

PAPER • OPEN ACCESS

Multi-tracer power spectra and bispectra: formalism

To cite this article: Dionysios Karagiannis *et al* JCAP03(2024)034

View the [article online](#) for updates and enhancements.

You may also like

- [A brief review on cosmological analysis of galaxy surveys with multiple tracers](#)
Yuting Wang and Gong-Bo Zhao
- [Fast Phase Spectrum Estimation Using the Parallel PartBispectrum Algorithm](#)
David W. Tyler and Kathy J. Schulze
- [The 21 cm kSZ–kSZ Bispectrum during the Epoch of Reionization](#)
Paul La Plante, Adam Lidz, James Aguirre et al.

Multi-tracer power spectra and bispectra: formalism

Dionysios Karagiannis^{ID},^a Roy Maartens,^{a,b,c} José Fonseca,^{d,e,a} Stefano Camera^{f,g,h,a}
and Chris Clarkson^{i,a}

^aDepartment of Physics & Astronomy, University of the Western Cape,
Cape Town 7535, South Africa

^bInstitute of Cosmology & Gravitation, University of Portsmouth,
Portsmouth PO1 3FX, United Kingdom

^cNational Institute for Theoretical and Computational Sciences (NITheCS),
Cape Town 7535, South Africa

^dInstituto de Astrofísica e Ciências do Espaço, Universidade do Porto CAUP,
Porto 4150-762, Portugal

^eDepartamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto,
Rua do Campo Alegre 687, Porto PT4169-007, Portugal

^fDipartimento di Fisica, Università degli Studi di Torino,
Torino 10125, Italy

^gINFN — Istituto Nazionale di Fisica Nucleare, Sezione di Torino,
Torino 10125, Italy

^hINAF — Istituto Nazionale di Astrofisica, Osservatorio Astrofisico di Torino,
Pino Torinese 10025, Italy

ⁱDepartment of Physics & Astronomy, Queen Mary University of London,
London E1 4NS, United Kingdom

E-mail: dakaragian@gmail.com, roy.maartens@port.ac.uk,
jose.fonseca@astro.up.pt, stefano.camera@unito.it,
chris.clarkson@qmul.ac.uk

ABSTRACT: The power spectrum and bispectrum of dark matter tracers are key and complementary probes of the Universe. Next-generation surveys will deliver good measurements of the bispectrum, opening the door to improved cosmological constraints and the breaking of parameter degeneracies, from the combination of the power spectrum and bispectrum. Multi-tracer power spectra have been used to suppress cosmic variance and mitigate the effects of nuisance parameters and systematics. We present a bispectrum multi-tracer formalism that can be applied to next-generation survey data. Then we perform a simple Fisher analysis to illustrate qualitatively the improved precision on primordial non-Gaussianity that is expected to come from the bispectrum multi-tracer. In addition, we investigate the parametric dependence of conditional errors from multi-tracer power spectra and multi-tracer bispectra, on the

differences between the biases and the number densities of two tracers. Our results suggest that optimal constraints arise from maximising the ratio of number densities, the difference between the linear biases, the difference between the quadratic biases, and the difference between the products $b_1 b_\Phi$ for each tracer, where b_Φ is the bias for the primordial potential.

KEYWORDS: cosmological parameters from LSS, galaxy surveys, inflation, power spectrum

ARXIV EPRINT: [2305.04028](https://arxiv.org/abs/2305.04028)

Contents

1	Introduction	1
2	Single-tracer power spectrum and bispectrum	2
3	Multi-tracer power spectra and bispectra	4
4	Multi-tracer covariance	6
5	Fisher information on f_{NL}	9
6	Summary and conclusion	20
A	Redshift-space kernels and damping factors	22
B	3-tracer case	23

1 Introduction

Spectacular advances in observations and computations have transformed our understanding of the Universe. At the same time, this has brought new puzzles to the fore, such as the still unknown nature of dark matter and dark energy that are core to the current standard model, and the emergence of tensions between observations that appear to challenge the foundations of the standard model. The next generation of large-scale structure surveys will deliver major advances in precision cosmology, cast new light on current tensions and anomalies, and open new windows on the primordial and late-time Universe.

Up to now, cosmological information has mainly been extracted via the galaxy power spectrum, but the galaxy bispectrum is set to play an increasingly important role, in combination with the power spectrum (see e.g. the recent work [1–7] and references therein). Recent measurements of the BOSS bispectrum [2, 3, 5], including the first measurement of the bispectrum quadrupole [8], prepare the way for next-generation surveys. These upcoming surveys are also predicted to detect the bispectrum dipole [9–11] and other relativistic effects in the bispectrum [12–15]. Wide-angle effects in the galaxy bispectrum [15–18] will also need to be taken into account.

The bispectrum is a direct probe of non-Gaussianity, and one of its key strengths lies in constraints on primordial non-Gaussianity (PNG). Unlike the power spectrum, it can constrain non-local types of PNG, while in combination with the power spectrum it can improve constraints on local PNG [19–24]. In order to further improve the precision of power spectrum constraints, one can use a multi-tracer analysis, which suppresses cosmic variance and also reduces the effects of nuisance parameters and systematics (see e.g. [25–28]). A multi-tracer analysis is particularly effective for constraints on local PNG and other quantities that have a signal on very large scales, where cosmic variance is strongest (see e.g. [29–41]).

This suggests the use of a multi-tracer bispectrum analysis, in combination with the corresponding multi-tracer power spectra. Here we develop for the first time a Fisher forecast multi-tracer formalism for the tree-level power spectra + bispectra combination. Multi-tracer power spectra have been combined with the auto-bispectrum of one of the tracers by [38]. We extend such approaches to include all the bispectra. A multi-tracer forecast has been applied to bispectra alone (in real space) by [42], on the basis of simplifying assumptions and without giving the covariances or other details. We fill the gap by providing a systematic analysis in redshift-space, including all clustering bias and noise parameters at tree level and giving explicit forms for the covariances.

In this first paper, we do not perform detailed forecasts, which will be presented in future work. However, we do show qualitatively the improvements that are potentially delivered by the full power spectrum and bispectrum multi-tracer of 2 tracers, by investigating the parametric behaviour of the conditional errors on PNG. To quantify the potential of this method we consider a comparison between the forecasts generated by the multi-tracer and each single-tracer, over a single volume and with the same redshift range. This is done in order to investigate the effect of combining the samples or not on primordial non-Gaussianity, without the effect of survey specifications for the different tracers, e.g. new volumes and different redshift overlap. To begin with, we do not include wide-angle or relativistic effects.

2 Single-tracer power spectrum and bispectrum

The power spectrum of the Bardeen gauge-invariant primordial gravitational potential is defined in Fourier space by

$$\langle \Phi(\mathbf{k}) \Phi(\mathbf{k}') \rangle = (2\pi)^3 \delta_{\mathbf{D}}(\mathbf{k} + \mathbf{k}') P_{\Phi}(k), \quad (2.1)$$

where $P_{\Phi}(k)$ is directly related to the power spectrum of the primordial curvature perturbations ζ (during the matter-dominated era, $\Phi = 3\zeta/5$), which are generated during inflation. In the case of the standard single-field slow-roll inflationary scenario, they have a nearly perfect Gaussian distribution, which means that they can be adequately characterised by their power spectrum. The primordial perturbations Φ are in turn related to the linear dark matter over-density field through the Poisson equation, $\delta(\mathbf{k}, z) = M(k, z) \Phi(\mathbf{k})$, where

$$P(k, z) = M(k, z)^2 P_{\Phi}(k),$$

$$M(k, z) = \frac{2D(z)}{3\Omega_{\text{m},0} H_0^2 g_{\text{dec}}} T(k) k^2. \quad (2.2)$$

Here $D(z)$ is the linear growth factor (since the linear fluid equations generate a linearly evolved matter density field), normalised to unity at $z = 0$, and $T(k)$ is the matter transfer function normalized to unity at large scales $k \rightarrow 0$. The factor g_{dec} is the Bardeen potential growth factor at decoupling, which ensures that f_{NL} is in the CMB convention [43, 44]. The linear power spectrum is computed with the numerical Boltzmann code CAMB [45].

Deviations from the standard inflationary model will produce a violation of Gaussian initial conditions [46]. PNG generates nonzero higher-order correlators in the primordial

curvature perturbations, starting with the bispectrum, i.e. the Fourier transform of the three-point function,

$$\langle \Phi(\mathbf{k}_1) \Phi(\mathbf{k}_2) \Phi(\mathbf{k}_3) \rangle = (2\pi)^3 \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_{\Phi}(k_1, k_2, k_3). \quad (2.3)$$

Higher-order correlators in the primordial density field are a product of nonlinear interactions during the inflationary or reheating stage. The primordial bispectrum arising from such interactions is characterized by a dimensionless amplitude parameter f_{NL} , giving the strength of the PNG signal, and a shape function F , i.e. $B_{\Phi}(k_1, k_2, k_3) = f_{\text{NL}} F(k_1, k_2, k_3)$. F describes the dependence of the bispectrum on various triangles, where different inflationary models generate signal that peaks at distinct triangle configurations. By eq. (2.2), the leading order PNG contribution to the matter density bispectrum is

$$B_{\text{PNG}}(k_1, k_2, k_3, z) = M(k_1, z) M(k_2, z) M(k_3, z) B_{\Phi}(k_1, k_2, k_3). \quad (2.4)$$

The matter bispectrum has additional contributions from nonlinearity induced by gravity, even at zeroth order. In other words, the dark matter density field is intrinsically non-Gaussian, even in the absence of PNG. Therefore, in order to disentangle the PNG information from late-time nonlinearity, an adequate description for the latter is needed. Here we model nonlinearity in the framework of Standard Perturbation Theory (SPT) (see e.g. [47] for a review). In addition to the nonlinear nature of the dark matter density field, the hierarchical formation of halos and the way they are populated by luminous matter induces non-Gaussianity in the distribution of observable structures. Hence, in order to extract the PNG signal from large-scale structure, we need the bias relationship between the observed tracers and the matter field.

We follow the perturbative approach of [48–50], in which the halo over-density field δ_h is described as a function of all possible local gravitational observables, which are introduced in the expansion in the form of renormalised operators. A complete Eulerian bias expansion can be built from the tensor $\partial_i \partial_j \Phi$, which contains the trace $\nabla^2 \Phi \propto \delta$ and the trace-free tidal field $s_{ij} = (\partial_i \partial_j - \delta_{ij} \nabla^2 / 3) \nabla^{-2} \delta$. Note that the term ‘halo’ here can be replaced by a general matter tracer.

We consider up to second-order terms in the expansion, which are sufficient for the spatial scales considered here, i.e. much larger scales than those involved in halo formation. For Gaussian initial conditions, the Eulerian halo density contrast is [44, 48–50]:

$$\delta^h(\mathbf{x}, \tau) = b_1(\tau) \delta(\mathbf{x}, \tau) + \varepsilon(\mathbf{x}, \tau) + \frac{b_2(\tau)}{2} \delta(\mathbf{x}, \tau)^2 + \frac{b_s(\tau)}{2} s(\mathbf{x}, \tau)^2 + \varepsilon_{\delta}(\mathbf{x}, \tau) \delta(\mathbf{x}, \tau). \quad (2.5)$$

Here τ is conformal time, \mathbf{x} are spatial comoving Eulerian coordinates, $s^2 = s_{ij} s^{ij}$ is the simplest scalar that can be formed from the tidal field, ε is the leading stochastic field [51–53] and ε_{δ} is the stochastic field associated with the linear bias. These fields take into account the stochastic relation between the tracer density and any large-scale field. The second-order tidal field bias coefficient, following the convention in [54], is given by $b_s = -4(b_1 - 1)/7$.

In the presence of PNG there is a scale-dependent correction to the linear bias b_1 , especially in the case of local PNG [55–60], since the local PNG bispectrum peaks in squeezed triangles.

(A similar scale-dependent bias correction can be derived for any general nonlocal quadratic non-Gaussianity template [61–64].) Then the non-Gaussian bias terms, linear in f_{NL} , are [65]:

$$\delta_{\text{PNG}}^h(\mathbf{x}, \tau) = b_{\Psi}(\tau) \Psi(\mathbf{q}) + b_{\Psi\delta}(\tau) \Psi(\mathbf{q}) \delta(\mathbf{x}, \tau) + \varepsilon_{\Psi}(\mathbf{x}, \tau) \Psi(\mathbf{q}), \quad (2.6)$$

where \mathbf{q} are the spatial coordinates in the Lagrangian frame and Ψ is a nonlocal transformation of Φ :

$$\Psi(\mathbf{q}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} k^\alpha \Phi(\mathbf{k}) e^{i\mathbf{q}\cdot\mathbf{k}}. \quad (2.7)$$

The parameter $\alpha = 2, 1, 0$ for equilateral, orthogonal and local shapes. Note that in the local case, $\Psi \rightarrow \Phi$. The stochastic counterpart of Ψ is ε_{Ψ} . The bias coefficients in eqs. (2.5) and (2.6) can be derived by using the peak-background split argument (see e.g. [21]).

The single-tracer power spectrum and bispectrum in redshift space are

$$P^h(\mathbf{k}, z) = \mathcal{D}_P(\mathbf{k}, z) Z_1(\mathbf{k}, z)^2 P(k, z) + P_{\varepsilon}(z), \quad (2.8)$$

$$B^h(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, z) = \mathcal{D}_B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, z) \left[2 Z_1(\mathbf{k}_1, z) Z_1(\mathbf{k}_2, z) Z_2(\mathbf{k}_1, \mathbf{k}_2, z) P(k_1, z) P(k_2, z) \right. \\ \left. + 2 \text{perm.} \right] + 2 P_{\varepsilon\varepsilon\delta}(z) \left[Z_1(\mathbf{k}_1, z) P(k_1, z) + 2 \text{perm.} \right] + B_{\varepsilon}(z). \quad (2.9)$$

At tree level, the first- and second-order kernels $Z_{1,2}$ include redshift space distortions (RSD) [66–68] and PNG. In order to deal with redshift errors and nonlinear RSD, we include the phenomenological factors $\mathcal{D}_{P,B}$. Explicit expressions are in appendix A. In the case of power spectra, an alternative to the phenomenological factor \mathcal{D}_P is to use the 1-loop EFT model, as in the multi-tracer power spectrum analysis of [69]. The stochastic terms P_{ε} , $P_{\varepsilon\varepsilon\delta}$, B_{ε} are generated by stochastic bias and their fiducial values are taken to be those predicted by Poisson statistics [70]:

$$P_{\varepsilon} = \frac{1}{\bar{n}_h}, \quad P_{\varepsilon\varepsilon\delta} = \frac{b_1}{2\bar{n}_h}, \quad B_{\varepsilon} = \frac{1}{\bar{n}_h^2}. \quad (2.10)$$

3 Multi-tracer power spectra and bispectra

In the multi-tracer framework (e.g. [25–41]) we consider the combined information of all correlations of different tracer samples in a given redshift bin and sky patch. The multi-tracer power spectra and bispectra are

$$\langle \delta^I(\mathbf{k}_1) \delta^J(\mathbf{k}_2) \rangle = (2\pi)^3 \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2) P^{IJ}(\mathbf{k}_1), \quad (3.1)$$

$$\langle \delta^I(\mathbf{k}_1) \delta^J(\mathbf{k}_2) \delta^K(\mathbf{k}_3) \rangle = (2\pi)^3 \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B^{IJK}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \quad (3.2)$$

where here and elsewhere we omit the z -dependence for brevity. The tracer indices I, J, K range over the tracer labels: $I, J, K = t, t', t'', \dots$. Note that the cross-power spectra are symmetric under permutations of the tracers, while the cross-bispectra require symmetrisation in order to avoid double counting of information:

$$P^{(IJ)} \equiv P^{IJ}, \quad (3.3)$$

$$B^{(IJK)} = \frac{1}{6} \left(B^{IJK} + B^{KJI} + B^{KIJ} + B^{JIK} + B^{JKI} + B^{IKJ} \right). \quad (3.4)$$

We will also use the short-hand notation

$$P^I \equiv P^{II}, \quad B^I \equiv B^{III}, \quad (3.5)$$

where convenient.

