The Covariance of Squeezed Bispectrum Configurations

Matteo Biagetti\textsuperscript{a,b,c,d}, Lina Castiblanco\textsuperscript{e}, Jorge Noreña\textsuperscript{e} and Emiliano Sefusatti\textsuperscript{c,d,a}

\textsuperscript{a}Institute for Fundamental Physics of the Universe, Via Beirut 2, 34151 Trieste, Italy
\textsuperscript{b}SISSA - International School for Advanced Studies, Via Bonomea 265, 34136 Trieste, Italy
\textsuperscript{c}Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Trieste, via Tiepolo 11, 34143 Trieste, Italy
\textsuperscript{d}Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, via Valerio 2, 34127 Trieste, Italy
\textsuperscript{e}Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4950, Valparaíso, Chile

Abstract. We measure the halo bispectrum covariance in a large set of N-body simulations and compare it with theoretical expectations. We find a large correlation among (even mildly) squeezed halo bispectrum configurations. A similarly large correlation can be found between squeezed triangles and the long-wavelength halo power spectrum. This shows that the diagonal Gaussian contribution fails to describe, even approximately, the full covariance in these cases. We compare our numerical estimate with a model that includes, in addition to the Gaussian one, only the non-Gaussian terms that are large for squeezed configurations. We find that accounting for these large terms in the modeling greatly improves the agreement of the full covariance with simulations. We apply these results to a simple Fisher matrix forecast, and find that constraints on primordial non-Gaussianity are degraded by a factor of \( \sim 2 \) when a non-Gaussian covariance is assumed instead of the diagonal, Gaussian approximation.
Contents

1 Introduction 2

2 The power spectrum and bispectrum covariances 4
  2.1 Estimators 4
  2.2 Power spectrum and bispectrum covariance 5
  2.3 Approximation of the mode-counting factors 7

3 The covariance of squeezed bispectrum configurations 10
  3.1 Order of magnitude estimates 10
  3.2 The covariance in terms of response functions 14
  3.3 Prescription for the covariance and its inverse 15

4 Comparison with N-body simulations 16
  4.1 N-body simulations 17
  4.2 Measurements and binning strategy 17
  4.3 Results 17
  4.4 A $\chi^2$ test for the inverse covariance 24

5 Impact on parameter constraints 25
  5.1 Local Primordial non-Gaussianity 27
  5.2 N-body simulations with non-Gaussian initial conditions 27
  5.3 Fisher matrix forecasts 28

6 Conclusions 30

A Integrals for open and flat triangles 32

B Cross-correlation matrix with correlated noise 33

C Half-inverse test 34
1 Introduction

The next generation of large scale structure surveys such as Euclid [1], the Dark Energy Spectroscopic Instrument (DESI) [2], the Legacy Survey of Space and Time (LSST) [3] and the Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (SPHEREx) [4] are going to map an unprecedented number of galaxies at high redshifts. These maps will describe the distribution of the large-scale structure with great accuracy.

Most of cosmological information is routinely extracted from the 2-point correlation function of the galaxy distribution, or its Fourier-space counterpart, the power spectrum. However, low-redshift galaxy density perturbations are a highly non-Gaussian random field. As such, their statistical properties are described, in addition to 2-point statistics, by higher-order correlation functions, starting with the 3-point correlation function. In this work, we consider in particular the galaxy bispectrum, i.e. the Fourier Transform of the 3-point correlation function. Significant effort has been made to provide an accurate theoretical description of this statistic [5–28], while several works attempted to quantify the additional information potentially provided by the bispectrum [29–32]. On the other hand, its measurement and analysis has been performed from early data-sets [7, 33] up to the latest BOSS survey [34–38].

Despite such efforts, the analysis of the galaxy bispectrum has not yet reached the same maturity as the treatment of the galaxy power spectrum. One of the main challenges is the accurate estimation of the bispectrum covariance properties. The straightforward method consists in a direct measurement from a large set of mocks, typically obtained with approximate large-scale structure distributions based on Lagrangian Perturbation Theory [39, 40]. This has been used to obtain a full covariance matrix for bispectrum measurements in tests of theoretical modeling or forecasts [13, 17, 32, 41–52], as well as in actual data analysis [7, 37]. The clear advantage is the natural inclusion of systematic and window effects in the case of redshift surveys. Moreover, it accounts for all non-Gaussian contributions to the bispectrum and power spectrum-bispectrum cross-covariance matrices. ¹ Yet, current sets of mock catalogs need a large number of realizations (on the order of a few thousands, see e.g. [53]) for the precise determination of the covariance of the power spectrum alone. Including thousands of bispectrum configurations require roughly an order of magnitude more realizations.

When a large set of mocks is not available it is common to limit the bispectrum covariance matrix to its diagonal. This can be estimated from a limited set of realizations (see e.g. [54] for a recent implementation) or approximated by its Gaussian expression. This Gaussian approximation is written in terms of the power spectrum alone, in turn, obtained from simulations or in linear theory (see e.g. [55, 56]). These are good approximations, at least for the analysis of large-scale measurements from simulations in boxes with periodic boundary conditions aiming at the determination of bias parameters [47].²

It is reasonable to expect, however, that non-Gaussian contributions could become relevant as we consider smaller scales [43, 48, 63, 64] or in combination with window (i.e. finite-volume) effects

¹However, approximate schemes based on second-order Lagrangian perturbation theory can reproduce the large-scale (tree-level) bispectrum, but provide an incomplete description of the trispectrum and higher-order correlators.
²An alternative approach to the high dimensionality of the covariance matrix relative to the number of mocks available has been to compress the bispectrum to make the covariance in this reduced space tractable purely with mocks; compression has been explored in [32, 57–61, 61, 62].
such as super-sample covariance and local-average (see [45, 65] or [66] and references therein for the power spectrum case). In addition, the leading contribution to the cross-covariance between power spectrum and bispectrum is non-Gaussian and can be quite relevant at large scales as well [44, 52, 67].

These arguments motivated a few works in recent years to explore in more detail an analytic description of the full joint power spectrum and bispectrum \((P + B)\) covariance. A first theoretical prediction of all non-Gaussian contributions to the \(P + B\) covariance matrix, including finite-volume effects, has been studied by [68, 69]. They studied weak lensing statistics in the context of the halo model, therefore focusing on matter correlators at relatively small scales. A comparison of a prediction of the non-Gaussian covariance for the 3D bispectrum of matter and halos is presented instead in [43] but limited to equilateral configurations, showing only a qualitative, overall agreement. It is shown that non-Gaussian contributions, for these triangles, are subdominant w.r.t. the Gaussian one with some exception in the case of sparse halo distributions, essentially due to shot-noise. In [45], the super-sample covariance is estimated for the bispectrum in the response function formalism [70] finding good agreement with simulations. They show it to be a small contribution to the bispectrum covariance, but it is particularly relevant for the power spectrum-bispectrum cross-covariance. In [63], the focus is on the covariance of squeezed triangular configurations of the matter bispectrum, following the results of [71]. For these triangles, the non-Gaussian contributions are shown to be significantly larger than the Gaussian one, already in the quasi-linear regime, but without direct comparison to simulation results. Ref. [49] computes a full perturbation theory prediction for the covariance of the redshift-space multipoles of the power spectrum and bispectrum, the latter measured with the estimator proposed in [24]. This prediction includes all correlators up to the 6-point correlation but in the approximation of linear bias and accounts for survey volume by a simple rescaling of the volume, without including super-sample effects. The comparison with the galaxy mocks produced for the BOSS survey by [53] shows a qualitatively good agreement, with some significant discrepancies particularly for the \(P – B\) cross-covariance. Finally, [48] also provides a comparison of an analytical covariance for the redshift-space \(P\) multipoles and the bispectrum monopole with the numerical estimate from the same light-cone BOSS mocks. In this case, predictions are calibrated against the mocks in terms of three free parameters accounting for shot-noise contributions and the overall amplitudes of the power spectrum and bispectrum submatrices. This set-up provides a qualitative prediction of the relative size of off-diagonal terms, except for extra contributions due to window function effects not included in the model.

This work presents an accurate model for the theoretical covariance of the galaxy bispectrum and its detailed comparison with estimates from numerical simulations. We explore the regime where non-Gaussian contributions to the covariance are most important, that is the squeezed triangular configurations where the longest-wavelength mode \(k_L\) is at least three times smaller than the other two. For such triangles, it is possible to express all relevant non-Gaussian contributions in terms of power spectrum and bispectrum configurations that can be directly measured in the simulations themselves (or accurately modeled in perturbation theory). This allows a rather accurate prediction of the bispectrum covariance and power spectrum-bispectrum cross-covariance matrices. We highlight three main features: first, the Gaussian approximation fails to order one [63, 72]; second, off-diagonal elements corresponding to triangles sharing a long mode are large (as we can expect); third, there are large cross power spectrum-bispectrum covariance terms where the power spectrum
shares the same long mode as the squeezed bispectrum.

