thus $[\mathcal{G}_2]^{(3)} = [\delta^2]^{(3)} = 0$. Therefore $\gamma_{2,\mathcal{L}}$ cannot appear in $b_{3,\mathcal{L}}^{\text{sph}}$, and, in fact, all this means that the only role of the γ_2 term in Eq. (C31) is to cancel the appearance of such terms in the evolved values of γ_3 and γ_{21} in Eq. (C21) and in b_3 and $\gamma_2^\times$ through $b_2 = b_2^{\text{sph}} + 4\gamma_2/3$. Indeed, evaluating Eq. (C31) yields

$$b_3^{\text{sph}} = b_{3,\mathcal{L}}^{\text{sph}} - \frac{13}{7}b_{2,\mathcal{L}}^{\text{sph}} - \frac{796}{1323}b_{1,\mathcal{L}}^{\text{sph}}, \quad (\text{C32})$$

where $b_{3,\mathcal{L}}^{\text{sph}} = b_{3,\mathcal{L}} - 4\gamma_{2,\mathcal{L}}^\times + (4/3)\gamma_{3,\mathcal{L}} + (8/3)\gamma_{21,\mathcal{L}}$. At quartic order, a similar subtlety arises with all nonlocal bias parameters of lower order than fourth. However, for the same reasons as in the cubic bias calculation, all such terms must cancel at the end. Therefore, a fast way of doing the calculation is to simply ignore all nonlocal quadratic and cubic bias parameter contributions to Eqs. (C25)–(C27) and in the relation between spherical bias parameters and local bias parameters. A quick calculation then gives

$$b_4^{\text{sph}} = b_{4,\mathcal{L}}^{\text{sph}} - \frac{40}{7} b_{3,\mathcal{L}}^{\text{sph}} + \frac{7220}{1323} b_{2,\mathcal{L}}^{\text{sph}} + \frac{476320}{305613} b_{1,\mathcal{L}}^{\text{sph}}, \quad (\text{C33})$$

where $b_{4,\mathcal{L}}^{\text{sph}} = b_{4,\mathcal{L}} - 16\gamma_{2,\mathcal{L}}^{\times\times} + (16/3)\gamma_{3,\mathcal{L}}^{\times} + (32/3)\gamma_{2,\mathcal{L}}^{\text{sq}} + (32/3)\gamma_{21,\mathcal{L}}^{\times} - (32/9)\gamma_{211,\mathcal{L}} - (64/9)\gamma_{22,\mathcal{L}} - (368/567)\gamma_{31,\mathcal{L}}$. The results in Eqs. (C32) and (C33) agree with those in [46] who corrected the cubic and quartic spherical collapse bias results in [73].

APPENDIX D: EVOLUTION OF RENORMALIZED BIAS

In Appendix C we discussed time evolution of bias parameters in terms of the standard expansion, as opposed to the bias parameters that appear in the multipoint-propagator

expansion, and therefore we ignored renormalization. One may ask whether those results are somehow changed due to renormalization. In fact, Ref. [12] contrasted the relation between renormalized and bare (final time) bias parameters to the time evolution of bare bias parameters found in [9], and appeared to suggest that the results of [9] may be affected by renormalization.

What matters, of course, is the time evolution of the *observable* bias parameters, i.e., how the renormalized late-time bias parameters are related to the renormalized initial-time bias parameters (this is what is often measured in numerical simulations). We therefore start from the initial conditions written using the multipoint propagator expansion up to fourth order,