We define the estimators:

$$\hat{P}^{IJ}(\mathbf{k}) = \frac{1}{V_s V_k} \sum_{\mathbf{q} \in k} \delta^I(\mathbf{q}) \delta^J(-\mathbf{q}), \quad (3.6)$$

$$\hat{B}^{(IJK)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{V_s V_{123}} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \delta^I(\mathbf{q}_1) \delta^J(\mathbf{q}_2) \delta^K(\mathbf{q}_3), \quad (3.7)$$

where $\mathbf{q}_{123} \equiv \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$. For the power spectra the sum runs over all wavenumbers \mathbf{q} in the spherical shell of radius k and width Δk . For the bispectra, the Kronecker delta symbol ensures the formation of ‘fundamental triangles’ with sides \mathbf{q}_i , satisfying the condition $\mathbf{q}_{123} = 0$, that fall in the ‘triangular bin’ defined by the triplet of bin centres (k_1, k_2, k_3) and width Δk . The volume in Fourier space of the power spectra shells and the bispectra fundamental triangle bins is, in the thin shell limit [71],

$$V_k = 4\pi k^2 \Delta k, \quad V_{123} \equiv \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) = 8\pi^2 k_1 k_2 k_3 \Delta k_1 \Delta k_2 \Delta k_3. \quad (3.8)$$

The estimators in eqs. (3.6) and (3.7) in the case of one sample ($I = J = K$) reduce to the well known auto-correlation results.

The two-tracer case. From now on we consider the case of two tracers, t and t' , so that $I, J, K = t, t'$. The data vector of the power spectra is

$$\mathbf{D}_P = [P^{tt}, P^{tt'}, P^{t't'}]. \quad (3.9)$$

The expected values of the spectra are:

$$\langle \hat{P}^{IJ} \rangle \equiv P^{IJ}(\mathbf{k}, z) = \mathcal{D}_P^{IJ}(\mathbf{k}, z) Z_1^I(\mathbf{k}, z) Z_1^J(\mathbf{k}, z) P(k, z) + \delta^{IJ} P_\epsilon^I(z), \quad (3.10)$$

where Z_1^I and \mathcal{D}_P^{IJ} are given in appendix A.

The data vector of the bispectrum is

$$\mathbf{D}_B = [B^{ttt}, B^{(ttt')}, B^{(tt't')}, B^{t't't'}]. \quad (3.11)$$

In a tree-level description, we consider terms up to second order PT theory, RSD and bias, while keeping only linear f_{NL} terms in the perturbative expansion. Then the expected values of the estimator for each bispectrum data vector entry are:

$$\begin{aligned} \langle \hat{B}^{ttt} \rangle \equiv B^{ttt}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \mathcal{D}_B^{ttt} \left[Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) B_{\text{PNG}}(k_1, k_2, k_3) \right. \\ &\quad \left. + \left\{ 2 Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_2^t(\mathbf{k}_1, \mathbf{k}_2) P(k_1) P(k_2) + 2 \mathbf{k}\text{-perm.} \right\} \right] \\ &\quad + 2 P_{\epsilon\epsilon\delta}^t \left[Z_1^t(\mathbf{k}_1) P(k_1) + 2 \mathbf{k}\text{-perm.} \right] + B_\epsilon^t, \end{aligned} \quad (3.12)$$

$$\begin{aligned}
 \langle \hat{B}^{(ttt')} \rangle \equiv B^{(ttt')}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \frac{1}{3} \left[\left\{ \mathcal{D}_B^{ttt'} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) + \mathcal{D}_B^{tt't} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) \right. \right. \\
 &\quad \left. \left. + \mathcal{D}_B^{t'tt} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) \right\} B_{\text{PNG}}(k_1, k_2, k_3) \right. \\
 &\quad \left. + 2 \left\{ [\mathcal{D}_B^{tt't} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) + \mathcal{D}_B^{t'tt} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2)] Z_2^t(\mathbf{k}_1, \mathbf{k}_2) \right. \right. \\
 &\quad \left. \left. + \mathcal{D}_B^{ttt'} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_2^t(\mathbf{k}_1, \mathbf{k}_2) \right\} P(k_1) P(k_2) + 2 \mathbf{k}\text{-perm.} \right] \\
 &\quad + \frac{1}{3} P_{\varepsilon\varepsilon\delta}^t \left[Z^t(\mathbf{k}_1) P(k_1) + 2 \mathbf{k}\text{-perm.} \right], \tag{3.13}
 \end{aligned}$$

$$\begin{aligned}
 \langle \hat{B}^{(t't't')} \rangle \equiv B^{(t't't')}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \frac{1}{3} \left[\left\{ \mathcal{D}_B^{t't't'} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) + \mathcal{D}_B^{t't't} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) \right. \right. \\
 &\quad \left. \left. + \mathcal{D}_B^{t'tt'} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) \right\} B_{\text{PNG}}(k_1, k_2, k_3) \right. \\
 &\quad \left. + 2 \left\{ [\mathcal{D}_B^{t't't'} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) + \mathcal{D}_B^{t'tt'} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2)] Z_2^t(\mathbf{k}_1, \mathbf{k}_2) \right. \right. \\
 &\quad \left. \left. + \mathcal{D}_B^{t't't} Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_2^t(\mathbf{k}_1, \mathbf{k}_2) \right\} P(k_1) P(k_2) + 2 \mathbf{k}\text{-perm.} \right] \\
 &\quad + \frac{1}{3} P_{\varepsilon\varepsilon\delta}^{t'} \left[Z^t(\mathbf{k}_1) P(k_1) + 2 \mathbf{k}\text{-perm.} \right], \tag{3.14}
 \end{aligned}$$

$$\begin{aligned}
 \langle \hat{B}^{t't't't'} \rangle \equiv B^{t't't't'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \mathcal{D}_B^{t't't't'} \left[Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_1^t(\mathbf{k}_3) B_{\text{PNG}}(k_1, k_2, k_3) \right. \\
 &\quad \left. + \left\{ 2 Z_1^t(\mathbf{k}_1) Z_1^t(\mathbf{k}_2) Z_2^t(\mathbf{k}_1, \mathbf{k}_2) P(k_1) P(k_2) + 2 \mathbf{k}\text{-perm.} \right\} \right] \\
 &\quad + 2 P_{\varepsilon\varepsilon\delta}^{t'} \left[Z_1^t(\mathbf{k}_1) P(k_1) + 2 \mathbf{k}\text{-perm.} \right] + B_{\varepsilon}^{t'}, \tag{3.15}
 \end{aligned}$$

where Z_2^I and \mathcal{D}_B^{IJK} are given in appendix A. Note that we omitted the redshift dependence in all terms, and also the k -dependence in \mathcal{D}_B^{IJK} . The stochastic terms have fiducial values:

$$P_{\varepsilon}^I = \frac{1}{\bar{n}_I}, \quad P_{\varepsilon\varepsilon\delta}^I = \frac{b_1^I}{2\bar{n}_I}, \quad B_{\varepsilon}^I = \frac{1}{\bar{n}_I^2}. \tag{3.16}$$

We assume that the cross shot noise of the power spectra and bispectra is zero:

$$P_{\varepsilon}^{IJ} = P_{\varepsilon\varepsilon\delta}^{IJ} = 0 \quad (I \neq J), \quad B_{\varepsilon}^{IJK} = 0 \quad (I, J, K \text{ unequal}). \tag{3.17}$$

The extension of the power spectrum and bispectrum multi-tracers from 2 to 3 tracers is given in appendix B.

Note that the multi-tracer tree-level expressions listed above are for a general form of the redshift-space kernels Z_1^I and Z_2^I . This means that additional effects that modify these kernels can be introduced fairly easily in the formalism. For example, we can incorporate local relativistic effects by replacing the kernels given in appendix A with those given in [14].

4 Multi-tracer covariance

The multi-tracer combination of power spectra and bispectra for two tracers has data vector

$$\mathbf{D}_{PB} = \left[P^{tt}, P^{tt'}, P^{t't'}, B^{ttt}, B^{(ttt')}, B^{(tt't')}, B^{t't't'} \right]. \tag{4.1}$$

The covariance of the data vector contains four blocks, coming from the covariance of the power spectra, the bispectra and the power spectra — bispectra covariance. In matrix form:

$$\mathbf{Cov}_{PB} = \begin{bmatrix} \text{Cov}(\mathbf{P}_i, \mathbf{P}_j) & \text{Cov}(\mathbf{P}_i, \mathbf{B}_b) \\ \text{Cov}(\mathbf{P}_i, \mathbf{B}_b)^\top & \text{Cov}(\mathbf{B}_a, \mathbf{B}_b) \end{bmatrix}, \quad (4.2)$$

where

$$\mathbf{P}_i = \mathbf{P}(\mathbf{k}_i), \quad \mathbf{B}_a = \mathbf{B}(\mathbf{k}_{a_1}, \mathbf{k}_{a_2}, \mathbf{k}_{a_3}). \quad (4.3)$$

We assume that the covariance between two power spectra is Gaussian, i.e. the off-diagonal contributions are neglected. Then the cross-sample covariance of the power spectra is a symmetric 3×3 diagonal block matrix:

$$\begin{aligned} \text{Cov}(\mathbf{P}_i, \mathbf{P}_j) &= \begin{pmatrix} \text{Cov}(P_i^{tt}, P_j^{tt}) & \text{Cov}(P_i^{tt}, P_j^{tt'}) & \text{Cov}(P_i^{tt}, P_j^{t't'}) \\ & \text{Cov}(P_i^{t't'}, P_j^{t't'}) & \text{Cov}(P_i^{t't'}, P_j^{t't't'}) \\ & & \text{Cov}(P_i^{t't't'}, P_j^{t't't'}) \end{pmatrix} \\ &= 2 \frac{(2\pi)^3}{V_s V_k} \delta_{ij} \begin{pmatrix} [P^{tt}(\mathbf{k}_i)]^2 & P^{tt}(\mathbf{k}_i) P^{t't'}(\mathbf{k}_i) & [P^{t't'}(\mathbf{k}_i)]^2 \\ & \frac{1}{2} \{P^{tt}(\mathbf{k}_i) P^{t't't'}(\mathbf{k}_i) + [P^{t't't'}(\mathbf{k}_i)]^2\} & P^{t't'}(\mathbf{k}_i) P^{t't't'}(\mathbf{k}_i) \\ & & [P^{t't't'}(\mathbf{k}_i)]^2 \end{pmatrix}, \end{aligned} \quad (4.4)$$

where each sub-block is a diagonal matrix in \mathbf{k}_i .

The covariance of the symmetrised bispectra is a symmetric 4×4 block matrix:

$$\begin{aligned} \text{Cov}(\mathbf{B}_a, \mathbf{B}_b) &= \frac{(2\pi)^6}{3 V_s V_{123}} s_{123} \delta_{ab} \\ &\times \begin{pmatrix} \text{Cov}(B_a^{ttt}, B_b^{ttt}) & \text{Cov}(B_a^{ttt}, B_b^{(tt't)}) & \text{Cov}(B_a^{ttt}, B_b^{(tt't't)}) & \text{Cov}(B_a^{ttt}, B_b^{t't't'}) \\ & \text{Cov}(B_a^{(tt't)}, B_b^{(tt't)}) & \text{Cov}(B_a^{(tt't)}, B_b^{(tt't't)}) & \text{Cov}(B_a^{(tt't)}, B_b^{t't't'}) \\ & & \text{Cov}(B_a^{(tt't't)}, B_b^{(tt't't)}) & \text{Cov}(B_a^{(tt't't)}, B_b^{t't't'}) \\ & & & \text{Cov}(B_a^{t't't'}, B_b^{t't't'}) \end{pmatrix}, \end{aligned} \quad (4.5)$$

where $s_{123} = 6, 2, 1$ for equilateral, isosceles and scalene triangles respectively. Each block is an $N_T \times N_T$ matrix, where N_T is the total number of triangles formed. The index a represents the triangle triplet $(\mathbf{k}_{a_1}, \mathbf{k}_{a_2}, \mathbf{k}_{a_3})$. The Kronecker delta indicates that each sub-block of the bispectrum-bispectrum covariance is a diagonal matrix, i.e. we only consider the Gaussian contribution.