We test our prescription against the QUIJOTE suite of simulations [73] and find that the agreement is at the 20% level for (even mildly) squeezed configurations and within 40% for all configurations. This is much better than the Gaussian approximation that has an error $\sim 100\%$ for squeezed configurations, and completely misses off-diagonal correlations.

This has an important impact on observables that are particularly sensitive to squeezed configurations, such as primordial non-Gaussianity, and in general for methods exploiting consistency relations in large scale structures [82–84]. Given that there are a large number of triangles in the mildly squeezed regime we consider, our findings are potentially relevant for constraints on any cosmological parameter.

We organize this work as follows: in Section 2 we define the power spectrum and bispectrum covariances and find an improved formula to compute non-Gaussian terms of the bispectrum covariance. In Section 3, we estimate the relative importance of non-Gaussian terms over Gaussian ones, determining that the former cannot be neglected for squeezed configurations. We then use response functions to propose a prescription to take these terms into account and find a formula to analytically invert the joint power spectrum-bispectrum covariance. In Section 4 we verify our findings against a large suite of N-body simulations. We show that indeed non-Gaussian terms are large in the regime we predicted and find good agreement of our prescription with the data. We then verify through a $\chi^2$ test that the inverse covariance of our theoretical prediction is reliable. In Section 5 we determine the impact of non-Gaussian terms using a Fisher matrix to determine the constraining power of primordial non-Gaussianity using joint power spectrum-bispectrum measurements. We confirm that non-Gaussian terms cannot be neglected, leading to a degradation in constraints by a factor of $\sim 2$. We finally conclude in Section 6.

2 The power spectrum and bispectrum covariances

In this section we briefly review the theoretical description of the covariance matrix of the power spectrum and the bispectrum (along with their cross-covariance) of a generic, non-Gaussian random field $\delta$. This can represent the matter, galaxy or halo distributions. In section 2.3, we show how some approximations used in the evaluation of the mode-counting factors in the theoretical covariance can lead to large errors, and propose an efficient method to fix such problems.

2.1 Estimators

In our comparison with simulation results we deal with finite-volume effects. Therefore, we start by introducing the Discrete Fourier Transform (DFT) of the density contrast $\delta(x)$ as

$$
\delta(k) \equiv \int_V \frac{d^3x}{(2\pi)^3} e^{-ik \cdot x} \delta(x),
$$

with the inverse given by the series

$$
\delta(x) \equiv \sum_k k^3 \delta(k).
$$

A breaking of adiabaticity, such as that induced by the presence of multiple light fields during inflation, is expected to leave a characteristic signal in the squeezed limit [74–78]. This has the same shape in that limit as the local template with amplitude $f_{\text{NL}}$ [79], which is often used to constrain this effect (see [80, 81] for recent results).
The $N$-point correlator $P_N(k_1, ..., k_N)$ in Fourier space is then generically defined as

$$\langle \delta(k_1) ... \delta(k_N) \rangle \equiv \frac{\delta_K(k_1...N)}{k_f^3} P_N(k_1, ..., k_N),$$

(2.3)

where we adopt the notation $k_1...N \equiv k_1 + ... + k_N$, $k_f = 2\pi/L$ is the fundamental frequency of a cubic box of volume $L^3$ while $\delta_K(k)$ stands for the Kronecker symbol equal to one when the argument vanish, zero otherwise. The cases of $N = 2$ and $3$ correspond to the power spectrum $P(k)$ and the bispectrum $B(k_1, k_2, k_3)$, but the full expression for the bispectrum covariance includes contributions from correlation functions of up to $N = 6$.

An unbiased estimator for the power spectrum of a catalog of particles in a box with periodic boundary conditions can be written as

$$\hat{P}(k) \equiv \frac{k^3}{N_k} \sum_{q \in k} \delta(q) \delta(-q)$$

(2.4)

where the sum runs over all wavenumbers $q$ in the shell of radius $k$, that is such that $k - \Delta k/2 \leq q < k + \Delta k/2$, $\Delta k$ being the radial size of the shell. The normalization factor $N_k$ gives the number of modes in the shell

$$N_k \equiv \sum_{q \in k}$$

(2.5)

and is often approximated in the “thin shell” limit $k \gg \Delta k$ by the integral

$$N_k \simeq \frac{1}{k_f^3} \int_{k - \Delta k/2}^{k + \Delta k/2} dq q^2 d\Omega = 4\pi \frac{k^2 \Delta k}{k_f^3} + O(\Delta k^3).$$

(2.6)

Similarly, an unbiased estimator for the bispectrum can be written as [85]

$$\hat{B}(k_1, k_2, k_3) \equiv \frac{k^3}{N_{tr}(k_1, k_2, k_3)} \sum_{q_1 \in k_1} \sum_{q_2 \in k_2} \sum_{q_3 \in k_3} \delta_K(q_{123}) \delta(q_1) \delta(q_2) \delta(q_3)$$

(2.7)

where the normalization factor $N_{tr}$ gives the number of “fundamental triangles” formed by the vectors $q_i$ satisfying the condition $q_{123} = 0$ that fall in the “triangle bin” defined by the triplet of bin centers $(k_1, k_2, k_3)$ and width $\Delta k$. This is given by

$$N_{tr}(k_1, k_2, k_3) \equiv \sum_{q_1 \in k_1} \sum_{q_2 \in k_2} \sum_{q_3 \in k_3} \delta_K(q_{123}).$$

(2.8)

We discuss how to approximate this sum in section 2.3. Note that, according to this definition, we allow for “open bins”, whose centers do not satisfy the triangle condition themselves (e.g. $k_3 > k_1 + k_2$), but contain fundamental triangles that do $(q_3 < q_1 + q_2$ and permutations) [47].

### 2.2 Power spectrum and bispectrum covariance

The power spectrum covariance is defined in terms of the estimator $\hat{P}(k_i) \equiv \hat{P}_i$ of eq. (2.4) as

$$C_{ij}^P \equiv \langle \hat{P}_i \hat{P}_j \rangle,$$

(2.9)
with \( \delta \hat{P}_i = \hat{P}_i - \langle \hat{P}_i \rangle \) and the indices \( i \) and \( j \) denoting the wavenumbers bins. For realizations of the density field in a box with periodic boundary conditions, it is well known [86] that the covariance matrix is the sum of a Gaussian and a non-Gaussian contributions,

\[
C_{ij}^P = C_{ij}^{P,(PP)} + C_{ij}^{P,(T)}.
\]  

(2.10)

The Gaussian term, depending only on the power spectrum of the distribution, is given by

\[
C_{ij}^{P,(PP)} = 2 \frac{\delta_{ij}^K}{N_{k_i}} \sum_{q \in k_i} P(q)^2 \simeq 2 \frac{\delta_{ij}^K}{N_{k_i}} P(k_i)^2,
\]

(2.11)

where \( \delta_{ij}^K \) is a Kronecker symbol vanishing when the bins \( k_i \) and \( k_j \) do not coincide and where in the second step we assume that \( P(q) \simeq P(k_i) \) for \( q \in k_i \), as expected in the thin-shell approximation. The non-Gaussian term depends instead on the trispectrum of the distribution \( T(q_1, q_2, q_3, q_4) \) and is given by

\[
C_{ij}^{P,(T)} = k_f^3 \bar{T}(k_i, k_j) \equiv k_f^3 \frac{N_{k_i} N_{k_j}}{N_{k_i} N_{k_j}} \sum_{q' \in k_i} \sum_{q'' \in k_j} T(q', -q', q'', -q'').
\]

(2.12)

It is easy to see that the two sums, along with the normalization factors \( N_{k_i} \), provide, in practice, an average of the trispectrum over the angle \( \theta \) between the two vectors \( q' \) and \( q'' \). In fact, for thin shells (\( \Delta k \ll k_i \)) we can approximate in the expression above \( T(q', -q', q'', -q'') \equiv T(q', q'', \theta) \simeq T(k_i, k_j, \theta) \). In addition, the sum is often replaced by an integral, as in eq. (2.6), so that

\[
\bar{T}(k_i, k_j) \simeq \frac{1}{2} \int d\cos \theta T(k_i, -k_i, k_j, -k_j).
\]

(2.13)

Similarly, the bispectrum covariance is defined in terms of the estimator \( \hat{B} \) in eq. (2.7) as

\[
C_{ij}^B \equiv \langle \delta \hat{B}_i \delta \hat{B}_j \rangle,
\]