$$\begin{aligned} \delta_g^{\text{IC}} = & b_{1,\mathcal{L}}\delta + \frac{b_{2,\mathcal{L}}}{2}[\delta^2 - \sigma^2] + \gamma_{2,\mathcal{L}}\mathcal{G}_2 + \frac{b_{3,\mathcal{L}}}{6}[\delta^3 - 3\delta\sigma^2] + \gamma_{3,\mathcal{L}}\mathcal{G}_3 + \gamma_{2,\mathcal{L}}^{\times}\left[\delta\mathcal{G}_2 + \frac{4}{3}\delta\sigma^2\right] + \gamma_{21,\mathcal{L}}\mathcal{G}_2(\varphi_2, \varphi_1) \\ & + \frac{b_{4,\mathcal{L}}}{24}[\delta^4 - 6\delta^2\sigma^2] + \gamma_{2,\mathcal{L}}^{\times\times}\left[\delta^2\mathcal{G}_2 + \frac{8}{3}\delta^2\sigma^2 - \mathcal{G}_2\sigma^2\right] + \gamma_{3,\mathcal{L}}^{\times}[\delta\mathcal{G}_3 + \mathcal{G}_2\sigma^2] + \gamma_{2,\mathcal{L}}^{\text{sq}}\left[\mathcal{G}_2^2 - \frac{32}{15}\delta^2\sigma^2 - \frac{8}{15}\mathcal{G}_2\sigma^2\right] \\ & + \gamma_{21,\mathcal{L}}^{\times}\left[\delta\mathcal{G}_2(\varphi_2, \varphi_1) - \frac{16}{5}\delta^2\sigma^2 + \frac{2}{5}\mathcal{G}_2\sigma^2\right] + \gamma_{211,\mathcal{L}}\mathcal{G}_3(\varphi_2, \varphi_1, \varphi_1) + \gamma_{22,\mathcal{L}}\mathcal{G}_2(\varphi_2, \varphi_2) + \gamma_{31,\mathcal{L}}\mathcal{G}_2(\varphi_3, \varphi_1) \\ & + \dots, \end{aligned} \quad (\text{D1})$$

where, unlike Eq. (78), the bias parameters here are renormalized, and we have neglected $\mathcal{O}(\sigma^4)$ contributions. Also, we have used that the loops in the multipoint propagators coming from nonlocal evolution operators cancel against the corresponding terms in the Gamma expansion to make the expression more compact.

The next step is to use Eq. (C16) to evolve this to find the late-time evolved δ_g and find the evolved renormalized bias parameters. To compare with and address the issue suggested by [12] we proceed as they did and write the nonlinear renormalization formulas for linear and quadratic bias parameters to $\mathcal{O}(\sigma^2)$, following the calculations we discussed in Sec. II C, that is,

$$b_1 = \bar{b}_1 + \left(\frac{1}{2}\bar{b}_3\right)\sigma^2 + \left(\frac{34}{21}\bar{b}_2\right)\sigma^2 + \mathcal{O}(\sigma^4) \quad (\text{D2})$$

for linear bias [8] and

$$b_2 = \bar{b}_2 + \left(\frac{1}{2}\bar{b}_4 - \frac{16}{3}\bar{\gamma}_2^{\times\times} + \frac{32}{15}\bar{\gamma}_{21}^{\times} + \frac{64}{15}\bar{\gamma}_2^{\text{sq}}\right)\sigma^2 + \left(\frac{8126}{2205}\bar{b}_2 - \frac{208}{35}\bar{\gamma}_2^{\times} + \frac{68}{21}\bar{b}_3\right)\sigma^2 + \mathcal{O}(\sigma^4), \quad (\text{D3})$$

$$\gamma_2 = \bar{\gamma}_2 + \left(\bar{\gamma}_2^{\times\times} - \frac{2}{5}\bar{\gamma}_{21}^{\times} - \bar{\gamma}_3^{\times} + \frac{8}{15}\bar{\gamma}_2^{\text{sq}}\right)\sigma^2 + \left(\frac{127}{2205}\bar{b}_2 + \frac{92}{105}\bar{\gamma}_2^{\times}\right)\sigma^2 + \mathcal{O}(\sigma^4) \quad (\text{D4})$$

for the quadratic bias parameters. Note that all bias parameters in Eqs. (D2)–(D4) are final-time quantities. The first parenthesis contributions in each of these expressions correspond to the renormalizations by operators that depend on the linear Gaussian fluctuations (that is, those that are straightforward to compute) and the second parenthesis contributions describe renormalizations induced by nonlinear evolution (discussed in Sec. II C) that were bypassed in the method we advocate elsewhere in the

main text. Apart from some sign typos and the renormalizations induced by $\bar{\gamma}_2^{\times}$ (what they call $b_{g_2\delta}^{(0)}$), Eqs. (D3) and (D4) agree with Eqs. (3.17) and (3.18) in [12].