These sub-block matrices are:

$$\begin{aligned}
 \text{Cov}\left(B_a^{ttt}, B_b^{ttt}\right) &= 3 P_{a_1}^{tt} P_{a_2}^{tt} P_{a_3}^{tt}, \\
 \text{Cov}\left(B_a^{ttt}, B_b^{(ttt')}\right) &= P_{a_1}^{tt} P_{a_2}^{tt} P_{a_3}^{t't'} + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{ttt}, B_b^{(t't't')}\right) &= P_{a_1}^{tt} P_{a_2}^{t't'} P_{a_3}^{t't'} + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{(ttt')}, B_b^{(ttt')}\right) &= \frac{1}{3} \left(P_{a_1}^{t't'} P_{a_2}^{t't'} P_{a_3}^{t't'} + 2 P_{a_1}^{tt} P_{a_2}^{t't'} P_{a_3}^{t't'} \right) + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{(ttt')}, B_b^{(t't't')}\right) &= \frac{1}{3} \left(P_{a_1}^{t't'} P_{a_2}^{t't'} P_{a_3}^{t't'} + P_{a_1}^{tt} P_{a_2}^{t't'} P_{a_3}^{t't'} + P_{a_1}^{t't'} P_{a_2}^{tt} P_{a_3}^{t't'} \right) + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{(t't't')}, B_b^{(t't't')}\right) &= P_{a_1}^{t't'} P_{a_2}^{t't'} P_{a_3}^{t't'} + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{(t't't')}, B_b^{(ttt')}\right) &= \frac{1}{3} \left(P_{a_1}^{t't'} P_{a_2}^{t't'} P_{a_3}^{t't'} + 2 P_{a_1}^{t't'} P_{a_2}^{tt} P_{a_3}^{t't'} \right) + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{(ttt')}, B_b^{(t't't')}\right) &= P_{a_1}^{t't'} P_{a_2}^{t't'} P_{a_3}^{t't'} + 2 \text{perm.}, \\
 \text{Cov}\left(B_a^{(t't't')}, B_b^{(t't't')}\right) &= 3 P_{a_1}^{t't'} P_{a_2}^{t't'} P_{a_3}^{t't'}. \tag{4.6}
 \end{aligned}$$

Finally, the cross-covariance between the power spectra and bispectra is itself a 3×4 block matrix:

$$\begin{aligned}
 \text{Cov}(\mathbf{P}_i, \mathbf{B}_b) &= \frac{(2\pi)^3}{V_s V_k} \tag{4.7} \\
 &\times \begin{pmatrix} \text{Cov}(P_i^{tt}, B_b^{ttt}) & \text{Cov}(P_i^{tt}, B_b^{(ttt')}) & \text{Cov}(P_i^{tt}, B_b^{(t't't')}) & \text{Cov}(P_i^{tt}, B_b^{t't't'}) \\ \text{Cov}(P_i^{t't'}, B_b^{ttt}) & \text{Cov}(P_i^{t't'}, B_b^{(ttt')}) & \text{Cov}(P_i^{t't'}, B_b^{(t't't')}) & \text{Cov}(P_i^{t't'}, B_b^{t't't'}) \\ \text{Cov}(P_i^{t't't'}, B_b^{ttt}) & \text{Cov}(P_i^{t't't'}, B_b^{(ttt')}) & \text{Cov}(P_i^{t't't'}, B_b^{(t't't')}) & \text{Cov}(P_i^{t't't'}, B_b^{t't't'}) \end{pmatrix},
 \end{aligned}$$

where

$$\begin{aligned}
 \text{Cov}(P_i^{tt}, B_b^{ttt}) &= 2 (\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) P^{tt}(\mathbf{k}_i) B^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}), \\
 \text{Cov}(P_i^{tt}, B_b^{(ttt')}) &= \frac{2}{3} \left[(\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) P^{t't'}(\mathbf{k}_i) B^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right. \\
 &\quad \left. + P^{tt}(\mathbf{k}_i) B_{2\delta}^{t't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right], \\
 \text{Cov}(P_i^{tt}, B_b^{(t't't')}) &= \frac{2}{3} \left[P^{tt}(\mathbf{k}_i) B_{1\delta}^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + P^{t't'}(\mathbf{k}_i) B_{2\delta}^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right], \\
 \text{Cov}(P_i^{tt}, B_b^{t't't'}) &= 2 P^{t't'}(\mathbf{k}_i) B_{1\delta}^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}), \\
 \text{Cov}(P_i^{t't'}, B_b^{ttt}) &= (\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) P^{t't'}(\mathbf{k}_i) B^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + P^{tt}(\mathbf{k}_i) B_{1\delta}^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}),
 \end{aligned}$$

$$\begin{aligned}
 \text{Cov}\left(P_i^{tt'}, B_b^{(ttt')}\right) &= \frac{1}{3} \left\{ (\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) \left[P^{tt'}(\mathbf{k}_i) B^{(ttt')}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right. \right. \\
 &\quad \left. \left. + P^{t't'}(\mathbf{k}_i) B^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right] + P^{tt'}(\mathbf{k}_i) B_{2\delta}^{tt't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right\}, \\
 \text{Cov}\left(P_i^{tt'}, B_b^{(tt't')}\right) &= \frac{1}{3} \left\{ (\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) \left[P^{tt'}(\mathbf{k}_i) B^{(tt't')}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right. \right. \\
 &\quad \left. \left. + P^{t't'}(\mathbf{k}_i) B^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right] + P^{t't'}(\mathbf{k}_i) B_{2\delta}^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right\}, \\
 \text{Cov}\left(P_i^{tt'}, B_b^{t't't'}\right) &= (\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) P^{tt'}(\mathbf{k}_i) B^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \\
 &\quad + P^{t't'}(\mathbf{k}_i) B_{1\delta}^{tt't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}), \\
 \text{Cov}\left(P_i^{t't'}, B_b^{ttt}\right) &= 2P^{t't'}(\mathbf{k}_i) B_{1\delta}^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}), \\
 \text{Cov}\left(P_i^{t't'}, B_b^{(ttt')}\right) &= \frac{2}{3} \left[P^{t't'}(\mathbf{k}_i) B_{1\delta}^{ttt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + P^{t't'}(\mathbf{k}_i) B_{2\delta}^{tt't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right], \\
 \text{Cov}\left(P_i^{t't'}, B_b^{(tt't')}\right) &= \frac{2}{3} \left[(\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) P^{t't'}(\mathbf{k}_i) B^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right. \\
 &\quad \left. + P^{t't'}(\mathbf{k}_i) B_{2\delta}^{tt't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \right], \\
 \text{Cov}\left(P_i^{t't'}, B_b^{t't't'}\right) &= 2(\delta_{ib_1} + \delta_{ib_2} + \delta_{ib_3}) P^{t't'}(\mathbf{k}_i) B^{t't't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}). \tag{4.8}
 \end{aligned}$$

Here the partially symmetric bispectra are given by:

$$\begin{aligned}
 B_{1\delta}^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) &= \delta_{ib_1} B^{t'tt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + \delta_{ib_2} B^{tt't}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \\
 &\quad + \delta_{ib_3} B^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}), \tag{4.9}
 \end{aligned}$$

$$\begin{aligned}
 B_{2\delta}^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) &= [\delta_{ib_1} + \delta_{ib_2}] B^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + [\delta_{ib_1} + \delta_{ib_3}] B^{tt't}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) \\
 &\quad + [\delta_{ib_2} + \delta_{ib_3}] B^{t'tt}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}). \tag{4.10}
 \end{aligned}$$

These bispectrum tracer indices are used for convenience and can be generalised for any 2-tracer combination, e.g. $tt't'$. Note that a nonzero power spectrum-bispectrum cross-covariance is generated by configurations in which the power spectrum \mathbf{k} mode coincides at least with one of the sides of a triangle configuration. The results presented above for the cross-covariance between power spectrum and bispectrum, consider only the dominant contribution, while the 5-point connected correlation function part is neglected (see [72] for a discussion). Note that the block matrix of eq. (4.8) and each of the sub-blocks are not square matrices.

The expressions for 3 tracers are in appendix B.

5 Fisher information on f_{NL}

The previous section presented the formalism for the multi-tracer power spectra, the multi-tracer bispectra and their combination. In this section, we use the formalism for a simplified qualitative analysis of the improvements from the multi-tracer constraints on f_{NL} relative to the single-tracer as baseline. In order to achieve this, we consider the same volume and redshift range for each single-tracer and their multi-tracer. This means that the improvement

shown here on local PNG only refers to the effect of combining the samples or not, without taking into account the effect of different volumes and redshift overlap of two different samples. To this end, we consider only the conditional errors, and we use nominal surveys without considering details and neglecting systematics. More realistic forecasts, that marginalise over relevant cosmological and nuisance parameters and use specifications for current and upcoming galaxy surveys, will be treated in a follow-up paper [73]. Here we also confine ourselves to separate investigations into the precision from the multi-tracer power spectra and from the multi-tracer bispectra — without combining the information obtained from the two correlators. The follow-up paper [73] will include this combination of multi-tracers as part of more comprehensive forecasts.

The Fisher information on f_{NL} in each redshift bin for the multi-tracer power spectra is given by

$$F_{f_{\text{NL}}}^{\mathbf{P}} = \int d\mathbf{k} \frac{\partial \mathbf{D}_{\mathbf{P}}(\mathbf{k})}{\partial f_{\text{NL}}} \text{Cov}_{\mathbf{P}}^{-1}(\mathbf{k}) \frac{\partial \mathbf{D}_{\mathbf{P}}^{\top}(\mathbf{k})}{\partial f_{\text{NL}}}. \quad (5.1)$$

For a rough estimate of the improvements on precision that are produced by the multi-tracer, we consider the error $\sigma^{\mathbf{P}}(f_{\text{NL}}) = (F_{f_{\text{NL}}}^{\mathbf{P}})^{-1/2}$. Then an analytic expression can be derived for the single-tracer power spectrum:

$$\sigma_{f_{\text{NL}}}^I = N_P^{1/2} \int d\mathbf{k} \frac{M(k) Z_1^I(\mathbf{k})}{2 b_{\Psi}^I k^{\alpha}} \frac{P^I(\mathbf{k})}{P^I(\mathbf{k}) - P_{\varepsilon}^I}, \quad N_P = 2 \frac{(2\pi)^3}{V_s V_k}, \quad (5.2)$$

where we omitted the redshift dependence. Here $\alpha = 1, 0$ for orthogonal and local shapes, respectively [44, 55, 56, 61, 74–76]. Note that equilateral PNG ($\alpha = 2$) cannot be constrained using the galaxy power spectrum, as shown in [65]. Using the linear redshift space kernel eq. (A.1), the analytic result eq. (5.2) shows that, for a fixed number density and μ , the tightest constraints on f_{NL} can be obtained from a sample that has $b_{\Psi}^I \gg b_1^I$.

For the multi-tracer case the analytic expression of the Fisher matrix on $f_{\text{NL}}^{\text{loc}}$ is given by:

$$\begin{aligned} F_{f_{\text{NL}}}^{\mathbf{P}} = \sum_{I,J} \int d\mathbf{k} \mathcal{N}^{-1}(\mathbf{k}) & \left\{ 4 P(k)^2 (b_{\Psi}^J)^2 [P_{\varepsilon}^I + P_{\varepsilon}^J R(\mathbf{k})^2]^2 [Z_1^J(\mathbf{k})]^2 \right. \\ & + 8 P(k)^2 P_{\varepsilon}^J b_{\Psi}^J [Z_{1,G}^J(\mathbf{k}) b_{\Psi}^I - Z_{1,G}^I(\mathbf{k}) b_{\Psi}^J] [P_{\varepsilon}^I + P_{\varepsilon}^J R(\mathbf{k})^2] R(\mathbf{k}) Z_1^J(\mathbf{k}) \\ & \left. + 2 P(k)^2 [Z_{1,G}^J(\mathbf{k}) b_{\Psi}^I - Z_{1,G}^I(\mathbf{k}) b_{\Psi}^J]^2 [P_{\varepsilon}^I P^J(\mathbf{k}) + P_{\varepsilon}^J \{P_{\varepsilon}^J + P^J(\mathbf{k})\} R^2(\mathbf{k})] \right\}, \end{aligned} \quad (5.3)$$

where

$$\mathcal{N}(\mathbf{k}) = N_P M(k)^2 k^{-2\alpha} [P_{\varepsilon}^I P^J(\mathbf{k}) + P_{\varepsilon}^J \{P^J(\mathbf{k}) - P_{\varepsilon}^J\} R(\mathbf{k})^2]^2 \quad \text{with} \quad R(\mathbf{k}) = \frac{Z_1^I(\mathbf{k})}{Z_1^J(\mathbf{k})}. \quad (5.4)$$

Here $Z_{1,G}^I(\mathbf{k})$ is the Gaussian version of eq. (A.1), i.e. for $f_{\text{NL}} = 0$. Note that this expression is valid for a general linear redshift kernel (e.g. including non-integrated relativistic corrections).

Our eq. (5.3) generalises the result of [40], which only considers local PNG, takes the cosmic variance limit and neglects RSD effects. At fixed redshift bin and scales, eq. (5.3) indicates that precision on f_{NL} is increased by

$$\text{increasing } |Z_{1,G}^J(\mathbf{k}) b_{\Psi}^I - Z_{1,G}^I(\mathbf{k}) b_{\Psi}^J| \quad \text{and} \quad R(\mathbf{k}). \quad (5.5)$$

One way to maximise the former is by choosing appropriately the line-of-sight angle μ to gain a boost from the presence of RSD, which is also evident in forecasts that utilise single tracers (e.g. [77]). If we neglect RSD (or fix μ) and consider local PNG, then eq. (5.5) implies maximising $|b_1^J b_{\Phi}^I - b_1^I b_{\Phi}^J|$, as shown also in [40].

For the multi-tracer bispectra, the Fisher matrix is

$$F_{f_{\text{NL}}}^{\mathbf{B}} = \int d^3\mathbf{k} \frac{\partial \mathbf{D}_B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{\partial f_{\text{NL}}} \mathbf{Cov}_B^{-1}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \frac{\partial \mathbf{D}_B^{\top}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{\partial f_{\text{NL}}}. \quad (5.6)$$

Analytical calculations for the bispectrum multi-tracer are long and uninformative, without assuming some limit, either for the shot noise or the triangle configurations. Consequently, we perform the bispectrum analysis numerically.

We consider two nominal spectroscopic surveys, with the following features:

- Overlap sky area = 10 000 deg², overlap redshift range $0.9 \leq z \leq 1.7$ (bin size $\Delta z = 0.1$).
- *Tracer t*:
 - high density, $\mathcal{O}(10^{-3})$ (h/Mpc)³; $\bar{n}_t(z)$ similar to a Euclid sample [78];
 - linear bias $b_1^t(z) = 1.3 + 0.6 z$.
- *Tracer t'*:
 - lower density, $\mathcal{O}(10^{-4})$ (h/Mpc)³; $\bar{n}_{t'}(z)$ similar to DESI sample [79];
 - linear bias $b_1^{t'}(z) = 0.84/D(z)$.

In order to derive the higher-order and non-Gaussian bias coefficients needed for the tree-level bispectrum, we use a simplified halo-occupation distribution (HOD) model from [80]. The HOD model has two free parameters which are calibrated from the values of the linear bias and number densities at each redshift.

This set-up for the nominal surveys exploits the findings of eq. (5.3) and will function as the benchmark case. The fiducial cosmology is a flat Λ CDM model, with $\omega_c = 0.1199$, $\omega_b = 0.02237$, $h = 0.6736$, $n_s = 0.96488$, $A_s = 2.1 \times 10^{-9}$, as measured by [81]. In addition, we assume non-Gaussian initial conditions, with an amplitude of $f_{\text{NL}} = 20$, for all three PNG types considered here. The scales used to perform the Fisher matrix analysis are well within the perturbative regime: $k_{\text{min}}(z) = k_f(z) \equiv 2\pi/V_s(z)^{1/3}$ and $k_{\text{max}}(z) = 0.75 k_{\text{NL}}(z)$,

$\sigma(f_{\text{NL}})(z)$: benchmark case. Figure 1 presents the relative change in the conditional f_{NL} errors from the multi-tracer approach with respect to the two nominal single-tracers t and t' , as a function of redshift. The results are for three types of PNG, i.e. local, equilateral and orthogonal, using the power spectrum and bispectrum. For both correlators, the multi-tracer delivers a noticeable overall improvement relative to each single-tracer survey.