(2.14)

where the indices \( i \) and \( j \) now denote triplets of wavenumbers, so that \( \hat{B}_i \equiv \hat{B}(k_1^i, k_2^i, k_3^i) \). Again, the full expression can be written in terms of a Gaussian and a non-Gaussian contribution, the latter depending on the density field bispectrum \( B(k_1, k_2, k_3) \), trispectrum \( T(k_1, \ldots, k_4) \) and pentaspectrum \( P_6(k_1, \ldots, k_6) \) [13], that is

\[
C_{ij}^B = C_{ij}^{B,(PPP)} + C_{ij}^{B,(BB)} + C_{ij}^{B,(PT)} + C_{ij}^{B,(P_6)}.
\]

(2.15)

The Gaussian contribution is given, in the thin-shell approximation, by

\[
C_{ij}^{B,(PPP)} \simeq \frac{\delta_{ij} s_B}{k_f^3 N_{tr}^i} P(k_1^i) P(k_2^i) P(k_3^i),
\]

(2.16)

where \( s_B = 6, 2, 1 \) for equilateral, isosceles and scalene triangles, respectively, \( N_{tr}^i \) is the number of fundamental triangles in the triangle bin \( \{ k_1^i, k_2^i, k_3^i \} \), eq. (2.8). Assuming that the correlators are slowly varying in the wavenumber shells, the non-Gaussian terms can be written as

\[
C_{ij}^{B,(BB)} \simeq B_i B_j \left( \Sigma_{ij}^{11} + 8 \ \text{perm.} \right),
\]

(2.17)

\[
C_{ij}^{B,(PT)} \simeq P(k_1^i) \bar{T}(k_2^i, k_3^i) \Sigma_{ij}^{11} + 8 \ \text{perm.},
\]

(2.18)

\[
C_{ij}^{B,(P_6)} \simeq k_f^3 \bar{P}_6(k_1^i, k_2^i, k_3^i),
\]

(2.19)
where we introduce the mode-counting factor

\[ \Sigma_{ij}^{ab} = \frac{1}{N_{tr}^i N_{tr}^j} \sum_{q_i^1 \in k_i^1} \ldots \sum_{q_i^3 \in k_i^3} \sum_{q_j^1 \in k_j^1} \ldots \sum_{q_j^3 \in k_j^3} \delta_K(q_1^{123}) \delta_K(q_2^{123}) \delta_K(q_0^a + q_0^b). \]  

(2.20)

This can be approximated as

\[ \Sigma_{ij}^{ab} \approx \frac{\delta_{k_i^i k_j^j}}{\Delta k} \]  

(e.g. [48]), although the approximation is very inaccurate for some configurations, as we see in section 2.3. Similarly to the case of the non-Gaussian contribution to the power spectrum covariance, eq. (2.12), \( \tilde{T} \) is an angle-averaged trispectrum defined in this case as

\[ \tilde{T}(k_i^2, k_i^3, k_j^2, k_j^3, q) \equiv \frac{1}{\Delta p} \int dp \ T(k_i^2, k_i^3, k_j^2, k_j^3, q, p), \]  

(2.21)

where we have written the trispectrum \( T(k_i^2, -k_i^3, k_j^2, -k_j^3) \) as a function of the sides of the quadrilateral formed by the momenta, and the two diagonals \( q = k_{12} \) and \( p = k_{14} \). The range of integration, of size \( \Delta p \), is over all allowed values of \( p \). Finally, \( \tilde{P}_5 \) is an angle average of the pentaspectrum (see e.g. [48]) that we ignore in our implementation, under the assumption that it is negligible w.r.t. the other contributions (we verify this in Section 3). The full covariance matrix for a data vector including both power spectrum and bispectrum measurements can be written as the block matrix

\[ C = \begin{pmatrix} C^P & C^{PB} \\ C^{BP} & C^B \end{pmatrix}, \]  

(2.22)

that, in addition to the power spectrum and bispectrum covariance matrices \( C^P \) and \( C^B \) described above, includes the cross-covariance between the two statistics \( C^{PB} \) (and its transpose \( C^{BP} \)). These are defined as

\[ C^{PB}_{ij} = C^{BP}_{ji} = \langle \delta \hat{P}_i \delta \hat{B}_j \rangle. \]  

(2.23)

The explicit expression is given by [13]

\[ C^{PB}_{ij} = C^{PB, (PB)}_{ij} + C^{PB, (P_b)}_{ij} \]

\[ \simeq \frac{2}{N_{k_i}} P(k_i) B_j \left( \delta^K_{k_i, k_i^1} + \delta^K_{k_i, k_i^2} + \delta^K_{k_i, k_i^3} \right) + \frac{2}{N_{k_i}} \hat{P}_5(k_i^1, k_i^2, k_i^3, k_i), \]  

(2.24)

where \( \hat{P}_5 \) is an angle-average of the tetraspectrum that we again assume negligible w.r.t. the first contribution, for the reasons that we discuss in Section 3.

### 2.3 Approximation of the mode-counting factors

The theoretical predictions presented above should be evaluated over the Fourier-space grid defined by the discrete modes \( q = n k_f \) for a proper comparison to N-body simulations. This would provide an exact determination of all factors accounting for the number of modes in each \( k \)-shell, or in the intersections of such shells when dealing with one or more triangular bins. In addition, the correlators themselves should, in principle, be calculated on the Fourier grid and the result summed over the shells. This procedure (see e.g. [47, 52]), however, is numerically rather expensive and sums over modes are often replaced by integrals over a continuum of modes.

In our approach, we assume the thin-shell approximation for all the relevant expressions as described in eq.s (2.11), (2.16)-(2.18) and (2.24). Moreover, we employ integrals instead of sums in
the evaluation of the mode-counting factors. We are particularly careful in the definition of such integrals and their integration limits, since a naive approach could lead to significant errors for some subsets of triangular configurations.

As an illustration of the problem, we present here the explicit calculation for the integral that approximates the number of fundamental triangles $N_{tr}$. Consider a fundamental triangle $\{q_1, q_2, q_3\}$ in a given triangular bin $\{k_1, k_2, k_3\}$ for a single measurement of the bispectrum estimator (2.7). From eq. (2.8) we have

$$N_{tr} \equiv \sum_{q_1 \in k_1} \sum_{q_2 \in k_2} \sum_{q_3 \in k_3} \delta_K(q_{123}) \approx \frac{1}{k^6_f} \int_{k_1-\Delta k/2}^{k_1+\Delta k/2} dq_1 \int_{k_2-\Delta k/2}^{k_2+\Delta k/2} dq_2 \int_{k_3-\Delta k/2}^{k_3+\Delta k/2} dq_3 \delta_D(q_{123})$$

$$= \frac{8\pi^2}{k^6_f} \int_{k_1-\Delta k/2}^{k_1+\Delta k/2} dq_1 \int_{k_2-\Delta k/2}^{k_2+\Delta k/2} dq_2 \int_{k_3-\Delta k/2}^{k_3+\Delta k/2} dq_3 \int_{-1}^{1} d\mu \, q_1 q_2 q_3 \delta_D\left(\mu - \frac{q_3^2 - q_1^2 - q_2^2}{2q_1 q_2}\right). \quad (2.25)$$

After taking advantage of the Dirac delta function to get rid of the integral on $\mu$, this leads to the usual result [85, 87]

$$N_{tr} \simeq \frac{8\pi^2}{k^6_f} \int_{k_1-\Delta k/2}^{k_1+\Delta k/2} dq_1 \int_{k_2-\Delta k/2}^{k_2+\Delta k/2} dq_2 \int_{k_3-\Delta k/2}^{k_3+\Delta k/2} dq_3 \, q_1 q_2 q_3 = \frac{8\pi^2}{k^6_f} k_1 k_2 k_3 \Delta k^3. \quad (2.26)$$

In these derivations, we implicitly assumed that all values of $q_1$, $q_2$ and $q_3$ within the integration limits do form a closed triangle corresponding to a value of $\mu$ also within the integration bounds.

Such assumption is not satisfied by all measured configurations. In fact, it breaks (by a large margin) for flattened triangle bins, i.e. those with $k_1 + k_2 = k_3$, or for those configurations where the values of $k_1$, $k_2$ and $k_3$ (the $k$-bin centers) cannot even form a closed triangle. These triangle bins do include closed fundamental triangles in them, that is with $q_{123} = 0$, that we want to keep in our analysis. We refer to these configurations, with some abuse of language, as “open” triangle bins.

For flattened and open triangles, the triangle condition can only be satisfied by the values of $q_1$, $q_2$ and $q_3$ in the regions depicted in Figure 1. The integral can still be performed analytically over those regions, giving

$$N_{tr}^\text{flattened} = \frac{\pi^2}{192k^6_f} \Delta k^3 \left[5\Delta k^3 + 104 \left(k_1^2 + k_2 k_1 + k_2^2\right) \Delta k + 768k_1 k_2 \left(k_1 + k_2\right)\right], \quad (2.27)$$

and

$$N_{tr}^\text{open} = \frac{\pi^2}{1152k^6_f} \Delta k^3 \left[24k_1^2 \left(3\Delta k + 8k_2\right) + \Delta k \left(72k_2 \Delta k + 17\Delta k^2 + 72k_2^2\right)\right]. \quad (2.28)$$

In order to simplify these expressions we used $k_3 = k_1 + k_2$ for flattened triangles, and $k_3 = k_1 + k_2 + \Delta k$ for open triangles.