Using Eq. (D1) in Eq. (C16), we can identify the evolved bare bias parameters as the coefficients of the operators in our bias basis up to quadratic order, and then renormalize according to Eqs. (D2)–(D4), finding the evolved renormalized bias parameters in terms of the initial renormalized bias parameters. In doing so we have taken into account all

renormalizations induced by operators up to fourth order (both in the initial bias parameters and in the final ones). As expected, all σ^2 dependencies go away (a nontrivial check on the nonlinear renormalization calculations in Eqs. (D3) and (D4), and we recover

$$b_1 = 1 + b_{1,\mathcal{L}}, \quad b_2 = b_{2,\mathcal{L}}, \quad \gamma_2 = -\frac{2}{7}b_{1,\mathcal{L}} + \gamma_{2,\mathcal{L}}, \quad (\text{D5})$$

i.e., the same results obtained from the evolution of the bare bias parameters, showing explicitly that the bare evolution results in [9] are immune to renormalization. Therefore

final-time renormalization formulas such as Eqs. (D2)–(D4) are not in conflict with time evolution of bare bias parameters, but rather guarantee that the time evolution results also hold for the renormalized bias parameters. In our approach advocated in the main text, the initial-time renormalization [equivalent to the first parenthesis in Eqs. (D2)–(D4)] is automatically carried out by the multi-point propagator expansion, while the second parenthesis renormalizations are never needed because time evolution of conserved tracers cannot generate such divergences (i.e., the evolved propagators already avoid spurious contributions proportional to σ^2).

-
- [1] N. Kaiser, On the spatial correlations of Abell clusters, *Astrophys. J.* **284**, L9 (1984).
 - [2] H. D. Politzer and M. B. Wise, Relations between spatial correlations of rich clusters of galaxies, *Astrophys. J.* **285**, L1 (1984).
 - [3] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, The statistics of peaks of Gaussian random fields, *Astrophys. J.* **304**, 15 (1986).
 - [4] H. J. Mo and S. D. M. White, An analytic model for the spatial clustering of dark matter haloes, *Mon. Not. R. Astron. Soc.* **282**, 347 (1996).
 - [5] R. K. Sheth and G. Tormen, Large-scale bias and the peak background split, *Mon. Not. R. Astron. Soc.* **308**, 119 (1999).
 - [6] V. Desjacques, D. Jeong, and F. Schmidt, Large-scale galaxy bias, *Phys. Rep.* **733**, 1 (2018).
 - [7] P. McDonald, Clustering of dark matter tracers: Renormalizing the bias parameters, *Phys. Rev. D* **74**, 103512 (2006).
 - [8] P. McDonald and A. Roy, Clustering of dark matter tracers: Generalizing bias for the coming era of precision LSS, *J. Cosmol. Astropart. Phys.* **08** (2009) 020.
 - [9] K. C. Chan, R. Scoccimarro, and R. K. Sheth, Gravity and large-scale nonlocal bias, *Phys. Rev. D* **85**, 083509 (2012).
 - [10] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, Evidence for quadratic tidal tensor bias from the halo bispectrum, *Phys. Rev. D* **86**, 083540 (2012).
 - [11] F. Schmidt, D. Jeong, and V. Desjacques, Peak-background split, renormalization, and galaxy clustering, *Phys. Rev. D* **88**, 023515 (2013).
 - [12] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, Renormalized halo bias, *J. Cosmol. Astropart. Phys.* **08** (2014) 056.
 - [13] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, Biased tracers and time evolution, *J. Cosmol. Astropart. Phys.* **07** (2015) 030.
 - [14] L. Fonseca de la Bella, D. Regan, D. Seery, and D. Parkinson, Impact of bias and redshift-space modelling for the halo power spectrum: Testing the effective field theory of large-scale structure, [arXiv:1805.12394](https://arxiv.org/abs/1805.12394).
 - [15] J. A. Frieman and E. Gaztanaga, The three-point function as a probe of models for large-scale structure, *Astrophys. J.* **425**, 392 (1994).
 - [16] E. Gaztanaga and J. A. Frieman, Bias and high-order galaxy correlation functions in the APM galaxy survey, *Astrophys. J.* **437**, L13 (1994).
 - [17] J. N. Fry, Gravity, Bias, and the Galaxy Three-Point Correlation Function, *Phys. Rev. Lett.* **73**, 215 (1994).
 - [18] S. Matarrese, L. Verde, and A. F. Heavens, Large-scale bias in the universe: Bispectrum method, *Mon. Not. R. Astron. Soc.* **290**, 651 (1997).
 - [19] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, Nonlinear evolution of the bispectrum of cosmological perturbations, *Astrophys. J.* **496**, 586 (1998).
 - [20] R. Scoccimarro, The bispectrum: From theory to observations, *Astrophys. J.* **544**, 597 (2000).
 - [21] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Cosmology and the bispectrum, *Phys. Rev. D* **74**, 023522 (2006).
 - [22] J. E. Pollack, R. E. Smith, and C. Porciani, Modelling large-scale halo bias using the bispectrum, *Mon. Not. R. Astron. Soc.* **420**, 3469 (2012).
 - [23] J. Bel, K. Hoffmann, and E. Gaztañaga, Non-local bias contribution to third-order galaxy correlations, *Mon. Not. R. Astron. Soc.* **453**, 259 (2015).
 - [24] S. Saito, T. Baldauf, Z. Vlah, U. Seljak, T. Okumura, and P. McDonald, Understanding higher-order nonlocal halo bias at large scales by combining the power spectrum with the bispectrum, *Phys. Rev. D* **90**, 123522 (2014).
 - [25] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, The power spectrum and bispectrum of SDSS DR11 BOSS galaxies—I. Bias and gravity, *Mon. Not. R. Astron. Soc.* **451**, 539 (2015).
 - [26] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres, and M. D. Olmstead, The clustering of galaxies in the SDSS-III baryon oscillation spectroscopic survey: RSD measurement from the power spectrum and bispectrum of the