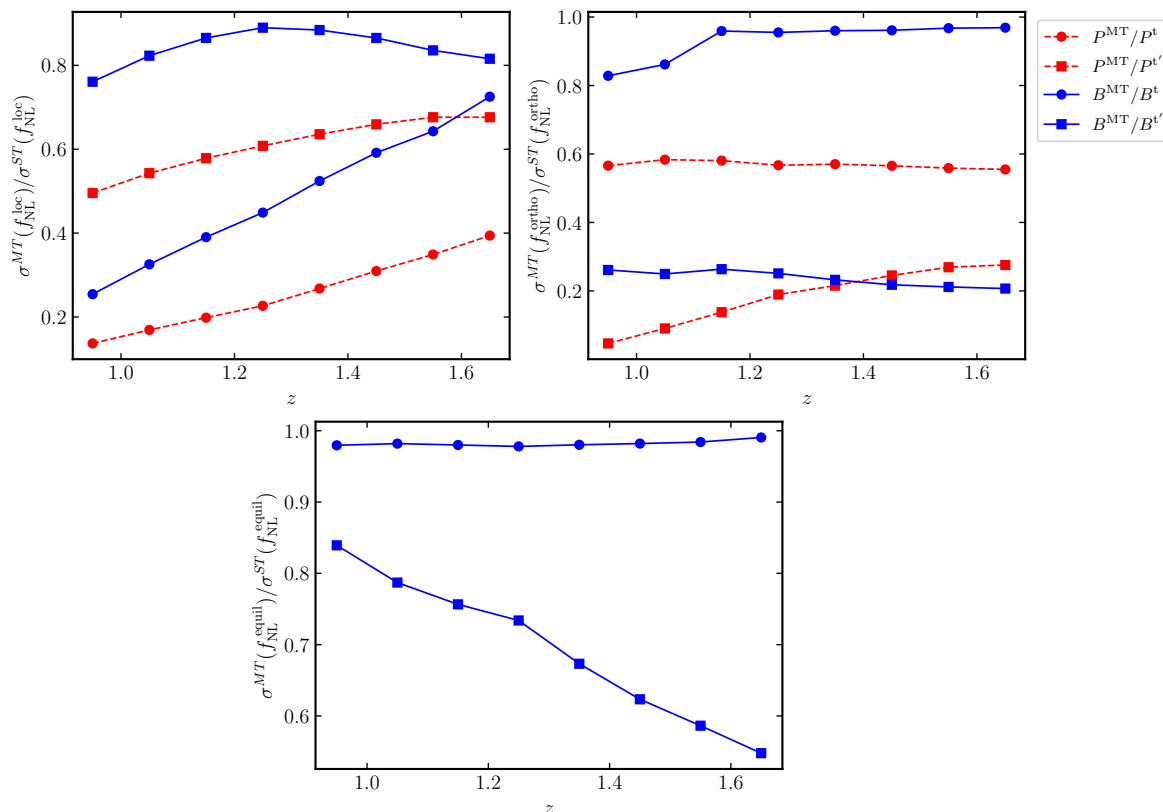


Figure 1. Relative change in the cumulative 1σ error on f_{NL} , provided by the multi-tracer (MT) approach relative to each single tracer (ST), as a function of redshift. Power spectrum (red lines) and bispectrum (blue lines) results, are shown for three PNG types. t and t' are tracers of the two nominal next-generation surveys considered here.

In the case of local PNG, the multi-tracer power spectrum improvement over the f_{NL} errors from a single tracer is significant — well above 30% across the redshift range. For the bispectrum, this applies only in the low-redshift bins of tracer t , while for tracer t' the gain is more modest. However, the overall enhancement of the bispectrum multi-tracer approach is still important and in the range 20–30%.

For orthogonal PNG, the improvement from the multi-tracer with respect to tracer t' is equivalent for both correlators ($\sim 70\%$), while regarding tracer t it is significantly more for P^{MT} than for B^{MT} . Note that the performances of the single tracers t and t' are opposite to the local case. Equation (5.2) shows that the single-tracer conditional error from the power spectrum on f_{NL} is parametrically determined by b_1 , b_Ψ and P_ε . Numerical tests and the analytic results show that the constraints are more sensitive to the values of b_Ψ , in particular for the case where $b_\Psi^t \gg b_1^t$. For the local case, $b_\Psi^{t'} > b_\Psi^t$ and hence we see that tracer t outperforms the constraints from t' , for both correlators. On the other hand, for orthogonal PNG, due to the nature of this shape (see [44] for a discussion), $b_\Psi^t > b_\Psi^{t'}$, leading to an opposite behaviour with respect to the local PNG.

For equilateral PNG, the multi-tracer bispectrum gives only a marginal improvement over the performance of tracer t , as in the orthogonal case. On the other hand, the gain for

tracer t' increases with redshift, reaching a substantial $\sim 50\%$. Note that the power spectrum does not provide any constraints on $f_{\text{NL}}^{\text{equil}}$ (see e.g. [65]).

The attributes of the nominal surveys were selected to produce a large $|b_1^J b_\Psi^I - b_1^I b_\Psi^J|$. The results in figure 1 indicate that in the general benchmark case — namely, with no additional assumptions — this is enough to achieve a significant gain from the multi-tracer approach in most cases, but for the rest it is insufficient.

Apart from the dependence on PNG type shown in figure 1, additional parameters have a significant effect on the level of improvement achieved by the multi-tracer. The main tracer-dependent variables that affect $\sigma(f_{\text{NL}})$, besides the linear and PNG biases (b_1^I and b_Ψ^I), are the number densities \bar{n}_I , which in turn affect the stochastic contributions in eq. (2.10). In addition, for the bispectrum, higher-order bias parameters such as b_2^I will have an effect on the constraining power.

Further examination is needed into the parametric behaviour of the multi-tracer power spectrum and bispectrum, in order to determine the tracer properties that maximise gain from combining their information. Here we confine attention to local PNG, and we fix the redshift at $z = 1$. Note that $b_\Psi \rightarrow b_\Phi$ for local PNG, as follows from eq. (2.7) with $\alpha = 0$.

$\sigma(f_{\text{NL}}^{\text{loc}})$: effect of number densities. We start with the effect of number densities on $f_{\text{NL}}^{\text{loc}}$ conditional errors. We vary \bar{n}_t , while fixing all other parameters of tracers t and t' , including $\bar{n}_{t'}$. The dependence of the $f_{\text{NL}}^{\text{loc}}$ errors on the ratio $\bar{n}_t/\bar{n}_{t'}$, at $z = 1$, is shown in the left panel of figure 2. In addition to the multi-tracer results (marked with the superscript MT), we present the forecasts from simply adding the Fisher matrices of the two tracers (marked with the superscript $t + t'$), without considering the cross-terms in the data vectors, eqs. (3.9) and (3.11), and block covariances, eqs. (4.4) and (4.5).

The left panel of figure 2 shows that for both correlators the multi-tracer gives maximal improvement on $f_{\text{NL}}^{\text{loc}}$ errors for $\bar{n}_t/\bar{n}_{t'} \gtrsim 10$. In other words, the number density of one tracer should be at least an order of magnitude larger than that of the other. In fact, if the ratio is beyond that point, the constraints do not improve further, as we remain shot-noise dominated by the low-density tracer. This behaviour occurs since only one of the two tracers is allowed to reach the limit where stochastic contributions can be neglected (i.e. the cosmic variance limit). The other tracer has a constant shot-noise component, that ultimately saturates the multi-tracer Fisher information. Nonetheless, in this setup, the multi-tracer power spectrum seems to be more sensitive to the cosmic variance limit of tracer t , than the bispectrum. This is an expected behaviour, since the effect of local PNG appears only on the largest scales for the tracer power spectrum, whereas the effect on the tracer bispectrum is from the squeezed configurations that are not purely dependent on the largest scales, since they couple large scales to small scales.

On the other hand, the constraints from just summing the Fisher matrices of the two tracers (i.e. neglecting the multi-tracer cross-terms), exhibit a minimal change as a function of the number density ratio. This indicates that the complete multi-tracer approach takes advantage of the increasing number density in an optimal way, by maximising the contribution of the cross-terms and hence providing a significant improvement on the $f_{\text{NL}}^{\text{loc}}$ constraints, i.e. up to 10% for $\bar{n}_t/\bar{n}_{t'} \gtrsim 10$. Moreover, for the low density regime ($\bar{n}_t/\bar{n}_{t'} < 1$), neglecting the multi-tracer cross-terms, when considering the summed information from two overlapping

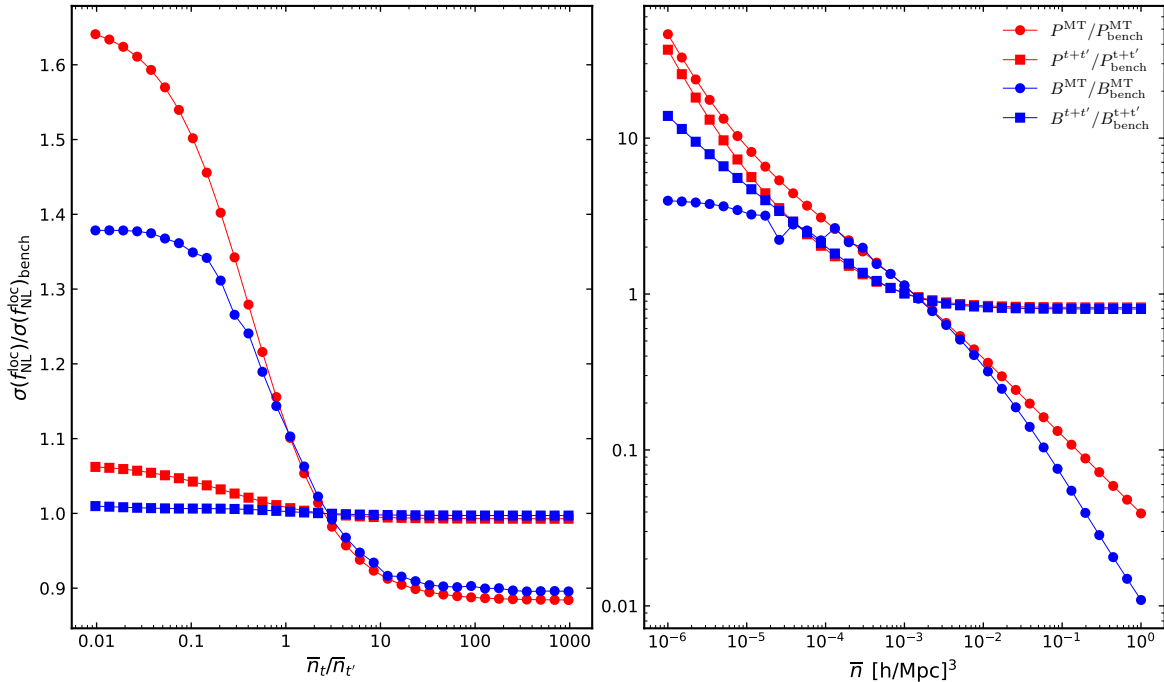


Figure 2. *Left:* the dependence of the 1σ conditional $f_{\text{NL}}^{\text{loc}}$ errors on the ratio $\bar{n}_t/\bar{n}_{t'}$, at fixed redshift $z = 1$. We vary only the number density of sample t , while fixing all other parameters of tracer t , together with those of sample t' . *Right:* testing the cosmic variance limit, by enforcing equal number densities, i.e. $\bar{n}_t = \bar{n}_{t'} = \bar{n}$, and varying \bar{n} (at $z = 1$). Both panels show the multi-tracer forecasts (circles) and the results from the simple combination of the two Fisher matrices (squares) (without considering the cross-terms in the data vectors and block covariances), divided by the benchmark results of figure 1 at $z = 1$, for both summary statistics.

tracers, can lead to severe overestimation of the $f_{\text{NL}}^{\text{loc}}$ errors (even up to 40–60%), in the case of both correlators.

Note that the scenario with equal number densities of tracers (i.e. $\bar{n}_t/\bar{n}_{t'} = 1$) is not an optimal set-up for a multi-tracer application, at least when neither of the tracers approaches the high-density limit.

In order to investigate further the limit where the shot noise can be neglected, we assume equal number densities. Our aim is to explore the low- and high-density limits. The dependence of the forecast error on $f_{\text{NL}}^{\text{loc}}$ as a function of number density is presented in the right panel of figure 2, for both correlators. This indicates that the multi-tracer constraints on local PNG exhibit an inverse power-law behaviour as the number densities of the tracers approach the cosmic variance limit, $\bar{n} \gg 10^{-3} (h/\text{Mpc})^3$. By contrast, the results of simply adding the Fisher matrices of the two tracers saturate around the same value, without exhibiting any further improvement as the number densities increase. This shows the clear advantage of the multi-tracer approach, where the cross-term contribution is enhanced substantially, due to the decreasing shot-noise terms, providing significant improvement on $f_{\text{NL}}^{\text{loc}}$.

Our power spectrum results clearly show the benefit of the multi-tracer, for the high density samples, as well as for tracers that approach the cosmic variance limit. This agrees with the findings of the original work on the multi-tracer approach [25], as well as the more

recent results of [40]. Moreover, here we show that the advantage of the multi-tracer towards the high-density limit can also be extended to the bispectrum, which exhibits a significant improvement compared to the power spectrum in the cosmic variance limited regime.

$\sigma(f_{\text{NL}}^{\text{loc}})$: effect of linear biases. Next we investigate the effect of the linear biases on the multi-tracer power spectrum and bispectrum. Due to the symmetrisation of the two correlators under tracer permutations, the outcome of the parametric analysis does not depend on which tracer's parameters we chose to vary. Note that the higher-order bias coefficients of each tracer depend on the HOD model, whose free parameters are calibrated by the value of the linear bias at each redshift. This means that varying b_1^t will, in the HOD framework, affect all the bias coefficients of tracer t . Here we are interested in isolating the effect of a change in linear bias and we vary only b_1^t for tracer t , while fixing all other parameters for tracers t and t' .

The relative change of the multi-tracer constraints on $f_{\text{NL}}^{\text{loc}}$ over the benchmark case, for the two correlators and at $z = 1$, are presented in figure 3 as a function of the difference between the linear biases, $b_1^t - b_1^{t'}$. The results indicate that when the difference between the two biases is $\gtrsim 2$, there is a significant improvement (up to $\sim 30\%$) with respect to the benchmark results, $\sigma(f_{\text{NL}}^{\text{loc}})/\sigma(f_{\text{NL}}^{\text{loc}})_{\text{bench}} = 1$, for the multi-tracer power spectrum. For the multi-tracer bispectrum, the gain is more modest. Beyond that point, the improvement slowly saturates ($|b_1^t - b_1^{t'}| > 5$) for both correlators. The results show that the constraining power on $f_{\text{NL}}^{\text{loc}}$ is $\propto b_1^t - b_1^{t'}$. Note that the case where the linear biases of the tracers are similar, $|b_1^t - b_1^{t'}| \sim 0$, corresponds to the least tight constraints on $f_{\text{NL}}^{\text{loc}}$, for both correlators. In this case, the power spectrum constraints degrade significantly ($\sim 60\%$) compared to the bispectrum ($\sim 10\%$).

$\sigma(f_{\text{NL}}^{\text{loc}})$: effect of $b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'}$. We proceed by allowing all other bias coefficients of tracer t to vary as well, through the dependence of the HOD free parameters on the b_1^t value, including $b_{\Phi}^t(b_1^t)$ via the universality relation assumption. The relative change of the $f_{\text{NL}}^{\text{loc}}$ forecasts, compared to the benchmark results, as a function of $b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'}$, is displayed in figure 4. The behaviour exhibited indicates that for both correlators, the gain on the $f_{\text{NL}}^{\text{loc}}$ forecasts increases as a function of $|b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'}|$, reaching significant levels.