We plot a comparison between our approximations for $N_{tr}$ and the exact sum in Figure 2. We see that using equation (2.26) for open or flattened configurations gives an error $\gtrsim O(100\%)$. Our approximations for flattened bins in equations (2.28), and the usual approximation of equation (2.26)

\footnote{It also breaks for other configurations, such as isosceles squeezed triangles. But we’ve checked that this gives a small correction in those other cases.}
Figure 1: Regions of a typical bin in the magnitudes of the momenta $q_1$, $q_2$, $q_3$ with a given bin center $k_1$, $k_2$, $k_3$, where the triangle condition is satisfied $q_3 < q_1 + q_2$. Axes are in units of $k_f$. **Left:** A typical “open” bin for which $k_3 = k_1 + k_2 + \Delta k$. **Right:** A typical “flattened” bin for which $k_3 = k_1 + k_2$. Note that in these cases, integrating over the whole cube gives an $O(1)$ error. Here we chose $k_1 = 6k_f$, $k_2 = 12k_f$, and $\Delta k = 3k_f$, but all such configurations give identical plots.

Figure 2: Comparison between analytical approximations and the exact sum for the number of triangles contributing to the bispectrum estimator in a bin. We present the percentage error incurred by using said analytical approximations. Here, our approximation is given by Eqs. (2.27) and (2.28) for “open” and “flattened” bins, respectively. This is usually approximated by Eq. (2.26) instead. All triangle bins with sides between $3k_f$ and $84k_f$ were considered. **Left:** Fractional error for “open” bins. **Right:** Fractional error for “flattened” bins.
for closed bins, give an error $\lesssim 5\%$. Our approximation for open bins, equation (2.27) gives an error $\lesssim 15\%$.

Similar considerations hold for the sums appearing in the expression for the bispectrum covariance. We explicitly computed the sums only for triangles that share the longest wavelength mode $k_1$. We see that they are the most relevant in our case. From eq. (2.20) we have

$$N_{tr}^i N_{tr}^j \Sigma_{11} \simeq \frac{16\pi^3}{k_f^9} \delta^K_{k_1^i k_1^j} \int_{1 - \Delta k/2}^{k_1+\Delta k/2} dq_1 \int_{k_1-\Delta k/2}^{k_1+\Delta k/2} dq_2^a \int_{k_1-\Delta k/2}^{k_1+\Delta k/2} dq_3^b \times \int_{k_2-\Delta k/2}^{k_2+\Delta k/2} dq_2^b \int_{k_3-\Delta k/2}^{k_3+\Delta k/2} dq_3^b \int_{k_2-\Delta k/2}^{k_2+\Delta k/2} dq_3^b \int_{k_3-\Delta k/2}^{k_3+\Delta k/2} dq_3^b \times \int_{-1}^{1} d\mu_a \delta_D \left[ \mu_a - \frac{(q_3^a)^2 - q_1^2 - (q_2^a)^2}{2q_1q_2^a} \right] \times \int_{-1}^{1} d\mu_b \delta_D \left[ \mu_b - \frac{(q_3^b)^2 - q_1^2 - (q_2^b)^2}{2q_1q_2^b} \right].$$

(2.29)

The region of integration is again given by Figure 1, with $q_1$ in the intersection of the two allowed regions. Integrating over the full range is a good approximation for typical bins, which simply gives

$$N_{tr}^i N_{tr}^j \Sigma_{11} = \frac{16\pi^3}{k_f^9} \delta^K_{k_1^i k_1^j} k_2^a k_3^b k_2^b k_3^b \Delta k_5.$$  \hspace{1cm} (2.30)

As in the previous case, using equation (2.30) to approximate sums involving flattened, or open bins gives an error $\gtrsim \mathcal{O}(1)$. The results of the integration over the appropriate regions for flattened and open triangular bins are given in Appendix A.

3 The covariance of squeezed bispectrum configurations

As already mentioned, we focus our attention on squeezed triangular configurations. For such triangles, in fact, we expect the non-Gaussian contribution to their covariance to be dominant. For these configurations, it is possible to obtain an approximate, but satisfactory, model of their covariance matrix from direct measurements of the power spectrum and bispectrum of a given distribution, without the need to fit for any free parameter. This exercise illustrates relevant aspects of the modeling of the bispectrum covariance, which is helpful for its full analytical description.

In this section, we quantify the relevance of the various non-Gaussian contributions to the covariance of the bispectrum in the squeezed limit. We then simplify some of them using response functions [88, 89].

3.1 Order of magnitude estimates

In this section we provide an order of magnitude estimate of the relevance of the non-Gaussian contributions to the covariance matrix w.r.t. the Gaussian one. For simplicity, we assume each higher-order correlator to be described by its tree-level expression in perturbation theory, even at small scales where this approximation breaks down. We validate these results with simulation measurements in the next section.
Figure 3: Dimensionless halo power spectrum for the halos taken from the QUIJOTE simulations, as described in Section 4.1.

**Power Spectrum.** In this case, we have only one non-Gaussian contribution depending on the trispectrum $\bar{T}(k_i, k_j)$. We compare it to the Gaussian diagonal contribution for $k_i = k_j = k$, so that $\bar{T}(k, k) \sim P^3(k)$. We have then

$$\frac{C^{P,(T)}}{C^{P,(PP)}} \sim \frac{k^3 N_k \bar{T}(k, k)}{2 P^2(k)} \sim 2\pi \left( \frac{\Delta k}{k} \right) k^3 P(k) \sim \left( \frac{\Delta k}{k} \right) \Delta^2(k), \quad (3.1)$$

where $\Delta^2(k) \equiv 4\pi k^3 P(k)$ is the dimensionless power spectrum. Note that this estimate holds also for halos taking shot noise into account. Indeed, in the shot noise dominated regime, for Poisson shot noise $\bar{T}(k, k) \sim 1/\bar{n}^3$ and $P(k) \sim 1/\bar{n}$, so that $\bar{T}(k, k) \sim P^3(k)$.

For the halos we consider (see Section 4.1), this is of order $\Delta^2(k) \sim \mathcal{O}(10^{-2})$ at large scales, close to the size of the bin $k \sim \Delta k$. At small scales, e.g. $k/\Delta k \approx 100$, we have instead $\Delta^2(k) \sim \mathcal{O}(10)$ (see Figure 3). Therefore, this contribution is negligible at large scales but only mildly suppressed at intermediate and small scales [86]. On scales that are shot-noise dominated, this ratio grows like $\propto k^2$ such that the non-Gaussian term can become dominant at small enough scales and for halos with low enough number density.

**Bispectrum.** In this case we have several non-Gaussian contributions, Eqs. (2.17), (2.18), (2.19). Let us estimate the size of each with respect to the Gaussian contribution. When all scales are comparable, we set $B(k, k, k) \sim P^2(k)$. This estimate holds also when including shot noise. Indeed, in the shot noise dominated regime for Poisson shot noise $B(k, k, k) \sim 1/\bar{n}^2$ and $P(k) \sim 1/\bar{n}$. Using this in Eq. (2.17) we obtain

$$\frac{C^{B,(BB)}}{C^{B,(PPP)}} \sim k^3 N \Sigma \frac{B^2(k, k, k)}{P^3(k)} \sim \left( \frac{\Delta k}{k} \right)^2 \Delta^2_n(k), \quad (3.2)$$
where $\Sigma_{ij}^{ab} = \delta_{k_i k_j} / N_{k_i}$ is the typical size of one of the mode-counting factors in Eq. (2.17). For the simulations and scales we consider, the contribution from $C^{B,(BB)}$ is expected to be subdominant when all scales are comparable. However, this term grows like $k$ in the shot noise dominated regime, such that it can become dominant at small enough scales.