- DR12 BOSS galaxies, *Mon. Not. R. Astron. Soc.* **465**, 1757 (2017).
- [27] P. Coles, Galaxy formation with a local bias, *Mon. Not. R. Astron. Soc.* **262**, 1065 (1993).
- [28] J. N. Fry and E. Gaztanaga, Biasing and hierarchical statistics in large-scale structure, *Astrophys. J.* **413**, 447 (1993).
- [29] P. Catelan, C. Porciani, and M. Kamionkowski, Two ways of biasing galaxy formation, *Mon. Not. R. Astron. Soc.* **318**, L39 (2000).
- [30] T. Fujita, V. Mauerhofer, L. Senatore, Z. Vlah, and R. Angulo, Very massive tracers and higher derivative biases, [arXiv:1609.00717](https://arxiv.org/abs/1609.00717).
- [31] E. O. Nadler, A. Perko, and L. Senatore, On the bispectra of very massive tracers in the effective field theory of large-scale structure, *J. Cosmol. Astropart. Phys.* **18** (2018) 058.
- [32] S. Pueblas and R. Scoccimarro, Generation of vorticity and velocity dispersion by orbit crossing, *Phys. Rev. D* **80**, 043504 (2009).
- [33] M. Pietroni, G. Mangano, N. Saviano, and M. Viel, Coarse-grained cosmological perturbation theory *J. Cosmol. Astropart. Phys.* **12** (2012) 019.
- [34] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, Cosmological non-linearities as an effective fluid, *J. Cosmol. Astropart. Phys.* **12** (2012) 051.
- [35] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, The effective field theory of cosmological large scale structures, *J. High Energy Phys.* **12** (2012) 82.
- [36] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer, The bispectrum in the effective field theory of large scale structure, *J. Cosmol. Astropart. Phys.* **15** (2015) 007.
- [37] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore, The one-loop matter bispectrum in the effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **15** (2015) 039.
- [38] R. J. Scherrer and D. H. Weinberg, Constraints on the effects of locally biased galaxy formation, *Astrophys. J.* **504**, 607 (1998).
- [39] A. Dekel and O. Lahav, Stochastic nonlinear galaxy biasing, *Astrophys. J.* **520**, 24 (1999).
- [40] M. Crocce and R. Scoccimarro, Renormalized cosmological perturbation theory, *Phys. Rev. D* **73**, 063519 (2006).
- [41] F. Bernardeau, M. Crocce, and R. Scoccimarro, Multipoint propagators in cosmological gravitational instability, *Phys. Rev. D* **78**, 103521 (2008).
- [42] F. Bernardeau, M. Crocce, and E. Sefusatti, Multipoint propagators for non-Gaussian initial conditions, *Phys. Rev. D* **82**, 083507 (2010).
- [43] F. Bernardeau, M. Crocce, and R. Scoccimarro, Constructing regularized cosmic propagators, *Phys. Rev. D* **85**, 123519 (2012).
- [44] M. Crocce, R. Scoccimarro, and F. Bernardeau, MPTBREEZE: A fast renormalized perturbative scheme, *Mon. Not. R. Astron. Soc.* **427**, 2537 (2012).
- [45] A. Taruya, T. Nishimichi, and F. Bernardeau, Precision modeling of redshift-space distortions from a multipoint propagator expansion, *Phys. Rev. D* **87**, 083509 (2013).
- [46] T. Matsubara, Nonlinear perturbation theory integrated with nonlocal bias, redshift-space distortions, and primordial non-Gaussianity, *Phys. Rev. D* **83**, 083518 (2011).