For the multi-tracer power spectrum, the improvement on $f_{\text{NL}}^{\text{loc}}$ constraints — i.e., the case where $\sigma(f_{\text{NL}}^{\text{loc}})/\sigma(f_{\text{NL}}^{\text{loc}})_{\text{bench}} < 1$ — saturates at $\sim 20\%$ when dealing with large positive values of $b_1^{t'}$, $b_{\Phi}^t - b_1^t$, $b_{\Phi}^{t'}$. The maximum gain seems to be achieved towards the negative values of $b_1^{t'}$, $b_{\Phi}^t - b_1^t$, $b_{\Phi}^{t'}$. The improvement towards the negative values is in fact significantly larger than the saturation regime towards the large positive values. Nonetheless, maximising $|b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'}|$ provides the highest gain from the multi-tracer power spectrum, verifying the analytical findings of eq. (5.3).

For the multi-tracer bispectrum, the tightest constraint on $f_{\text{NL}}^{\text{loc}}$ is provided once $b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'} > 0$. Beyond that, the improvement relative to the benchmark saturates. This indicates the different behaviour of the two correlators as a function of $b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'}$, once we allow for all bias parameters of tracer t to change and $b_{\Phi}^t \propto b_1^t$. Moreover, the multi-tracer power spectrum results for $b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'} \approx 0$ exhibit a degradation ($\sim 20\%$) relative to the benchmark case, while for the bispectrum the constraints are close to the optimal outcome.

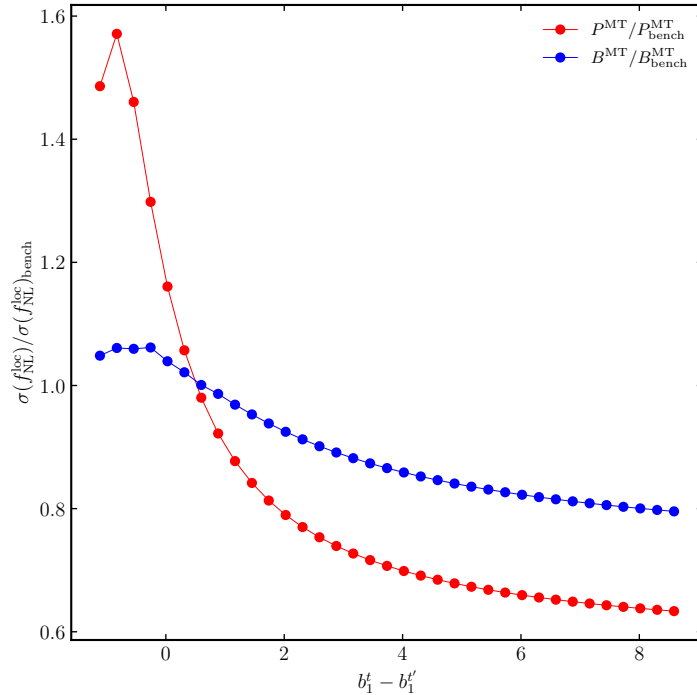


Figure 3. The dependence of the 1σ conditional $f_{\text{NL}}^{\text{loc}}$ error on the linear bias difference $b_1^t - b_1^{t'}$, at fixed redshift $z = 1$, for the multi-tracer power spectrum and bispectrum. Here we vary only b_1^t , while fixing all other tracer-dependent parameters for both t and t' . The results are normalised to the benchmark forecasts of figure 1.

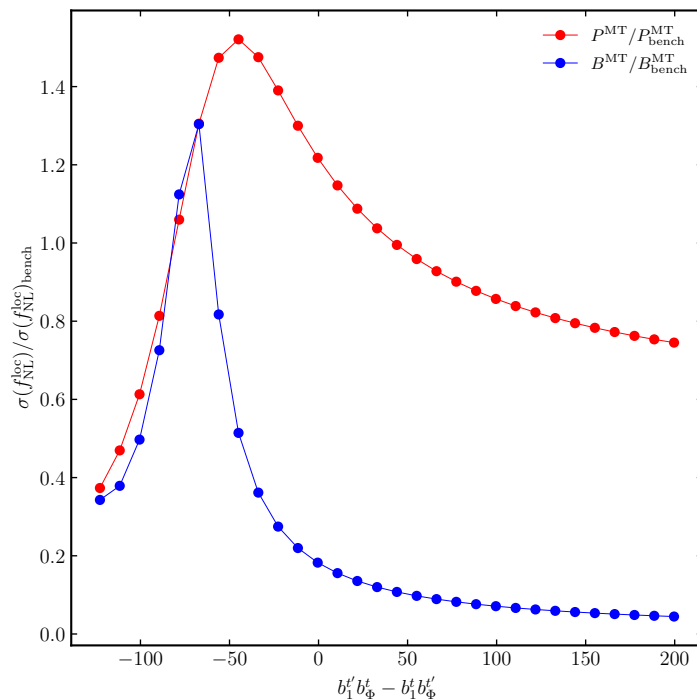


Figure 4. Same as figure 3, where we not only vary b_1^t , but we allow all the other bias parameters of tracer t to change in accordance with the dependence of the HOD model on the linear bias, including b_Φ^t . The relative change, with respect to the benchmark results, is shown as a function of $b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}$.

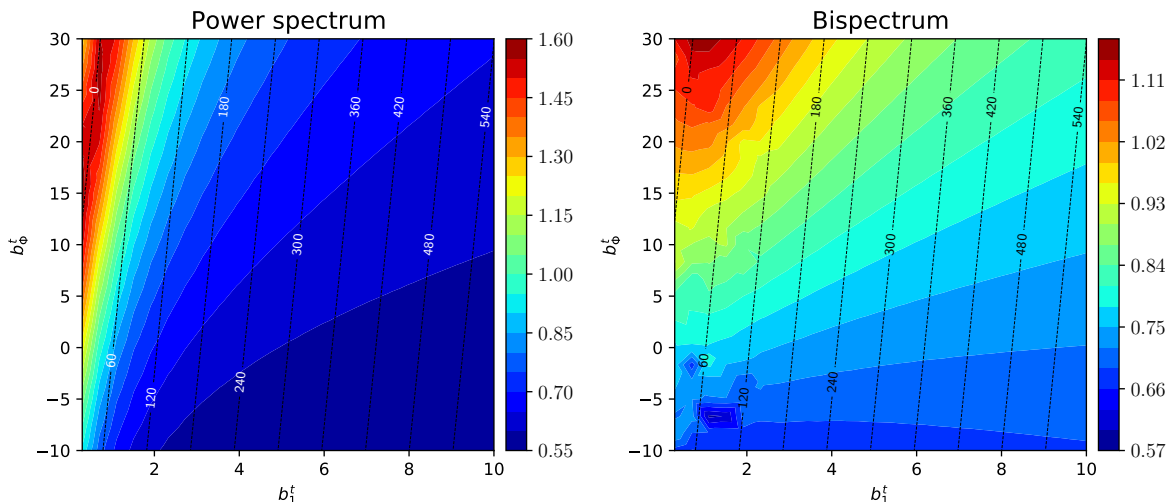


Figure 5. The ratio $\sigma(f_{\text{NL}}^{\text{loc}})/\sigma(f_{\text{NL}}^{\text{loc}})_{\text{bench}}$, with values given by the colour bar, in the (b_Φ^t, b_1^t) plane, for multi-tracer power spectrum (*left*) and bispectrum (*right*). We vary b_1^t and b_Φ^t independently, without assuming a universality relation, while fixing all other parameters for both tracers. Black dotted lines are contours of constant $b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}$.

The different parametric behaviour of the multi-tracer power spectrum and bispectrum can also be seen in the contour plots of figure 5. In this case, we vary b_1^t and b_Φ^t , without assuming any relation between the two (i.e. no universality relation), while fixing all other tracer-dependent parameters for both tracers. The black dotted contours are lines of constant $b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}$. The results show that for the multi-tracer power spectrum, the case $b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'} \approx 0$ corresponds to the weakest constraints, where errors increase compared to the benchmark case. The maximal improvement on the forecast precision originates from the case of maximal $b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}$ values — in particular for $-10 \leq b_\Phi \leq 10$ and for $b_1 \geq 2$. The behavior of the multi-tracer bispectrum is somewhat similar, but in the $b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'} \approx 0$ regime, which corresponds also to weaker constraints than the benchmark, the performance is better than the power spectrum equivalent. As shown in figure 4, the optimal constraints from the multi-tracer bispectrum arise for an increasing $|b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}|$. In particular, in order to achieve a significant improvement, growing b_1^t values and very negative b_Φ^t are required (blue shaded region in the right panel of figure 5).

The results presented in figures 3, 4 and 5 indicate the potential of the multi-tracer to significantly improve constraints on local PNG by carefully choosing the two tracers. For the multi-tracer power spectrum and bispectrum, increasing $|b_1^t - b_1^{t'}|$ or $|b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}|$ gives the optimal results, where the latter leads to overall substantial enhancement.

$\sigma(f_{\text{NL}}^{\text{loc}})$: effect of second-order biases. In order to investigate further the parametric behaviour of the multi-tracer tree-level bispectrum, we vary the second-order bias coefficient b_2^t , while all other parameters for both tracers are fixed. The relative improvement results are shown in figure 6 as a function of the difference $b_2^t - b_2^{t'}$ for the complete multi-tracer approach, and for the sum of the individual Fisher matrices. Comparing the two shows once again the importance of utilising the full multi-tracer in order to optimise the gain on $f_{\text{NL}}^{\text{loc}}$ constraints compared to the summed information of the bispectrum of two tracers.

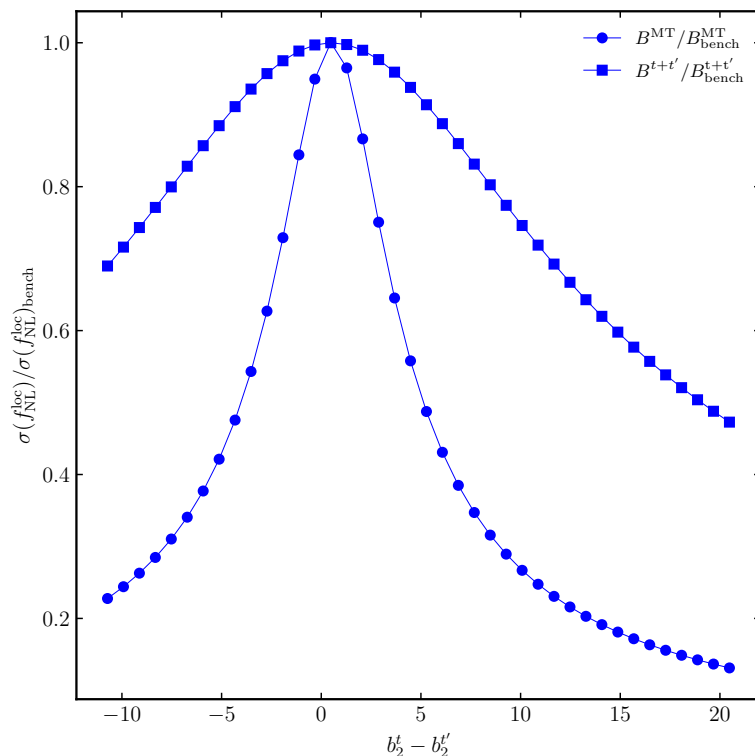


Figure 6. $\sigma(f_{\text{NL}}^{\text{loc}})/\sigma(f_{\text{NL}}^{\text{loc}})_{\text{bench}}$ as a function of the difference $b_2^t - b_2^{t'}$, varying b_2^t , while keeping fixed all other parameters for both tracers. Square points correspond to the summation of the two Fishers, while dots represent the full multi-tracer results.

In addition, varying b_2^t shows a peak around the case where the nonlinear bias parameters of tracers t and t' are similar, i.e. the difference $|b_2^t - b_2^{t'}|$ is small. This indicates the regime where the multi-tracer bispectrum offers minimal improvement with respect to the benchmark results. Therefore, the constraining power on $f_{\text{NL}}^{\text{loc}}$ is $\propto |b_2^t - b_2^{t'}|$. In fact, any value outside of the range $-1 < |b_2^t - b_2^{t'}| < 5$ can have an improvement of $> 20\%$. Moreover, increasing the value of $|b_2^t - b_2^{t'}|$ improves the gain on $f_{\text{NL}}^{\text{loc}}$ as much as, or even more than, the case of an increasing $|b_1^t - b_1^{t'}|$ or $|b_1^{t'} b_{\Phi}^t - b_1^t b_{\Phi}^{t'}|$ (see figures 3 and 4). This reveals the importance of the nonlinear bias parameter in utilising the full potential of the multi-tracer bispectrum.

$\sigma(f_{\text{NL}}^{\text{loc}})$: effect on scales. The benefits of the multi-tracer approach for local PNG constraints beyond the nominal scenario studied here, should be explored. Although this is left for a future work, we can provide in this section some preliminary indication of whether the improvement shown in figure 1 would persist in any upcoming survey. To achieve this we investigate the sensitivity of the results to the minimum and maximum accessible scales. This is done in order to explore whether or not the multi-tracer gain originates from the expected scale range — and hence if it is likely to persist in a realistic set-up.

The information on the amplitude of local PNG, in the case of the power spectrum, comes from the scale-dependent correction to the linear bias (see section 2), while for the bispectrum it mostly originates from squeezed configurations that couple long and short modes. The latter has an additional source of information, similar to the scale-dependent bias correction in the

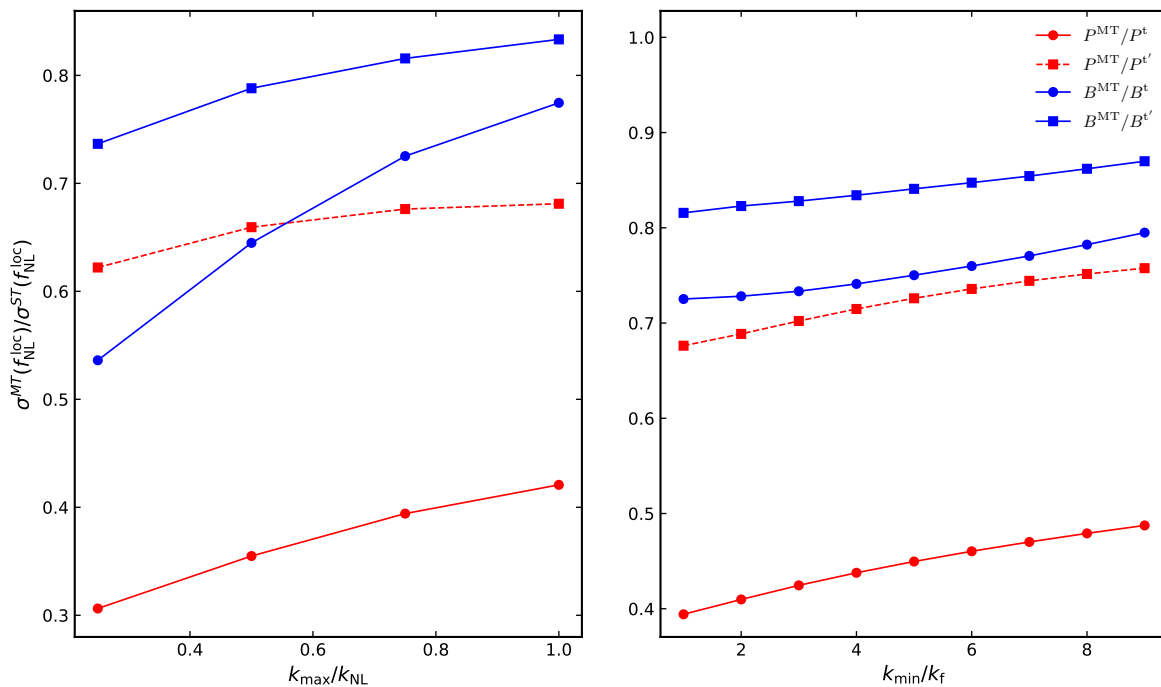


Figure 7. The relative change of the multi-tracer compared to single tracers t and t' , for the power spectrum (red) and bispectrum (blue), as a function of the maximum wavenumber k_{\max} (left panel) and minimum k_{\min} (right panel). The $f_{\text{NL}}^{\text{loc}}$ errors correspond to the cumulative contribution from all redshift bins, while k_{NL} and k_f are given in the text under figure 1. Note that the benchmark results in figure 1 correspond to $k_{\max}/k_{\text{NL}} = 0.75$ and $k_{\min}/k_f = 1$ respectively.

power spectrum (see e.g. [77]). While the power spectrum PNG constraints are very sensitive to these scales, the bispectrum can still provide competitive results in the absence of the very large scales, as long as enough squeezed triangles are probed by the survey [19, 21, 77].