On the other hand, the bispectrum couples different scales. In the extreme case of squeezed configurations, we write $B(k_L, k_s, k_s) \sim P(k_L)P(k_s)$, where we are assuming that $P(k_L) > P(k_s)$. This estimate still holds. For example, when the short scales are shot-noise dominated the dominant contribution is $B(k_L, k_s, k_s) \sim P(k_L)/\bar{n}$ (see e.g. [43]). Furthermore, the mode-counting factors depend on whether the triangles correlated share the long mode or the short mode. In this way we obtain

$$
\frac{C^{B,(BB)}}{C^{B,(PPP)}} \sim k_j^3 \frac{B^2(k_L, k_s, k_s)}{P(k_L)P^2(k_s)} \left(N_{tr} \Sigma_{ij}^{LL} + 4N_{tr} \Sigma_{ij}^{ss} \right) \sim \left(\frac{\Delta k}{k_L} \right)^2 \Delta_h^2(k_L) \left(\frac{k_s^2}{k_L^2} + 4\right),
$$

(3.3)

The first term in the parenthesis is the one proportional to $\Sigma_{ij}^{11}$ in equation (2.15), which corresponds to a pair of triangles for which the long mode coincides. The second term in the parenthesis corresponds to all other permutations in $C^{B,(BB)}$. Since $\Delta_h^2(k_L) \sim O(10^{-2})$, this contribution to the covariance is generically suppressed. However, there is a quadratic enhancement by the ratio $k_s/k_L$ for the first term, which can easily compensate for this suppression if $k_s/k_L \gtrsim 10$. We thus see that the contribution to $C^{B,(BB)}$ for squeezed configurations that share the long mode can become dominant in the covariance. This generates off-diagonal elements in the covariance of the same order of magnitude as the elements in the diagonal, inducing a large correlation among those triangles.

The trispectrum term is analogous. When all scales are comparable $\tilde{T}(k, k, k, k) \sim P^3(k)$ (which again holds in the presence of shot noise), and we get from Eq. (2.18)

$$
\frac{C^{B,(PT)}}{C^{B,(PPP)}} \sim k_j^3 N_{tr} \Sigma \frac{T(k, k, k, k)P(k)}{P^3(k)} \sim \left(\frac{\Delta k}{k_L} \right)^2 \Delta_h^2(k). \tag{3.4}
$$

For squeezed triangles we instead estimate $\tilde{T}(k, k, k, k, k) \sim P(k_L)P^2(k_s)$. From the explicit expression for the trispectrum in the presence of shot noise (see e.g. [43]), one can check that this estimate still holds. For example, when the short scales are shot-noise dominated $\tilde{T} \sim P(k_L)/\bar{n}^2$. We thus obtain,

$$
\frac{C^{B,(PT)}}{C^{B,(PPP)}} \sim k_j^3 \left(N_{tr} \frac{P(k_L)\tilde{T}(k_s, k_s, k_s, k_s, k_s)}{P(k_L)P^2(k_s)} \Sigma_{ij}^{LL} + 4N_{tr} \frac{P(k_s)\tilde{T}(k_s, k_L, k_s, k_s, k_s)}{P(k_L)P^2(k_s)} \Sigma_{ij}^{ss} \right)
$$

$$
\sim \left(\frac{\Delta k}{k_L} \right)^2 \Delta_h^2(k_L) \left(\frac{k_s^2}{k_L^2} + 4\right). \tag{3.5}
$$

Again, the contribution to $C^{B,(PT)}$ corresponding to triangles that share the long mode can be dominant in the covariance.\footnote{It is worth noticing that in previous literature these non-Gaussian terms were also calculated, either in the context of the position-dependent mass and halo power spectrum [72] or using response functions for the matter bispectrum [63]. Both these analyses identified that these non-Gaussian terms are indeed large, although they did not compare with simulations, and do not discuss non-squeezed triangles.}
For the pentaspectrum contribution, Eq. (2.19), when all scales are comparable, we similarly estimate \( \tilde{P}_6 \sim P^5(k) \) and obtain
\[
\frac{C_{B,(P)}^{B,(P)}}{C_{B,(PPP)}^{B,(PPP)}} \sim \frac{1}{16\pi^2} \left( \frac{k_f}{k} \right)^6 \left( \Delta_h^2(k) \right)^2 , \tag{3.6}
\]
For squeezed configurations we write \( \tilde{P}_6 \sim P^2(k_L)P^3(k_s) \), such that
\[
\frac{C_{B,(P)}^{B,(P)}}{C_{B,(PPP)}^{B,(PPP)}} \sim k_f^6 P(k_s) P(k_L) \sim \frac{1}{16\pi^2} \frac{k_f^3 n^2}{k_L^3 k_s^3} \Delta_h^2(k_s) \Delta_h^2(k_L) . \tag{3.7}
\]
Again, from the explicit expression of \( P_6 \) (e.g. from [43]), these estimates should hold even in the presence of shot-noise. In both cases, even in the shot-noise dominated regime, this contribution is suppressed.

**Bispectrum power spectrum cross covariance.** In order to evaluate if the cross covariance is important, let us estimate the size of the correlation coefficients. As before, we start by considering comparable scales, using Eq. (2.24)
\[
\frac{C_{PB,(PB)}^{PB,(PB)}}{\sqrt{C_{P,(PP)}^{P,(PP)} C_{B,(PPP)}^{B,(PPP)}}} \sim \frac{\sqrt{2} N_{tr}^{1/2} k_f^{3/2}}{N_k^{1/2} k_p^{3/2}} \frac{B(k,k,k)}{P^{3/2}(k)} \sim \frac{\Delta k}{k} (\Delta^2(k))^{1/2} . \tag{3.8}
\]
For the scales and halos we consider, this contribution is subdominant. However, in the shot noise dominated regime coefficient grows slowly as \( \propto k^{1/2} \).

Let us discuss the correlation between a squeezed bispectrum \( B(k_L, k_s, k_s) \) and a power spectrum evaluated at an arbitrary scale \( P(k) \). We get different results depending on whether \( k \) is equal to \( k_s \) or \( k_L \). We estimate
\[
\frac{C_{PB,(PB)}^{PB,(PB)}}{\sqrt{C_{P,(PP)}^{P,(PP)} C_{B,(PPP)}^{B,(PPP)}}} \sim \frac{\sqrt{2} N_{tr}^{1/2} k_f^{3/2}}{N_k^{1/2} k_p^{3/2}} \frac{B(k_L, k_s, k_s)}{P^{1/2}(k_L) P(k_s)} \sim \frac{k_s \Delta k}{k_L} (\Delta^2(k_L))^{1/2} . \tag{3.9}
\]
When \( k = k_s \) this correlation coefficient is always mildly suppressed by the dimensionless power spectrum. On the other hand, when \( k = k_L \) it is enhanced by the squeezing and can easily become \( \mathcal{O}(1) \). Thus, there is a large correlation between squeezed configurations and the power spectrum evaluated at the long mode.

We have assumed that the cross-covariance is dominated by \( C_{PB,(PB)}^{PB,(PB)} \). In order to check this assumption, we compare the two contributions in Eq. (2.24) when all scales are comparable, estimating \( \tilde{P}_5 \sim P^4(k) \). The explicit expressions for \( P_5 \) in the presence of shot noise can be found in [43]. We write
\[
\frac{C_{PB,(P)}^{PB,(P)}}{C_{PB,(PB)}^{PB,(PB)}} \sim \frac{k_f^3 N_k}{2} \frac{R(k,k,k,k,k)}{P(k)B(k,k,k)} \sim \frac{\Delta k}{k} \Delta_h^2(k) , \tag{3.10}
\]
This is similar to the ratio between the non-Gaussian and the Gaussian contributions to the power spectrum covariance. Thus, the tetrascpectrum contribution to the cross-covariance is important at the same scales for which the trispectrum contribution is important for the power spectrum covariance. For the scales and halos we study, this contribution is subdominant. However, in the
shot-noise dominated regime, this ratio scales as $\propto k^2$, with the corresponding correlation coefficient growing strongly as $\propto k^{5/2}$.

For squeezed configurations, we again take the power spectrum to be evaluated at an arbitrary scale $k$ that can either close to $k_L$ or $k_s$. We then estimate $\bar{P}_6(k_L, k_s, k_s, k, k) \sim P(k)P(k_L)P^2(k_s)$, which is valid in both cases. We thus get

$$\frac{C^{PB,(P_b)}}{C^{PB,(PB)}} \sim \frac{k^3 N_k}{2} \frac{\bar{P}_5(k_L, k_s, k_s, k, k)}{P(k)B(k_L, k_s, k_s)} \sim \frac{\Delta k}{k} \Delta^2_h(k),$$  \hspace{1cm} (3.11)

When $k = k_s$, this ratio is the same as before. The corresponding correlation coefficient is subdominant for the halos and scales we consider, but scales like $\propto k_s^2$ in the shot-noise dominated regime. On the other hand, when $k = k_L$, this ratio is always suppressed. Thus, the tetraspectrum can always be ignored for those configurations for which the correlation coefficient is the largest.