- [47] T. Matsubara, Integrated perturbation theory and one-loop power spectra of biased tracers, *Phys. Rev. D* **90**, 043537 (2014).
- [48] T. Matsubara and V. Desjacques, Impacts of biasing schemes in the one-loop integrated perturbation theory, *Phys. Rev. D* **93**, 123522 (2016).
- [49] R. E. Smith, R. Scoccimarro, and R. K. Sheth, Scale dependence of halo and galaxy bias: Effects in real space, *Phys. Rev. D* **75**, 063512 (2007).
- [50] M. Crocce and R. Scoccimarro, Memory of initial conditions in gravitational clustering, *Phys. Rev. D* **73**, 063520 (2006).
- [51] A. S. Szalay, Constraints on the biasing of density fluctuations, *Astrophys. J.* **333**, 21 (1988).
- [52] Y. P. Jing, Accurate determination of the Lagrangian bias for the dark matter halos, *Astrophys. J.* **515**, L45 (1999).
- [53] R. K. Sheth and G. Lemson, Biasing and the distribution of dark matter haloes, *Mon. Not. R. Astron. Soc.* **304**, 767 (1999).
- [54] C. Porciani, P. Catelan, and C. Lacey, How dark matter Halos cluster in Lagrangian space, *Astrophys. J.* **513**, L99 (1999).
- [55] A. Elia, A. D. Ludlow, and C. Porciani, The spatial and velocity bias of linear density peaks and protohaloes in the Λ cold dark matter cosmology, *Mon. Not. R. Astron. Soc.* **421**, 3472 (2012).
- [56] R. K. Sheth, K. C. Chan, and R. Scoccimarro, Nonlocal Lagrangian bias, *Phys. Rev. D* **87**, 083002 (2013).
- [57] M. Biagetti, K. Chuen Chan, V. Desjacques, and A. Paranjape, Measuring non-local Lagrangian peak bias, *Mon. Not. R. Astron. Soc.* **441**, 1457 (2014).
- [58] K. C. Chan, R. K. Sheth, and R. Scoccimarro, Effective window function for Lagrangian halos, *Phys. Rev. D* **96**, 103543 (2017).
- [59] C. Modi, E. Castorina, and U. Seljak, Halo bias in Lagrangian space: Estimators and theoretical predictions, *Mon. Not. R. Astron. Soc.* **472**, 3959 (2017).
- [60] T. Lazeyras and F. Schmidt, Beyond LIMD bias: A measurement of the complete set of third-order halo bias parameters, *J. Cosmol. Astropart. Phys.* **18** (2018) 008.
- [61] M. M. Abidi and T. Baldauf, Cubic halo bias in Eulerian and Lagrangian space, *J. Cosmol. Astropart. Phys.* **07** (2018) 029.
- [62] F. Schmidt, F. Elsner, J. Jasche, N. Minh Nguyen, and G. Lavaux, A rigorous EFT-based forward model for large-scale structure, *J. Cosmol. Astropart. Phys.* **19** (2019) 042.
- [63] M. Schmittfull, M. Simonović, V. Assassi, and M. Zaldarriaga, Modeling biased tracers at the field level, [arXiv:1811.10640](https://arxiv.org/abs/1811.10640).
- [64] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Large-scale structure of the Universe and cosmological perturbation theory, *Phys. Rep.* **367**, 1 (2002).
- [65] F. R. Bouchet, R. Juszkiewicz, S. Colombi, and R. Pellat, Weakly nonlinear gravitational instability for arbitrary Omega, *Astrophys. J.* **394**, L5 (1992).
- [66] F. R. Bouchet, S. Colombi, E. Hivon, and R. Juszkiewicz, Perturbative Lagrangian approach to gravitational instability, *Astron. Astrophys.* **296**, 575 (1995).
- [67] J. N. Fry, The galaxy correlation hierarchy in perturbation theory, *Astrophys. J.* **279**, 499 (1984).