The sensitivity of the multi-tracer improvement in $f_{\text{NL}}^{\text{loc}}$ constraints, as a function of the largest accessible scales $k_{\min}(z)$, is presented in the right panel of figure 7. Here the expression for $k_{\max}(z)$ remains the same as in the benchmark case. The results are presented for both correlators, where the 1σ errors originate from the cumulative information over the whole redshift range and each point assumes a fixed value of the ratio $k_{\min}(z)/k_f(z)$.

In the case of the power spectrum, the relative gain of the multi-tracer with respect to the single-tracer exhibits only a small loss (up to $\sim 10\%$ reduction), once we reduce the access to the large scales (large values of the ratio $k_{\min}(z)/k_f(z)$). Although power spectrum constrains on $f_{\text{NL}}^{\text{loc}}$ are the most sensitive to these scales, the multi-tracer approach successfully maintains the relative improvement over single-tracer results, at more or less the same level.

The multi-tracer bispectrum follows a similar trend, where a $\sim 10\%$ reduction to the benchmark ($k_{\min}(z)/k_f(z) = 1$) improvement is observed. This decrease can be important for surveys like those considered here, where a maximum 20–30% gain from the multi-tracer bispectrum over the single-tracer (see figure 1) is provided. However, in the case where the parametric behaviour (discussed in the previous subsections) is used to advantage in the tracer selection process, the multi-tracer bispectrum could still provide a significant improvement to the local PNG constraints.

To investigate further the sensitivity of the multi-tracer improvement seen in the benchmark case to the small clustering scales, we vary the ratio $k_{\max}(z)/k_{\text{NL}}(z)$, while fixing $k_{\min}(z)$ to the fiducial choice. The results are presented in the left panel of figure 7. The power spectrum exhibits a reduction, in the overall gain, that does not surpass $\sim 10\%$. The constraints on $f_{\text{NL}}^{\text{loc}}$ in this case do not depend on the information within the small scales. The observed decrease in the relative gain is simply attributed to the improvement of the $f_{\text{NL}}^{\text{loc}}$ constraints from each single-tracer, while the multi-tracer errors are more or less unaffected.

On the other hand, limiting substantially the accessibility to small scales, i.e. in the regime $k_{\max}(z)/k_{\text{NL}}(z) < 0.5$, reduces the number of squeezed triangles, affecting the performance of the bispectrum in constraining $f_{\text{NL}}^{\text{loc}}$. This is true for both the multi-tracer and single-tracer analysis. However, in this case we see that the gain of the multi-tracer bispectrum is greater, managing to maintain the relative improvement relative to each single tracer at significant levels (within the margin 20–50%). This means that reducing the small-scale access affects the single-tracer analysis substantially more than the multi-tracer. In other words, the multi-tracer bispectrum does much better in a regime which is not shot-noise dominated. The gain exhibited reaches the benchmark values as soon as the small scale access is restored to the fiducial choice (i.e. $k_{\max}(z)/k_{\text{NL}}(z) \gtrsim 0.75$).

The overall trends suggests that the $f_{\text{NL}}^{\text{loc}}$ constraints behave differently in the single and multi-tracer case. The single-tracer case is more sensitive to k_{\max} than the multi-tracer, which leads to a larger improvement at lower k_{\max} cuts. This indicates the significance of the multi-tracer approach, and especially for the bispectrum, in upcoming surveys. In cases where the small scales are lost due to the survey’s design and systematics, the multi-tracer approach becomes ever more important.

6 Summary and conclusion

In this work we investigate the potential of the multi-tracer power spectrum and multi-tracer bispectrum analysis in improving the constraints on the amplitude of PNG, with a particular focus on local PNG. We consider two nominal next-generation surveys and use the Fisher matrix to probe the information on $f_{\text{NL}}^{\text{loc}}$, within the power spectra and bispectra of the two tracers. The main goal is to quantify the impact of the various tracer-dependent parameters on the constraining power on $f_{\text{NL}}^{\text{loc}}$, in order to maximise the gain from the multi-tracer technique and provide the tracer-selection criteria for future galaxy surveys. This is done by fixing the survey specifications for each single-tracer and the multi-tracer, which effectively means comparing the potential gain from combining the samples or not over the same volume and redshift range. The main findings can be summarised as follows:

- The complete multi-tracer approach takes advantage of increasing number density in an optimal way, by maximising the contribution of the cross-terms. This is particularly important for high-density samples, as well as for tracers that approach the cosmic variance limit. The scenario where the number densities of the tracers are similar does not provide the tightest constraints, except for high-density samples.
- The constraining power of the power spectrum and bispectrum multi-tracers is $\propto |b_1^t b_\Phi^{t'} - b_1^{t'} b_\Phi^t|$, where the power spectrum exhibits a much larger gain compared to the

bispectrum. In particular the optimal regime is for an increasing difference between the linear biases of the two tracers, as well as for a negative b_Φ for one of the two tracers.

- Maximising the quantity $|b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}|$ is not enough for to fully exploit the potential of the multi-tracer bispectrum. Increasing as well the second-order bias difference $|b_2^t - b_2^{t'}|$ provides a significant gain in the bispectrum constraints, compared to only increasing $|b_1^{t'} b_\Phi^t - b_1^t b_\Phi^{t'}|$.
- The findings of this work demonstrate a substantial improvement in constraining f_{NL} through the utilization of the multi-tracer bispectrum. However, realizing this benefit necessitates satisfying the criteria outlined in section 5, similar to the specifications employed by the nominal surveys in this analysis. These criteria involve maximising the number density ratio, as well as the difference between the linear biases, quadratic biases, and the products b_1 , b_Φ for each individual tracer.
- The advantage exhibited by the multi-tracer approach over a single-tracer analysis, for both power spectrum and bispectrum, could be maintained in a realistic setup, where the accessible scales can be limited. In particular, a bispectrum multi-tracer analysis could improve single-tracer constraints on $f_{\text{NL}}^{\text{loc}}$ up to $\sim 50\%$, for an analysis restricted within the very large linear and not shot-noise dominated scales.

The results presented here do not consider correlations between the various parameters, within the Fisher matrix formalism, but address only the case of the conditional $f_{\text{NL}}^{\text{loc}}$ error. Correlations are important since the power of the multi-tracer approach can be applied to reducing or breaking the various degeneracies between parameters. In order to fully investigate the parametric behaviour of the multi-tracer technique we need to perform a complete Fisher analysis, including the various parameters of the model. This is the subject of the follow-up paper [73].

We considered only the full shape of the power spectrum and bispectrum. We leave the study of the possible gains when using different types of data compression of the bispectrum for follow-up work. Future work will also investigate the inclusion of relativistic and wide-angle effects, which have been shown to be important in the multi-tracer power spectrum and single-tracer bispectrum.

Acknowledgments

We acknowledge an anonymous reviewer for comments that helped us significantly improve the paper. We thank Daisuke Yamauchi for helpful clarifications of their work [30]. DK and RM are supported by the South African Radio Astronomy Observatory and the National Research Foundation (Grant No. 75415). JF is supported by Fundação para a Ciência e a Tecnologia (FCT) through the research grants UIDB/04434/2020 and UIDP/04434/2020 and through the Investigador FCT Contract No. 2020.02633.CEECIND/CP1631/CT0002. SC acknowledges support from the ‘Departments of Excellence 2018-2022’ Grant (L. 232/2016) awarded by the Italian Ministry of University and Research (MUR) and from the ‘Ministero degli Affari Esteri della Cooperazione Internazionale (MAECI) — Direzione Generale per la

Promozione del Sistema Paese Progetto di Grande Rilevanza ZA18GR02. CC is supported by the U.K. Science & Technology Facilities Council Consolidated Grant ST/T000341/1. JF, SC and CC thank the University of the Western Cape for supporting a visit to Cape Town, during which this work was initiated. This work made use of the South African Centre for High Performance Computing, under the project *Cosmology with Radio Telescopes*, ASTRO-0945.

A Redshift-space kernels and damping factors

The general non-Gaussian redshift kernels up to second order for tracer I , neglecting $\mathcal{O}(f_{\text{NL}}^2)$ terms, are given by (see e.g. [77, 82, 83]):

$$Z_1^I(\mathbf{k}_i) = b_1^I + f \mu_i^2 + f_{\text{NL}} \frac{b_{\Phi}^I k_i^\alpha}{M(k_i)}, \quad (\text{A.1})$$

$$\begin{aligned} Z_2^I(\mathbf{k}_i, \mathbf{k}_j) &= b_1^I F_2(\mathbf{k}_i, \mathbf{k}_j) + f \mu_{ij}^2 G_2(\mathbf{k}_i, \mathbf{k}_j) + \frac{b_2^I}{2} + \frac{b_s^I}{2} S_2(\mathbf{k}_i, \mathbf{k}_j) \\ &+ \frac{f \mu_{ij} k_{ij}}{2} \left[\frac{\mu_i}{k_i} Z_1^I(\mathbf{k}_j) + \frac{\mu_j}{k_j} Z_1^I(\mathbf{k}_i) \right] + f_{\text{NL}} \frac{b_{\Psi\delta}^I}{2} \left[\frac{k_i^\alpha}{M(k_i)} + \frac{k_j^\alpha}{M(k_j)} \right] \\ &- f_{\text{NL}} b_{\Psi}^I N_2(\mathbf{k}_i, \mathbf{k}_j). \end{aligned} \quad (\text{A.2})$$

Here $f(z)$ is the linear growth rate, $\mu_i = \hat{\mathbf{k}}_i \cdot \hat{\mathbf{z}}$, with $\hat{\mathbf{z}}$ the line-of-sight direction, $\mu_{ij} = (\mu_i k_i + \mu_j k_j)/k_{ij}$, and $k_{ij}^2 = (\mathbf{k}_i + \mathbf{k}_j)^2$. The kernels $F_2(\mathbf{k}_i, \mathbf{k}_j)$ and $G_2(\mathbf{k}_i, \mathbf{k}_j)$ are the second-order symmetric SPT kernels (e.g. [47]), while $S_2(\mathbf{k}_1, \mathbf{k}_2) = (\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)^2 - 1/3$ is the tidal kernel [54, 84]. The symmetrised N_2 kernel is

$$N_2(\mathbf{k}_i, \mathbf{k}_j) = \frac{1}{2} \left[\frac{k_i^\alpha}{k_j^2 M(k_i)} + \frac{k_j^\alpha}{k_i^2 M(k_j)} \right] \mathbf{k}_i \cdot \mathbf{k}_j, \quad (\text{A.3})$$

which encodes the coupling of the PNG potential to the Eulerian-to-Lagrangian displacement field [74, 82]. The parameter $\alpha = 2, 1$ for equilateral, orthogonal shapes [44, 61, 75, 76] and $\alpha = 0$ for local shapes [55, 56, 74].

The non-perturbative RSD fingers-of-god effect, which models the reduction in clustering power due to nonlinear velocities, is described by a phenomenological model, following [85, 86]. Here we use an exponential damping factor, which in Fourier space is a radial convolution, $\delta_{\mathbf{k}} \rightarrow \delta_{\mathbf{k}} \exp(-k^2 \mu^2 \sigma/2)$. Furthermore, in the case of an optical survey, redshift errors smear the galaxy density field along the line-of-sight, which in Fourier space can be described by $\delta_{\mathbf{k}} \rightarrow \delta_{\mathbf{k}} \exp(-k^2 \mu^2 \sigma_r/2)$ [87]. The redshift errors propagate into comoving distance errors, where $\sigma_r(z) = c/H(z)\sigma_z(z)$ and $\sigma_z(z)$ is the redshift error of the survey. Following the most general approach, we assume that the damping amplitude is tracer-dependent, so that

$$\mathcal{D}_P^{IJ}(\mathbf{k}) = \exp \left[-\frac{1}{2} k^2 \mu^2 \left\{ (\sigma_P^I)^2 + (\sigma_r^I)^2 + (\sigma_P^J)^2 + (\sigma_r^J)^2 \right\} \right], \quad (\text{A.4})$$

$$\begin{aligned} \mathcal{D}_B^{IJK}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \exp \left[-\frac{1}{2} \left\{ k_1^2 \mu_1^2 \left[(\sigma_B^I)^2 + (\sigma_r^I)^2 \right] + k_2^2 \mu_2^2 \left[(\sigma_B^J)^2 + (\sigma_r^J)^2 \right] \right. \right. \\ &\quad \left. \left. + k_3^2 \mu_3^2 \left[(\sigma_B^K)^2 + (\sigma_r^K)^2 \right] \right\} \right], \end{aligned} \quad (\text{A.5})$$

where the redshift dependence has been suppressed for brevity. The amplitude parameters for RSD and redshift errors, for power spectra and bispectra, are independent free parameters. The fiducial values for RSD are given by the velocity dispersion, i.e. $\sigma_P^I(z) = \sigma_B^I(z) = \sigma_v(z)$.