### 3.2 The covariance in terms of response functions

We saw in the previous section that the bispectrum covariance is dominated by the Gaussian term $C^{B,(PPP)}$ plus two contributions involving higher-order correlation functions $C^{B,(BB)}$ and $C^{B,(PT)}$. The first can be computed at non-linear scales by using the measured bispectrum or a fitting function [8, 90, 91]. The second contribution is more challenging: trispectrum measurements are difficult (see [32, 42, 92–94]), and we are not aware of fitting formulae for the trispectrum at small scales. However, in the squeezed limit they can both be written in terms of response functions. Specifically, the coupling between short wavelength modes and a long wavelength perturbation of the gravitational potential can be written as

$$\lim_{k_L \to 0} B(k_L, k_1, k_2) = P(k_L) \frac{\partial}{\partial \delta_L} \langle \delta(k_1)\delta(k_2) \rangle'.$$  \hspace{1cm} (3.12)

Here, a prime denotes that the Dirac delta of momentum conservation is dropped, and $\partial/\partial \delta_L$ is the response to a change of the long-wavelength curvature. Such a change also correlates two-point functions at different points. This gives a contribution to the trispectrum in the limit in which the sum of two momenta goes to zero (which we call an internal squeezed limit)\(^6\)

$$\lim_{|k_1 + k_2| \to 0} T(k_1, k_2, k_3, k_4, |k_1 + k_2|, |k_4 + k_4|) \approx P(|k_1 + k_2|) \left( \frac{\partial}{\partial \delta_L} \langle \delta(k_1)\delta(k_2) \rangle' \right) \left( \frac{\partial}{\partial \delta_L} \langle \delta(k_3)\delta(k_4) \rangle' \right).$$  \hspace{1cm} (3.13)

This contribution is dominant when the power spectrum is the largest when evaluated at $k_L = |k_1 + k_2|$ compared to other combinations of the momenta.\(^7\) In our case, we take the long mode to be $k_L \sim 0.01$ Mpc\(^{-1}\) and the short modes to be $k_s \gtrsim 3k_L$. It can be checked that $P(k_L) \gg P(k_s)$, such that we can use this approximation for the trispectrum.

\(^6\)This response function diverges at unequal times, and this divergence is fixed by symmetry [95]. We discuss here the subleading terms in the response, which do not vanish at equal times.

\(^7\)It can be checked in perturbation theory that this contribution can be rendered subdominant in the squeezed limit, with respect to other terms that were discarded, when $P(k_s) \geq P(k_L)$. 

– 14 –
The response functions of short scale power spectra should, in principle, be measured in simulations. However, we can use these expressions to simplify the angle-averaged trispectrum appearing in the bispectrum covariance

\[
\tilde{T}(k_1, k_2, k_3, k_4, q) \approx P(k_L) \left( \frac{\partial}{\partial \delta_L} \langle \delta(k_1) \delta(k_2) \rangle' \right) \left( \frac{\partial}{\partial \delta_L} \langle \delta(k_3) \delta(k_4) \rangle' \right) \frac{1}{\Delta p} \int dp
\]

where we have used the fact that the response functions do not depend on the momentum combination being integrated. Using this, \( C_{ij}^{B, (PT)} \) can be written as

\[
C_{ij}^{B, (PT)} \approx P(k_1^i) \tilde{T}(k_2^i, k_3^i, k_2^j, k_3^j) \Sigma_{ij}^{11}
\]

\[
\approx B(k_1^i, k_2^j, k_3^j) B(k_1^i, k_2^j, k_3^j) \Sigma_{ij}^{11}
\]

\[
\approx C_{ij}^{B, (BB)}.
\]  

(3.14)

3.3 Prescription for the covariance and its inverse

We conclude that the following expression for the bispectrum covariance is a good approximation for squeezed triangles

\[
C_{ij}^{B} \approx \frac{\delta_{ij} s_B}{k_j^i N_{tr}^i} P(k_1^i) P(k_2^j) P(k_3^j) + \frac{2B(k_1^i, k_2^j, k_3^j) B(k_1^i, k_2^j, k_3^j)}{N_{tr}^i N_{tr}^j} \sum_{q \in \{k\}} \delta_{q_1 + q_2 + q_3, q_1 + q_2 + q_3} \delta_{q_1 + q_2 + q_3, q_1 + q_2 + q_3}.
\]

(3.16)

where as before \( s_B = 6, 2, 1 \) for equilateral, isosceles and scalene triangles, respectively, \( N_{tr}^i \) is the number of fundamental triangles in the triangle bin \( \{k_1^i, k_2^i, k_3^i\} \), eq. (2.8), and the sum can be approximated as described in section 2.3. This expression is expected to work reasonably well even for non-squeezed triangles, which are dominated by the Gaussian covariance. In the next section we show that the non-Gaussian terms in this expression are large even for mildly squeezed triangles for which \( k_L \lesssim 3k_s \).

Furthermore, the correlation coefficient between a squeezed triangle and the corresponding long-mode power spectrum is order \( \mathcal{O}(1) \) and dominated by the \( C_{PB}^{BB} \) term.

We derive equation (3.16) under the following assumptions:

- The 6-point function can be ignored. This is reasonable when the long mode is linear. The 6-point function is expected to be important only when all scales involved are non-linear.
- Correlations among triangles sharing the long mode are found to be the largest ones. Other correlations, such as when all scales are short, become important only when all those short scales are deep in the non-linear regime.
- The trispectrum appearing in these correlations can be approximated by the squeezed bispectrum. This is expected to hold as long as the triangles are squeezed and \( P(k_L) > P(k_s) \), as shown in Section 3.2.

\footnote{Note that the analysis in this section holds also in the presence of shot noise.}
For any procedure involving constraints on a parameter, like Fisher matrix and likelihood analyses, we need to invert the covariance matrix. If, as in our case, the covariance matrix has sizable elements far outside the diagonal, the inversion operation can be tricky from a numerical point of view. The off-diagonal terms we introduce have a particular structure that allows for analytic inversion using a generalization of the Sherman-Morrison formula (see [96, 97], also used recently in a scenario similar to ours [72]).

Both from our expectations and from our measurements, we see that the dominant term in the cross-covariance is when the long mode of a squeezed triangle is the same as the momentum of the power spectrum. In that case, the cross-covariance can be written as a sum of outer products of vectors

$$\left( \sum_q b_q a_q^T \right),$$

with the sum running over the values of the momentum $q$. The vectors are

$$a_q = \frac{2}{N_k} \delta_{q,k} P(k), \quad a_q^{(k_1,k_2,k_3)} = 0,$$

$$b_q^k = 0, \quad b_q^{(k_1,k_2,k_3)} = \delta_{k_1,q} B(k_1,k_2,k_3).$$

We labeled the elements of the data vector corresponding to the power spectrum with a single index representing its momentum, and those corresponding to the bispectrum with the three momenta at which it is evaluated.

Our generalization of the Sherman-Morrison formula then gives the following inverse

$$\left( (C^P)^{-1} + \sum_q x \beta_q (C^P)^{-1} a_q a_q^T (C^P)^{-1} - \sum_q x (C^B)^{-1} b_q b_q^T (C^B)^{-1} (C^B)^{-1} + \sum_q x \alpha_q (C^B)^{-1} b_q b_q^T (C^B)^{-1} \right)^{-1},$$

where

$$\alpha_q = a_q^T (C^P)^{-1} a_q, \quad \beta_q = b_q^T (C^B)^{-1} b_q, \quad x = 1/(1 - \alpha_q \beta_q),$$

and we used the diagonal approximation for $C^P$ and the fact that $C^B$ in Eq. (3.16) is block-diagonal. We numerically invert the bispectrum covariance $C^B$, and use this formula for the inverse of the total covariance.

4 Comparison with N-body simulations

In this section we provide a detailed comparison between theoretical models of the power spectrum and bispectrum covariance discussed in the previous section and measurements from numerical simulations. We analyze catalogs of dark matter halos in a cubic box at fixed redshift $z = 0$. Our setup allows for a comparison where the shot-noise contribution is relevant, even if its level is not realistic for typical galaxy redshift surveys. Working in this simplified, but controlled, scenario allows us to verify the accuracy of the model in Sect. 3.3 with little in the way of observational systematic errors.
| Name   | $n_s$ | $h$ | $\Omega_b$ | $\Omega_m$ | $\sigma_8$ | # simulations | $N_p^{1/3}$ | $L_{\text{box}}$ | $V_{\text{tot}}$ | $m_p$ |
|--------|-------|-----|-------------|-------------|------------|---------------|--------------|-----------------|-----------------|------|
| QUIJOTE | 0.9624 | 0.6711 | 0.049 | 0.3175 | 0.834 | 2377 | 512 | 1000 | 2377 | 65.6 |

Table 1: Cosmological and structural parameters for the QUIJOTE simulations.