- [68] M. H. Goroff, B. Grinstein, S.-J. Rey, and M. B. Wise, Coupling of modes of cosmological mass density fluctuations, *Astrophys. J.* **311**, 6 (1986).
- [69] J. N. Fry, The evolution of bias, *Astrophys. J.* **461**, L65 (1996).
- [70] L. Senatore, Bias in the effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **15** (2015) 007.
- [71] R. Scoccimarro and J. Frieman, Loop corrections in nonlinear cosmological perturbation theory, *Astrophys. J. Suppl.* **105**, 37 (1996).
- [72] K. M. Hoffman and R. Kunze, *Linear Algebra* (Prentice-Hall, Englewood Cliffs, 1971).
- [73] H. J. Mo, Y. P. Jing, and S. D. M. White, High-order correlations of peaks and haloes: A step towards understanding galaxy biasing, *Mon. Not. R. Astron. Soc.* **284**, 189 (1997).
- [74] A. Paranjape and R. K. Sheth, Peaks theory and the excursion set approach, *Mon. Not. R. Astron. Soc.* **426**, 2789 (2012).
- [75] L. Kofman and D. Pogosyan, Dynamics of gravitational instability is nonlocal, *Astrophys. J.* **442**, 30 (1995).
- [76] A. Eggemeier, R. Scoccimarro, M. Crocce, A. G. Sanchez, and R. E. Smith (to be published).
- [77] T. Buchert, Lagrangian perturbation theory—A key-model for large-scale structure, *Astron. Astrophys.* **267**, L51 (1993).
- [78] T. Buchert, Lagrangian theory of gravitational instability of Friedman-Lemaître cosmologies—A generic third-order model for nonlinear clustering, *Mon. Not. R. Astron. Soc.* **267**, 811 (1994).
- [79] T. Matsubara, Diagrammatic methods in statistics and biasing in the large-scale structure of the universe, *Astrophys. J. Suppl.* **101**, 1 (1995).
- [80] A. Paranjape, E. Sefusatti, K. Chuen Chan, V. Desjacques, P. Monaco, and R. K. Sheth, Bias deconstructed: Unravelling the scale dependence of halo bias using real-space measurements, *Mon. Not. R. Astron. Soc.* **436**, 449 (2013).
- [81] M. Schmittfull, T. Baldauf, and U. Seljak, Near optimal bispectrum estimators for large-scale structure, *Phys. Rev. D* **91**, 043530 (2015).
- [82] K. Hoffmann, J. Bel, and E. Gaztañaga, Linear and nonlinear bias: Predictions versus measurements, *Mon. Not. R. Astron. Soc.* **465**, 2225 (2017).
- [83] M. Biagetti, V. Desjacques, A. Kehagias, and A. Riotto, Halo velocity bias, *Phys. Rev. D* **90**, 103529 (2014).
- [84] P. Catelan, F. Lucchin, S. Matarrese, and C. Porciani, The bias field of dark matter haloes, *Mon. Not. R. Astron. Soc.* **297**, 692 (1998).
- [85] R. Scoccimarro, Transients from initial conditions: A perturbative analysis, *Mon. Not. R. Astron. Soc.* **299**, 1097 (1998).
- [86] F. Schmidt, D. Jeong, and V. Desjacques, Peak-background split, renormalization, and galaxy clustering, *Phys. Rev. D* **88**, 023515 (2013).
- [87] A. Aviles, Renormalization of Lagrangian bias via spectral parameters, *Phys. Rev. D* **98**, 083541 (2018).