B 3-tracer case

For three tracers, I, J, K take values t, t', t'' , and the power spectrum data vector is

$$\mathbf{D}_P = [P^{tt}, P^{tt'}, P^{t't'}, P^{t't''}, P^{t''t''}, P^{tt''}], \quad (\text{B.1})$$

where the expected values of the entries can be derived as in the case of two tracers, by changing accordingly the tracer-dependent quantities in eq. (3.10). Likewise for the bispectrum data vector in the case of three tracers:

$$\mathbf{D}_B = [B^{ttt}, B^{(ttt')}, B^{(tt't')}, B^{t't't'}, B^{(t't't'')}, B^{(t''t''t'')}, B^{(ttt'')}, B^{(tt't'')}, B^{(tt't'')}], \quad (\text{B.2})$$

where the expected values for the first 9 entries can be derived from by exchanging accordingly the tracer indexes in the tracer-dependent quantities. The expected value of the last entry can be defined as

$$\begin{aligned} \langle \hat{B}^{(tt't'')} \rangle \equiv B^{(tt't'')}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \frac{1}{6} \left[\hat{B}^{tt't''}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \hat{B}^{t''t't}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \hat{B}^{t't't''}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \right. \\ &\quad \left. + \hat{B}^{t''t't'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \hat{B}^{t't't''}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \hat{B}^{t''t't'}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \right], \end{aligned} \quad (\text{B.3})$$

where

$$\begin{aligned} \langle \hat{B}^{t't't''} \rangle \equiv \hat{B}^{t't't''}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= \mathcal{D}_B^{t't't''}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \left[Z_1^t(\mathbf{k}_1) Z_1^{t'}(\mathbf{k}_2) Z_1^{t''}(\mathbf{k}_3) B_{\text{PNG}}(k_1, k_2, k_3) \right. \\ &\quad + 2 Z_1^t(\mathbf{k}_1) Z_1^{t'}(\mathbf{k}_2) Z_2^{t''}(\mathbf{k}_1, \mathbf{k}_2) P(k_1) P(k_2) \\ &\quad + 2 Z_1^{t'}(\mathbf{k}_2) Z_1^{t''}(\mathbf{k}_3) Z_2^t(\mathbf{k}_2, \mathbf{k}_3) P(k_2) P(k_3) \\ &\quad \left. + 2 Z_1^t(\mathbf{k}_1) Z_1^{t''}(\mathbf{k}_3) Z_2^{t'}(\mathbf{k}_3, \mathbf{k}_1) P(k_3) P(k_1) \right], \end{aligned} \quad (\text{B.4})$$

and similarly for the other terms, by appropriately permutating the tracer indexes. Note that the symmetrised bispectrum of eq. (B.3) has no stochastic terms due to the assumption $P_{\varepsilon\varepsilon\delta}^{IJ} = B_{\varepsilon}^{IJK} = 0$.

The data vector of the combined power spectra and bispectra is defined by combining the data vectors of eqs. (B.1) and (B.2), creating a 16 element array. The covariance of this vector is again a 2×2 block matrix, as in the two-tracer case of eq. (4.2), but now each individual block will have different dimensions. To be more precise:

- $\text{Cov}(\mathbf{P}_i, \mathbf{P}_j)$ is a 6×6 matrix,
- $\text{Cov}(\mathbf{B}_i, \mathbf{B}_b)$ is a 10×10 matrix,
- $\text{Cov}(\mathbf{P}_i, \mathbf{B}_b)$ is a 6×10 matrix.

As in the case of two tracers, we consider only the Gaussian part of the covariances, i.e. each sub-block is a diagonal matrix. The expression for each sub-block in the $\text{Cov}(\mathbf{P}_i, \mathbf{P}_j)$ block can be easily derived from eq. (4.4), by permuting the tracer indexes accordingly. The same follows for most of the sub-blocks of $\text{Cov}(\mathbf{B}_i, \mathbf{B}_b)$, where all the expressions for the diagonal sub-blocks can be derived from eq. (4.6) by interchanging the tracer indices accordingly. The only sub-blocks left to define are those that correspond to the correlations with the symmetrised bispectrum $B^{(tt't'')}$:

$$\begin{aligned}
 \text{Cov}\left(B_a^{ttt}, B_b^{(tt't'')}\right) &= \frac{1}{2}\left(P_{a_1}^{tt} P_{a_2}^{tt'} P_{a_3}^{tt''} + P_{a_1}^{tt} P_{a_2}^{tt''} P_{a_3}^{tt'}\right) + 2 \text{ perm.}, \\
 \text{Cov}\left(B_a^{(ttt')}, B_b^{(tt't'')}\right) &= \frac{1}{6}\left[P_{a_1}^{tt}\left(P_{a_2}^{tt''} P_{a_3}^{t't'} + P_{a_2}^{t't'} P_{a_3}^{tt''}\right) + P_{a_1}^{tt}\left(P_{a_2}^{tt'} P_{a_3}^{t't''} + P_{a_2}^{t't''} P_{a_3}^{tt'}\right)\right. \\
 &\quad \left.+ 2P_{a_1}^{tt'} P_{a_2}^{tt''} P_{a_3}^{t't''}\right] + 2 \text{ perm.}, \\
 \text{Cov}\left(B_a^{(tt't'')}, B_b^{(tt't'')}\right) &= \frac{1}{6}\left[P_{a_1}^{tt'}\left(P_{a_2}^{tt''} P_{a_3}^{t't''} + P_{a_2}^{t't''} P_{a_3}^{tt''}\right) + \frac{1}{2}P_{a_1}^{tt}\left(P_{a_2}^{t't'} P_{a_3}^{t't''} + P_{a_2}^{t't''} P_{a_3}^{t't'}\right)\right. \\
 &\quad \left.+ P_{a_1}^{tt''} P_{a_2}^{t't''} P_{a_3}^{t't'} + P_{a_1}^{tt} P_{a_2}^{t't''} P_{a_3}^{t't''} + P_{a_1}^{tt'} P_{a_2}^{t't'} P_{a_3}^{t't''}\right] + 2 \text{ perm.}.
 \end{aligned} \tag{B.5}$$

The remaining sub-block covariances can be retrieved by changing accordingly the tracer indices in the above expressions.

The additional expressions needed to derive the power spectra-bispectra cross-covariance in the case of three tracers are:

$$\begin{aligned}
 \text{Cov}\left(P_i^{tt}, B_b^{(tt't'')}\right) &= \frac{1}{3}\left[P^{tt''}(\mathbf{k}_i) B_{2\delta}^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + P^{tt'}(\mathbf{k}_i) B_{2\delta}^{ttt''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right. \\
 &\quad \left.+ P^{tt}(\mathbf{k}_i) B_{1\delta}^{tt't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right], \\
 \text{Cov}\left(P_i^{tt'}, B_b^{(tt't'')}\right) &= \frac{1}{6}\left[P^{tt'}(\mathbf{k}_i) B_{2\delta}^{tt't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + P^{tt}(\mathbf{k}_i) B_{1\delta}^{t't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right. \\
 &\quad \left.+ P^{tt''}(\mathbf{k}_i) B_{1\delta}^{tt't'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + P^{t't'}(\mathbf{k}_i) B_{1\delta}^{ttt''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right. \\
 &\quad \left.+ P^{t't''}(\mathbf{k}_i) B_{1\delta}^{ttt'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right],
 \end{aligned} \tag{B.6}$$

where

$$\begin{aligned}
 B_{1\delta}^{tt't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) &= \delta_{ib_1}\left[B^{tt't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + B^{tt''t'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right] \\
 &\quad + \delta_{ib_2}\left[B^{t't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + B^{t''t'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right] \\
 &\quad + \delta_{ib_3}\left[B^{t''t't}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + B^{t't''t}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right],
 \end{aligned} \tag{B.7}$$

$$\begin{aligned}
 B_{2\delta}^{tt't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) &= (\delta_{ib_1} + \delta_{ib_2})\left[B^{tt't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + B^{t't''}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right] \\
 &\quad + (\delta_{ib_1} + \delta_{ib_3})\left[B^{tt''t'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + B^{t''t't}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right] \\
 &\quad + (\delta_{ib_2} + \delta_{ib_3})\left[B^{t''t't}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3}) + B^{t't''t'}(\mathbf{k}_{b_1}, \mathbf{k}_{b_2}, \mathbf{k}_{b_3})\right].
 \end{aligned} \tag{B.8}$$

References

- [1] S. Novell-Masot, D. Gualdi, H. Gil-Marín and L. Verde, *GEO-FPT: a model of the galaxy bispectrum at mildly non-linear scales*, *JCAP* **11** (2023) 044 [[arXiv:2303.15510](#)] [[INSPIRE](#)].
- [2] N.S. Sugiyama et al., *New constraints on cosmological modified gravity theories from anisotropic three-point correlation functions of BOSS DR12 galaxies*, *Mon. Not. Roy. Astron. Soc.* **523** (2023) 3133 [[arXiv:2302.06808](#)] [[INSPIRE](#)].
- [3] M.M. Ivanov, O.H.E. Philcox, G. Cabass, T. Nishimichi, M. Simonović and M. Zaldarriaga, *Cosmology with the galaxy bispectrum multipoles: Optimal estimation and application to BOSS data*, *Phys. Rev. D* **107** (2023) 083515 [[arXiv:2302.04414](#)] [[INSPIRE](#)].
- [4] D. Karagiannis, R. Maartens and L.F. Randrianjanahary, *Cosmological constraints from the power spectrum and bispectrum of 21 cm intensity maps*, *JCAP* **11** (2022) 003 [[arXiv:2206.07747](#)] [[INSPIRE](#)].
- [5] O.H.E. Philcox and M.M. Ivanov, *BOSS DR12 full-shape cosmology: Λ CDM constraints from the large-scale galaxy power spectrum and bispectrum monopole*, *Phys. Rev. D* **105** (2022) 043517 [[arXiv:2112.04515](#)] [[INSPIRE](#)].
- [6] M.M. Ivanov, O.H.E. Philcox, T. Nishimichi, M. Simonović, M. Takada and M. Zaldarriaga, *Precision analysis of the redshift-space galaxy bispectrum*, *Phys. Rev. D* **105** (2022) 063512 [[arXiv:2110.10161](#)] [[INSPIRE](#)].
- [7] A. Moradinezhad Dizgah, M. Biagetti, E. Sefusatti, V. Desjacques and J. Noreña, *Primordial Non-Gaussianity from Biased Tracers: Likelihood Analysis of Real-Space Power Spectrum and Bispectrum*, *JCAP* **05** (2021) 015 [[arXiv:2010.14523](#)] [[INSPIRE](#)].
- [8] N.S. Sugiyama, S. Saito, F. Beutler and H.-J. Seo, *A complete FFT-based decomposition formalism for the redshift-space bispectrum*, *Mon. Not. Roy. Astron. Soc.* **484** (2019) 364 [[arXiv:1803.02132](#)] [[INSPIRE](#)].
- [9] C. Clarkson, E.M. de Weerd, S. Jolicoeur, R. Maartens and O. Umeh, *The dipole of the galaxy bispectrum*, *Mon. Not. Roy. Astron. Soc.* **486** (2019) L101 [[arXiv:1812.09512](#)] [[INSPIRE](#)].
- [10] R. Maartens, S. Jolicoeur, O. Umeh, E.M. De Weerd, C. Clarkson and S. Camera, *Detecting the relativistic galaxy bispectrum*, *JCAP* **03** (2020) 065 [[arXiv:1911.02398](#)] [[INSPIRE](#)].
- [11] S. Jolicoeur, R. Maartens, E.M. De Weerd, O. Umeh, C. Clarkson and S. Camera, *Detecting the relativistic bispectrum in 21 cm intensity maps*, *JCAP* **06** (2021) 039 [[arXiv:2009.06197](#)] [[INSPIRE](#)].
- [12] O. Umeh, S. Jolicoeur, R. Maartens and C. Clarkson, *A general relativistic signature in the galaxy bispectrum: the local effects of observing on the lightcone*, *JCAP* **03** (2017) 034 [[arXiv:1610.03351](#)] [[INSPIRE](#)].
- [13] E. Di Dio et al., *Non-Gaussianities due to Relativistic Corrections to the Observed Galaxy Bispectrum*, *JCAP* **03** (2017) 006 [[arXiv:1611.03720](#)] [[INSPIRE](#)].
- [14] R. Maartens, S. Jolicoeur, O. Umeh, E.M. De Weerd and C. Clarkson, *Local primordial non-Gaussianity in the relativistic galaxy bispectrum*, *JCAP* **04** (2021) 013 [[arXiv:2011.13660](#)] [[INSPIRE](#)].
- [15] M. Noorikuhani and R. Scoccimarro, *Wide-angle and relativistic effects in Fourier-space clustering statistics*, *Phys. Rev. D* **107** (2023) 083528 [[arXiv:2207.12383](#)] [[INSPIRE](#)].
- [16] E. Di Dio, R. Durrer, R. Maartens, F. Montanari and O. Umeh, *The Full-Sky Angular Bispectrum in Redshift Space*, *JCAP* **04** (2019) 053 [[arXiv:1812.09297](#)] [[INSPIRE](#)].

- [17] K. Garcia and Z. Slepian, *Improving the line of sight for the anisotropic 3-point correlation function of galaxies: Centroid and Unit-Vector-Average methods scaling as $\mathcal{O}(N^2)$* , *Mon. Not. Roy. Astron. Soc.* **515** (2022) 1199 [[arXiv:2011.03503](#)] [[INSPIRE](#)].
- [18] K. Pardede, E. Di Dio and E. Castorina, *Wide-angle effects in the galaxy bispectrum*, *JCAP* **09** (2023) 030 [[arXiv:2302.12789](#)] [[INSPIRE](#)].
- [19] D. Karagiannis, A. Slosar and M. Liguori, *Forecasts on Primordial non-Gaussianity from 21 cm Intensity Mapping experiments*, *JCAP* **11** (2020) 052 [[arXiv:1911.03964](#)] [[INSPIRE](#)].
- [20] S. Bharadwaj, A. Mazumdar and D. Sarkar, *Quantifying the Redshift Space Distortion of the Bispectrum. Part I. Primordial Non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **493** (2020) 594 [[arXiv:2001.10243](#)] [[INSPIRE](#)].
- [21] D. Karagiannis, J. Fonseca, R. Maartens and S. Camera, *Probing primordial non-Gaussianity with the power spectrum and bispectrum of future 21 cm intensity maps*, *Phys. Dark Univ.* **32** (2021) 100821 [[arXiv:2010.07034](#)] [[INSPIRE](#)].
- [22] G. D’Amico, M. Lewandowski, L. Senatore and P. Zhang, *Limits on primordial non-Gaussianities from BOSS galaxy-clustering data*, [arXiv:2201.11518](#) [[INSPIRE](#)].
- [23] W.R. Coulton et al., *Quijote-PNG: The Information Content of the Halo Power Spectrum and Bispectrum*, *Astrophys. J.* **943** (2023) 178 [[arXiv:2206.15450](#)] [[INSPIRE](#)].
- [24] G. Jung et al., *Quijote-PNG: Quasi-maximum Likelihood Estimation of Primordial Non-Gaussianity in the Nonlinear Halo Density Field*, *Astrophys. J.* **948** (2023) 135 [[arXiv:2211.07565](#)] [[INSPIRE](#)].
- [25] U. Seljak, *Extracting primordial non-Gaussianity without cosmic variance*, *Phys. Rev. Lett.* **102** (2009) 021302 [[arXiv:0807.1770](#)] [[INSPIRE](#)].
- [26] P. McDonald and U. Seljak, *How to measure redshift-space distortions without sample variance*, *JCAP* **10** (2009) 007 [[arXiv:0810.0323](#)] [[INSPIRE](#)].
- [27] N. Hamaus, U. Seljak and V. Desjacques, *Optimal Constraints on Local Primordial Non-Gaussianity from the Two-Point Statistics of Large-Scale Structure*, *Phys. Rev. D* **84** (2011) 083509 [[arXiv:1104.2321](#)] [[INSPIRE](#)].
- [28] L.R. Abramo and K.E. Leonard, *Why multi-tracer surveys beat cosmic variance*, *Mon. Not. Roy. Astron. Soc.* **432** (2013) 318 [[arXiv:1302.5444](#)] [[INSPIRE](#)].
- [29] L.D. Ferramacho, M.G. Santos, M.J. Jarvis and S. Camera, *Radio galaxy populations and the multitracers technique: pushing the limits on primordial non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **442** (2014) 2511 [[arXiv:1402.2290](#)] [[INSPIRE](#)].
- [30] D. Yamauchi, K. Takahashi and M. Oguri, *Constraining primordial non-Gaussianity via a multitracers technique with surveys by Euclid and the Square Kilometre Array*, *Phys. Rev. D* **90** (2014) 083520 [[arXiv:1407.5453](#)] [[INSPIRE](#)].
- [31] D. Alonso and P.G. Ferreira, *Constraining ultralarge-scale cosmology with multiple tracers in optical and radio surveys*, *Phys. Rev. D* **92** (2015) 063525 [[arXiv:1507.03550](#)] [[INSPIRE](#)].
- [32] J. Fonseca, S. Camera, M. Santos and R. Maartens, *Hunting down horizon-scale effects with multi-wavelength surveys*, *Astrophys. J. Lett.* **812** (2015) L22 [[arXiv:1507.04605](#)] [[INSPIRE](#)].
- [33] J. Fonseca, R. Maartens and M.G. Santos, *Probing the primordial Universe with MeerKAT and DES*, *Mon. Not. Roy. Astron. Soc.* **466** (2017) 2780 [[arXiv:1611.01322](#)] [[INSPIRE](#)].
- [34] J. Gleyzes, R. de Putter, D. Green and O. Doré, *Biasing and the search for primordial non-Gaussianity beyond the local type*, *JCAP* **04** (2017) 002 [[arXiv:1612.06366](#)] [[INSPIRE](#)].