4.1 N-body simulations

We use the publicly available suite of simulations QUIJOTE\(^9\), run using the GADGET-3 code \cite{98}. We consider a subset of 2377 realizations from the fiducial cosmology (see Table 1), which are run on cubic boxes of 1000 Mpc/$h$ side length with 512\(^3\) particles in them. The initial conditions are set at $z_i = 127$ using the code 2LPTic \cite{99}. The linear transfer function is obtained using the Boltzmann code CAMB \cite{100}. The halo catalogs are generated using a Friends-of-Friends (FoF) algorithm with a linking length $\lambda = 0.2$. We require that halos are constituted by a minimum of 50 particles, implying a number density in each box of $\bar{n} \sim 5 \cdot 10^{-6} (h/\text{Gpc})^3$. In Table 1 we summarize the specifications and the fiducial cosmology for QUIJOTE.

4.2 Measurements and binning strategy

We measure the matter and halo power spectra and bispectra using the estimators Eqs. (2.4) and (2.7). We implement a fourth-order density interpolation and the interlacing scheme described in \cite{101}. Bins have width of $\Delta k = 3k_f$, with $k_f \simeq 0.006 h/\text{Mpc}$. We provide bispectrum measurements based on estimates of the halo number density on two different grids of different sizes, as described and summarized in Table 2:

- From a small grid with a linear size of 256, we measure the power spectrum and all bispectrum configurations up to $k_{\text{max}} = 0.528 h/\text{Mpc}$. We call these measurements Q1.

- From a larger grid of linear size 450 we only measure the bispectrum on squeezed triangles, since measuring all the triangles would have required an exceedingly large amount of memory. The selected triangles are composed of long modes in the range 0.0018 $h$/Mpc to 0.075 $h$/Mpc, while the short modes are between 0.641 $h$/Mpc and 0.942 $h$/Mpc, \(^{10}\) for a total of 4 bins for long modes and 16 bins for short modes. We refer to these measurements as Q2.

Note that we do not subtract shot noise from any of these measurements. This allows us to make use of the theoretical expressions for the covariance in section 3 assuming that all correlators in such expressions include the usual shot-noise contributions. As pointed out in \cite{102}, and extended for the bispectrum case in \cite{24}, if shot noise is subtracted from the measurements then the non-Gaussian contributions are reduced to some extent.

4.3 Results

We estimate the covariance for halo power spectrum $P$, bispectrum $B$ and their cross-covariance directly from the simulations using Eq. (2.23). We notice that measurements of the covariance from

\(^9\)Information on QUIJOTE simulations can be found at https://quijote-simulations.readthedocs.io/en/latest/ and on the reference paper \cite{73}.

\(^{10}\)Note that, for such short modes, the shot noise dominates in our setup, although that might not be the case for realistic galaxy surveys.
Table 2: Binning strategy for Quijote simulations. All bins have width $\Delta k = 3k_f$, where $k_f \simeq 0.006 \, h/\text{Mpc}$. For the Q2 sets of measurements, only squeezed triangles are measured. The triangles we consider are composed of long modes in the range $0.009 \, h/\text{Mpc}$ to $0.075 \, h/\text{Mpc}$, while the short modes are between $0.641 \, h/\text{Mpc}$ and $0.942 \, h/\text{Mpc}$. 

| Name | $N_{\text{grid}}$ | $k_{\text{min}}(h/\text{Mpc})$ | $k_{\text{max}}(h/\text{Mpc})$ | # bins | # triangles |
|------|------------------|------------------|------------------|--------|------------|
| Q1   | 256              | 0.018            | 0.528            | 28     | 2513       |
| Q2   | 450              | 0.018            | 0.942            | 50     | 254        |

a limited number of realizations suffer from correlated numerical noise. Since it is correlated, this noise can be easily confused with additional structure in the matrix. An illustration and a brief discussion of this is found in Appendix B. Our goal is to describe the covariance to an accuracy of roughly $\sim 20\%$. The model of Sect. 3.3 is able to achieve this goal, as we show from the comparison to N-body measurements we present in this section. This leads us to believe that the numerical error is also below this threshold.

The model we are plotting in this section contains the following terms:

- For the power spectrum covariance we only consider the first term in Eq. (2.10), $C_P^{(PP)}$. We expect a small contribution from the term $C_P^{(T)}$ at large scales. At small scales, this term is not negligible, but the power spectrum is not the main focus of this work. For a recent study of the power spectrum covariance, we refer the reader to [66] and references within.

- For the bispectrum covariance we use the result in Eq. (3.16). We call the first term “PPP” and the second one “2BB”, where the factor of 2 accounts for the “PT” term as explained in Section 3.2.

- For the bispectrum-power spectrum cross covariance we take into account only the first term in Eq. 2.24 as we expect the 5-point function to have a small contribution.

For all these predictions, we used the halo power spectrum and bispectrum measured from the simulations themselves. We have also tested results using perturbation theory predictions in place of the measured ones, finding good agreement at large scales. When entering the non-linear regime for the most squeezed triangle, the agreement degrades significantly. An alternative could be using a nonlinear model or fitting function calibrated on simulations [8, 90, 91].

In order to compare the model with measurements, we start by plotting the variance. We then compare the off-diagonal terms of the covariance by plotting the correlation matrix. Finally, we check the inverse covariance by computing $\chi^2$ values using our approximation.

### 4.3.1 Bispectrum variance

In Figures 4 and 5 we show the percentage error between the measurements and the theoretical variance of the bispectrum. We compare our results by considering only the Gaussian approximation “PPP” and the model of Section 3.3, for the Q1 and Q2 grids respectively.

In Figure 4, we present the percentage error of the variance between measurements and the theoretical prediction considering only the “PPP” term (left) and the model (right) for the full set of
triangle configurations in the $Q_1$ grid (containing squeezed and non-squeezed configurations). For triangles that are not squeezed, especially at large scales, the “PPP” term is expected to dominate. For squeezed triangles, non-Gaussian terms become relevant. Indeed, Figure 4 clearly shows that only including the “PPP” contribution leads to errors up to 100%.

Using our model of Eq. (3.16), we recover the measured variance to within a $\lesssim 40\%$ error. We notice that the error is largest at smaller scales (higher triangle index), where non-Gaussian terms beyond the ones we consider are expected to become important according to the estimates of Section 3. We also highlight those triangles that are squeezed, that are described by our approximation within a $20\%$ error.

In Figure 5 we present the percentage error of the variance for the $Q_2$ grid. Since the triangles measured are very squeezed, our approximation works very well. The percentage error is within 20% even for very small non-perturbative scales. The Gaussian approximation fails to approximate the variance by more than $O(1)$. Even if the short scales are shot-noise dominated, the non-Gaussian terms are crucial in computing the variance.

For ease of visualization, we focus on a subset of triangles in Figure 6. We consider two types of triangles: The left panel shows the bispectrum variance and percentage error for squeezed triangles, where we fix the long mode to $q = 0.018 h \text{ Mpc}^{-1}$ and the remaining sides satisfy the condition $k_2 = k_3 = k$. The error here is within the 20% range. We observe that as the triangle becomes squeezed (toward the right part of the Figure), the “2BB” term dominates over the “PPP” term. In the right panel we consider triangles that are not squeezed, with one of the sides fixed at $k_1 = 0.226 h \text{ Mpc}^{-1}$ and the other two satisfying the condition $k_2 = k_3 = k$. In this case the error is somewhat larger than before. The Gaussian “PPP” is a good approximation in all the range considered, and the error is slightly larger than 20% for a few points.
Figure 5: Percentage error for the bispectrum variance for $Q_2$ grid. We compare the variance obtained from N-body simulations with the theoretical variance considering only “PPP” term and the “Model” from (3.16). Note that all triangles for this grid are squeezed. Left: Percentage error between simulation measurements and “PPP” covariance. Right: Percentage error between simulation measurements and “Model” theoretical covariance. As for the previous plot, a Gaussian model for the bispectrum variance, i.e. only including the “PPP” contribution, clearly fails at predicting the variance for squeezed triangles.

In Figure 7, we show similar plots for very squeezed triangles involving highly non-linear scales in the $Q_2$ grid. We see that, as expected, the model works very well, while the Gaussian approximation fails to order $O(1)$. In this Figure, we show “flattened” configurations for which the usual approximation to the mode-counting factor $\Sigma_{ij}$ fails. We are able to recover the measured covariance using the approximation described in section 2.3.

We confirm that our approximations recover the variance measured from simulations within a 20% error for squeezed triangles involving scales up to $k \lesssim 1 \text{ h Mpc}^{-1}$. For these triangles, the non-Gaussian “2BB” term is dominant, and the Gaussian approximation fails. For non-squeezed triangles, we used the larger scales frequently considered in the literature, and find an agreement of 40%. For these configurations, as seen in Figure 6, the non-Gaussian terms we keep are subdominant, and this thus quantifies how good the Gaussian approximation is.

4.3.2 Correlation matrix and off-diagonal terms

It is useful to plot the correlation matrix of the halo power spectrum and bispectrum, defined as

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$

(4.1)

where $C_{ij}$ is the covariance matrix as defined in Eq. (2.22). This allows us to highlight the importance of off-diagonal elements.