- [88] E. Pajer and M. Zaldarriaga, On the renormalization of the effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **08** (2013) 37.
- [89] D. Blas, M. Garny, and T. Konstandin, Cosmological perturbation theory at three-loop order *J. Cosmol. Astropart. Phys.* **14** (2014) 010.
- [90] F. Bernardeau, A. Taruya, and T. Nishimichi, Cosmic propagators at two-loop order, *Phys. Rev. D* **89**, 023502 (2014).
- [91] T. Baldauf, L. Mercolli, and M. Zaldarriaga, The effective field theory of large scale structure at two loops: The apparent scale dependence of the speed of sound, *Phys. Rev. D* **92**, 123007 (2015).
- [92] R. Scoccimarro, Cosmological perturbations: Entering the nonlinear regime, *Astrophys. J.* **487**, 1 (1997).
- [93] T. Lazeyras, C. Wagner, T. Baldauf, and F. Schmidt, Precision measurement of the local bias of dark matter halos, *J. Cosmol. Astropart. Phys.* **16** (2016) 018.
- [94] G. D’Amico, M. Crocce, and R. Scoccimarro (to be published).
- [95] S. Pueblas, Dark matter clustering in precision cosmology, Ph.D. thesis, New York University, 2008.
- [96] V. Desjacques, M. Crocce, R. Scoccimarro, and R. K. Sheth, Modeling scale-dependent bias on the baryonic acoustic scale with the statistics of peaks of Gaussian random fields, *Phys. Rev. D* **82**, 103529 (2010).
- [97] A. F. Heavens, S. Matarrese, and L. Verde, The non-linear redshift-space power spectrum of galaxies, *Mon. Not. R. Astron. Soc.* **301**, 797 (1998).
- [98] F. Schmidt, Towards a self-consistent halo model for the nonlinear large-scale structure, *Phys. Rev. D* **93**, 063512 (2016).
- [99] D. Ginzburg, V. Desjacques, and K. Chuen Chan, Shot noise and biased tracers: A new look at the halo model, *Phys. Rev. D* **96**, 083528 (2017).
- [100] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980).
- [101] M. Biagetti, V. Desjacques, A. Kehagias, and A. Riotto, Nonlocal halo bias with and without massive neutrinos, *Phys. Rev. D* **90**, 045022 (2014).
- [102] A. G. Sánchez *et al.*, The clustering of galaxies in the completed SDSS-III baryon oscillation spectroscopic survey: Cosmological implications of the configuration-space clustering wedges, *Mon. Not. R. Astron. Soc.* **464**, 1640 (2017).
- [103] R. Angulo, M. Fasiello, L. Senatore, and Z. Vlah, On the statistics of biased tracers in the effective field theory of large scale structures, *J. Cosmol. Astropart. Phys.* **15** (2015) 029.
- [104] J. Carlson, B. Reid, and M. White, Convolution Lagrangian perturbation theory for biased tracers, *Mon. Not. R. Astron. Soc.* **429**, 1674 (2013).
- [105] Z. Vlah, E. Castorina, and M. White, The Gaussian streaming model and convolution Lagrangian effective field theory, *J. Cosmol. Astropart. Phys.* **16** (2016) 007.
- [106] F. Bernardeau, The gravity-induced quasi-Gaussian correlation hierarchy, *Astrophys. J.* **392**, 1 (1992).