- [35] M. Ballardini, W.L. Matthewson and R. Maartens, *Constraining primordial non-Gaussianity using two galaxy surveys and CMB lensing*, *Mon. Not. Roy. Astron. Soc.* **489** (2019) 1950 [[arXiv:1906.04730](#)] [[INSPIRE](#)].
- [36] Z. Gomes, S. Camera, M.J. Jarvis, C. Hale and J. Fonseca, *Non-Gaussianity constraints using future radio continuum surveys and the multitracer technique*, *Mon. Not. Roy. Astron. Soc.* **492** (2020) 1513 [[arXiv:1912.08362](#)] [[INSPIRE](#)].
- [37] J.-A. Viljoen, J. Fonseca and R. Maartens, *Multi-wavelength spectroscopic probes: prospects for primordial non-Gaussianity and relativistic effects*, *JCAP* **11** (2021) 010 [[arXiv:2107.14057](#)] [[INSPIRE](#)].
- [38] A. Barreira, *Predictions for local PNG bias in the galaxy power spectrum and bispectrum and the consequences for f_{NL} constraints*, *JCAP* **01** (2022) 033 [[arXiv:2107.06887](#)] [[INSPIRE](#)].
- [39] S. Jolicoeur, R. Maartens and S. Dlamini, *Constraining primordial non-Gaussianity by combining next-generation galaxy and 21 cm intensity mapping surveys*, *Eur. Phys. J. C* **83** (2023) 320 [[arXiv:2301.02406](#)] [[INSPIRE](#)].
- [40] A. Barreira and E. Krause, *Towards optimal and robust f_{NL} constraints with multi-tracer analyses*, *JCAP* **10** (2023) 044 [[arXiv:2302.09066](#)] [[INSPIRE](#)].
- [41] J.M. Sullivan, T. Priyon and U. Seljak, *Learning to concentrate: multi-tracer forecasts on local primordial non-Gaussianity with machine-learned bias*, *JCAP* **08** (2023) 004 [[arXiv:2303.08901](#)] [[INSPIRE](#)].
- [42] D. Yamauchi, S. Yokoyama and K. Takahashi, *Multitracer technique for galaxy bispectrum: An application to constraints on nonlocal primordial non-Gaussianities*, *Phys. Rev. D* **95** (2017) 063530 [[arXiv:1611.03590](#)] [[INSPIRE](#)].
- [43] S. Camera, M.G. Santos and R. Maartens, *Probing primordial non-Gaussianity with SKA galaxy redshift surveys: a fully relativistic analysis*, *Mon. Not. Roy. Astron. Soc.* **448** (2015) 1035 [[arXiv:1409.8286](#)] [[INSPIRE](#)].
- [44] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rep.* **733** (2018) 1 [[arXiv:1611.09787](#)] [[INSPIRE](#)].
- [45] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [[astro-ph/9911177](#)] [[INSPIRE](#)].
- [46] E. Komatsu et al., *Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe*, *Astro2010: The Astronomy and Astrophysics Decadal Survey*, Science White Papers volume 158, [arXiv:0902.4759](#) [[INSPIRE](#)].
- [47] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, *Phys. Rep.* **367** (2002) 1 [[astro-ph/0112551](#)] [[INSPIRE](#)].
- [48] V. Assassi, D. Baumann, D. Green and M. Zaldarriaga, *Renormalized Halo Bias*, *JCAP* **08** (2014) 056 [[arXiv:1402.5916](#)] [[INSPIRE](#)].
- [49] L. Senatore, *Bias in the Effective Field Theory of Large Scale Structures*, *JCAP* **11** (2015) 007 [[arXiv:1406.7843](#)] [[INSPIRE](#)].
- [50] M. Mirbabayi, F. Schmidt and M. Zaldarriaga, *Biased Tracers and Time Evolution*, *JCAP* **07** (2015) 030 [[arXiv:1412.5169](#)] [[INSPIRE](#)].
- [51] A. Dekel and O. Lahav, *Stochastic nonlinear galaxy biasing*, *Astrophys. J.* **520** (1999) 24 [[astro-ph/9806193](#)] [[INSPIRE](#)].

- [52] A. Taruya and J. Soda, *Stochastic biasing and galaxy mass density relation in the weakly nonlinear regime*, *Astrophys. J.* **522** (1999) 46 [[astro-ph/9809204](#)] [[INSPIRE](#)].
- [53] T. Matsubara, *Stochasticity of bias and nonlocality of galaxy formation: Linear scales*, *Astrophys. J.* **525** (1999) 543 [[astro-ph/9906029](#)] [[INSPIRE](#)].
- [54] T. Baldauf, U. Seljak, V. Desjacques and P. McDonald, *Evidence for Quadratic Tidal Tensor Bias from the Halo Bispectrum*, *Phys. Rev. D* **86** (2012) 083540 [[arXiv:1201.4827](#)] [[INSPIRE](#)].
- [55] N. Dalal, O. Doré, D. Huterer and A. Shirokov, *The imprints of primordial non-Gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects*, *Phys. Rev. D* **77** (2008) 123514 [[arXiv:0710.4560](#)] [[INSPIRE](#)].
- [56] A. Slosar, C. Hirata, U. Seljak, S. Ho and N. Padmanabhan, *Constraints on local primordial non-Gaussianity from large scale structure*, *JCAP* **08** (2008) 031 [[arXiv:0805.3580](#)] [[INSPIRE](#)].
- [57] S. Matarrese and L. Verde, *The effect of primordial non-Gaussianity on halo bias*, *Astrophys. J. Lett.* **677** (2008) L77 [[arXiv:0801.4826](#)] [[INSPIRE](#)].
- [58] L. Verde and S. Matarrese, *Detectability of the effect of Inflationary non-Gaussianity on halo bias*, *Astrophys. J. Lett.* **706** (2009) L91 [[arXiv:0909.3224](#)] [[INSPIRE](#)].
- [59] N. Afshordi and A.J. Tolley, *Primordial non-Gaussianity, statistics of collapsed objects, and the Integrated Sachs-Wolfe effect*, *Phys. Rev. D* **78** (2008) 123507 [[arXiv:0806.1046](#)] [[INSPIRE](#)].
- [60] V. Desjacques and U. Seljak, *Primordial non-Gaussianity in the large scale structure of the Universe*, *Adv. Astron.* **2010** (2010) 908640 [[arXiv:1006.4763](#)] [[INSPIRE](#)].
- [61] F. Schmidt and M. Kamionkowski, *Halo Clustering with Non-Local Non-Gaussianity*, *Phys. Rev. D* **82** (2010) 103002 [[arXiv:1008.0638](#)] [[INSPIRE](#)].
- [62] R. Scoccimarro, L. Hui, M. Manera and K.C. Chan, *Large-scale Bias and Efficient Generation of Initial Conditions for Non-Local Primordial Non-Gaussianity*, *Phys. Rev. D* **85** (2012) 083002 [[arXiv:1108.5512](#)] [[INSPIRE](#)].
- [63] V. Desjacques, D. Jeong and F. Schmidt, *Accurate Predictions for the Scale-Dependent Galaxy Bias from Primordial Non-Gaussianity*, *Phys. Rev. D* **84** (2011) 061301 [[arXiv:1105.3476](#)] [[INSPIRE](#)].
- [64] F. Schmidt, D. Jeong and V. Desjacques, *Peak-Background Split, Renormalization, and Galaxy Clustering*, *Phys. Rev. D* **88** (2013) 023515 [[arXiv:1212.0868](#)] [[INSPIRE](#)].
- [65] V. Assassi, D. Baumann and F. Schmidt, *Galaxy Bias and Primordial Non-Gaussianity*, *JCAP* **12** (2015) 043 [[arXiv:1510.03723](#)] [[INSPIRE](#)].
- [66] W.L.W. Sargent and E.L. Turner, *A statistical method for determining the cosmological density parameter from the redshifts of a complete sample of galaxies*, *Astrophys. J.* **212** (1977) L3.
- [67] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1 [[INSPIRE](#)].
- [68] A.J.S. Hamilton, *Linear redshift distortions: A Review*, in *The Evolving Universe, Astrophysics and Space Science Library* **231**, D. Hamilton ed., Springer, Dordrecht, The Netherlands (1998), pp. 185–275 [[DOI:10.1007/978-94-011-4960-0_17](#)] [[astro-ph/9708102](#)] [[INSPIRE](#)].
- [69] T. Mergulhão, H. Rubira, R. Voivodic and L.R. Abramo, *The effective field theory of large-scale structure and multi-tracer*, *JCAP* **04** (2022) 021 [[arXiv:2108.11363](#)] [[INSPIRE](#)].
- [70] F. Schmidt, *Towards a self-consistent halo model for the nonlinear large-scale structure*, *Phys. Rev. D* **93** (2016) 063512 [[arXiv:1511.02231](#)] [[INSPIRE](#)].

- [71] E. Sefusatti, M. Crocce, S. Pueblas and R. Scoccimarro, *Cosmology and the Bispectrum*, *Phys. Rev. D* **74** (2006) 023522 [[astro-ph/0604505](#)] [[INSPIRE](#)].
- [72] M. Biagetti, L. Castiblanco, J. Noreña and E. Sefusatti, *The covariance of squeezed bispectrum configurations*, *JCAP* **09** (2022) 009 [[arXiv:2111.05887](#)] [[INSPIRE](#)].
- [73] D. Karagiannis, R. Maartens, J. Fonseca, S. Camera and C. Clarkson, *Multi-tracer power spectra and bispectra: Forecasts on primordial non-Gaussianity*, to appear.
- [74] T. Giannantonio and C. Porciani, *Structure formation from non-Gaussian initial conditions: multivariate biasing, statistics, and comparison with N-body simulations*, *Phys. Rev. D* **81** (2010) 063530 [[arXiv:0911.0017](#)] [[INSPIRE](#)].
- [75] T. Giannantonio, C. Porciani, J. Carron, A. Amara and A. Pillepich, *Constraining primordial non-Gaussianity with future galaxy surveys*, *Mon. Not. Roy. Astron. Soc.* **422** (2012) 2854 [[arXiv:1109.0958](#)] [[INSPIRE](#)].
- [76] G. Cabass, E. Pajer and F. Schmidt, *Imprints of Oscillatory Bispectra on Galaxy Clustering*, *JCAP* **09** (2018) 003 [[arXiv:1804.07295](#)] [[INSPIRE](#)].
- [77] D. Karagiannis, A. Lazanu, M. Liguori, A. Raccanelli, N. Bartolo and L. Verde, *Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys*, *Mon. Not. Roy. Astron. Soc.* **478** (2018) 1341 [[arXiv:1801.09280](#)] [[INSPIRE](#)].
- [78] EUCLID collaboration, *Euclid preparation. Part VII. Forecast validation for Euclid cosmological probes*, *Astron. Astrophys.* **642** (2020) A191 [[arXiv:1910.09273](#)] [[INSPIRE](#)].
- [79] DESI collaboration, *The DESI Experiment. Part I. Science, Targeting, and Survey Design*, [arXiv:1611.00036](#) [[INSPIRE](#)].
- [80] V. Yankelevich and C. Porciani, *Cosmological information in the redshift-space bispectrum*, *Mon. Not. Roy. Astron. Soc.* **483** (2019) 2078 [[arXiv:1807.07076](#)] [[INSPIRE](#)].
- [81] PLANCK collaboration, *Planck 2018 results. Part VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [82] T. Baldauf, U. Seljak and L. Senatore, *Primordial non-Gaussianity in the Bispectrum of the Halo Density Field*, *JCAP* **04** (2011) 006 [[arXiv:1011.1513](#)] [[INSPIRE](#)].
- [83] M. Tellarini, A.J. Ross, G. Tasinato and D. Wands, *Galaxy bispectrum, primordial non-Gaussianity and redshift space distortions*, *JCAP* **06** (2016) 014 [[arXiv:1603.06814](#)] [[INSPIRE](#)].
- [84] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS*, *JCAP* **08** (2009) 020 [[arXiv:0902.0991](#)] [[INSPIRE](#)].
- [85] J.A. Peacock and S.J. Dodds, *Reconstructing the linear power spectrum of cosmological mass fluctuations*, *Mon. Not. Roy. Astron. Soc.* **267** (1994) 1020 [[astro-ph/9311057](#)] [[INSPIRE](#)].
- [86] W.E. Ballinger, J.A. Peacock and A.F. Heavens, *Measuring the cosmological constant with redshift surveys*, *Mon. Not. Roy. Astron. Soc.* **282** (1996) 877 [[astro-ph/9605017](#)] [[INSPIRE](#)].
- [87] H.-J. Seo and D.J. Eisenstein, *Probing dark energy with baryonic acoustic oscillations from future large galaxy redshift surveys*, *Astrophys. J.* **598** (2003) 720 [[astro-ph/0307460](#)] [[INSPIRE](#)].