In Figure 8 we show the simulation measurements and theoretical cross correlation matrix for the halo bispectrum covariance and the power spectrum bispectrum cross covariance for the $Q_1$ and $Q_2$ grids. The indexes $i_P, j_P, i_B, j_B$ in these figures refer to elements in the data vector. We ordered the bispectra by their smaller momentum first, then its next-to-smallest, and then the largest. In
Figure 6: Bispectrum variance and percentage error between simulation measurements and our theoretical prediction, considering only PPP term (orange) and Model (red). *Left:* Squeezed triangles, where we fix the long mode to $q = 0.018$ h Mpc$^{-1}$ as a function of one of the two remaining sides that satisfy the condition $k_2 = k_3$. *Right:* Generic triangles, where we fix one side to $k = 0.226$ h Mpc$^{-1}$ as a function of one of the two remaining sides that satisfy the condition $k_2 = k_3$. Measurements are performed using the Q1 grid. These plots show that our model of Eq. (3.16) predicts the bispectrum variance to within 20% for any given triangle configuration.

This way, the data vector is such that triangles that share their smallest momentum are situated close to each other. We see large off-diagonal elements both in the simulation measurements and in our model. These correspond to triangles that share a long mode. Perhaps not surprisingly, they show the largest values in the correlation matrix. For the cross-covariance, there are large correlations between squeezed triangles and large-scale power spectra. We can see that our model is in qualitative agreement with the covariance measured on simulations.

In order to better quantify the model accuracy on the off-diagonal elements of the correlation matrix, we plot several rows of $r_{ij}$ in Figure 9. These figures show very squeezed configurations (left panels), a mildly squeezed configuration (upper-right panel), and a non-squeezed configuration (lower-right panel). For squeezed configurations, the simulation measurements agree well with the theoretical model, which captures the dominant structure in the matrix. We notice that even when the triangles do not share one of the momenta, the bispectrum covariance is systematically above or below zero by a small amount. For these configurations, the “PPP”, “BB” and “PT” terms are zero. This correlation can either be due to correlated noise due to the numerical uncertainty in our measurement (see Appendix B) or a manifestation of the 6-point function. We notice that such contributions are, with few exceptions, of order 10% in the correlation matrix. We show a non-squeezed configuration in the lower-right panel. We see that, as expected, there are no large correlations between this triangle and others (they are all below $\sim 20\%$). The additional structure can be due to terms we ignore, such as other $BB$ or $PT$ terms in Eq. 2.19, the 6-point function,
Figure 7: Bispectrum variance and percentage error between N-body simulations and theoretical prediction, considering only PPP term (orange) and Model (red). Left: Very squeezed triangles, where we fix the long mode to $q = 0.018\, h\, \text{Mpc}^{-1}$ as a function of the two remaining sides that satisfy the condition $k_3 = k_2 + q$ (“flattened” triangles). Right: Less squeezed triangles, where we fix the long mode to $q = 0.075\, h\, \text{Mpc}^{-1}$ as a function of the two remaining sides that satisfy the condition $k_3 = k_2 + q$ (“flattened” triangles). Measurements are performed using the Q2 grid. These plots show that our model of Eq. (3.16) predicts the bispectrum variance to within 20% even for very squeezed configurations.

or numerical uncertainty. In the upper-right panel we show a mildly squeezed configuration. We see that it has a large correlation with triangles sharing the long-wavelength mode (towards the left of the plot). We seem to be missing the correlation with the short-wavelength modes, showing at $j_B \sim 500$, but it is $\lesssim 20\%$.

We plot some rows of the power spectrum-bispectrum cross-correlation matrix $r_{ij}^{PB}$ in Figure 10. We see again that our model captures the structure of the cross-correlation. It is large when the power spectrum is evaluated at the long-wavelength mode of the bispectrum. The correlation is systematically different from zero even when the power spectrum does not share a mode with the bispectrum. Since the “PB” contribution is zero for these cases, this can be due to numerical uncertainty (see Appendix B) or the 5-point function.

This comparison corroborates the model described in Section 3.3. It describes the correlation matrix with $\sim 10\%$ accuracy, and its main features are verified:

1. In the bispectrum covariance sub-matrix, the “PPP” term is not the only sizable contribution to the diagonal for squeezed triangles. In fact, it gets corrected by up to $O(1)$ factors because of the “BB” and “PT” terms.

2. Squeezed triangles sharing the same long mode in the bispectrum covariance are highly correlated, such that its off-diagonal component is not negligible.
Figure 8: Top: Cross-correlation matrix $r_{ij}^B$ of the halo bispectrum covariance from simulation measurements (left) and the model (right) in the Q1 grid. Middle: Cross-correlation matrix $r_{ij}^{PB}$ of the halo bispectrum power spectrum cross covariance from simulation measurements (left) and model (right) using Q1 grid. Bottom: simulation measurements (left) and the model (right), using the Q2 grid. For both panels, the upper-left quadrant shows the power spectrum cross-correlation matrix $r_{ij}^{PP}$, the bottom-right quadrant shows the halo bispectrum cross correlation $r_{ij}^B$, and the off-diagonal quadrants show the cross correlation between power spectrum and bispectrum $r_{ij}^{PB}$. Note that colors are saturated where correlations are close to one. This comparison shows qualitatively how our model captures the largest off-diagonal contributions of the cross power-spectrum-bispectrum and bispectrum covariance. Differences between N-body and model in the $r_{ij}^{PP}$ quadrant are related to the fact that we do not model the non-Gaussian covariance of the power spectrum, as the focus of this work is in the bispectrum and cross-bispectrum covariance.
Figure 9: Rows of the bispectrum correlation matrix. In each plot, there is a point with $r = 1$ corresponding to the correlation of a triangle with itself. **Upper-left panel**: very squeezed configuration for the $Q_2$ grid. **Upper-right panel**: squeezed configuration for the $Q_1$ grid. **Bottom-left panel**: Flat squeezed configuration for the $Q_1$ grid. **Bottom-right panel**: Non-squeezed configuration for the $Q_1$ grid. These plots show the agreement of our model of Eq. (3.16) with off-diagonal terms of the bispectrum covariance.

3. Squeezed triangles are very correlated with power spectra sharing the same long mode. As such, the cross-covariance between the power spectrum and the bispectrum is not negligible.

4.4 A $\chi^2$ test for the inverse covariance

We have been comparing the measured covariance with the prescription of Eq. (3.16), in order to prove that triangles are correlated and this correlation is expected from theory. However, what is used in Fisher forecasts and likelihood analyses is actually the precision matrix, i.e. the inverse of the covariance matrix. An accurate numerical inversion of the N-body covariance requires a large number of realizations. With the number of realizations we are using this inverse is very noisy (keeping only squeezed triangles) or even impossible to compute (keeping all triangles).

In order to check the accuracy of the precision matrix obtained from the model of the covariance
Figure 10: Rows of the bispectrum-power spectrum cross covariance correlation matrix. 
Left panel: Very squeezed triangles for the Q2 grid. Right panel: Triangles from the Q1 grid. These plots show the agreement of our model of Eq. (3.16) with off-diagonal terms of the cross bispectrum-power spectrum covariance.

in Section 3.3, we instead compute$^{11}$

$$
\chi^2_{th,i} = (D_i - D_{\text{mean}}) C^{-1}_{th} (D_i - D_{\text{mean}})^T,
$$

where $D_i = (P_i, B_i)$ is the vector of measurements for the $i$-th realization, $D_{\text{mean}}$ is the mean over 2377 realizations and $C_{th}$ is the theoretical covariance, inverted as described in section 3.3. If the theoretical covariance is close to the true covariance, $\chi^2_{th}$ should follow a $\chi^2$ distribution with degrees of freedom equal to the dimension of $D_i$.\(^{12}\)

In Figure 11 we plot this check for the configurations in the Q1, and Q2 grids. We show the histogram of $\chi^2_{th}$ values computed with Eq. (4.2). Different colors correspond to using Eq. (3.16) (Red) or using the Gaussian diagonal covariance (orange), and are compared to the expected $\chi^2$ (black solid line). The panels show the $\chi^2$ values for all triangles, a squeezing factor of 3 and a squeezing factor of 10 for the Q1 grid, and a squeezing factor of 30 for the Q2 grid.

When keeping all triangles, we see that the Gaussian approximation to the covariance fails at reproducing the expected distribution. On the other hand, our prescription produces a distribution that is about 20% off from the expected central $\chi^2$ value. For squeezed triangles, as expected, our distribution fares even better, while the failure of the Gaussian approximation is even more evident. This is a clear indication that correlations among triangles cannot be neglected.

5 Impact on parameter constraints

In the previous section, we showed that one cannot neglect all off-diagonal terms in the bispectrum covariance, nor in the cross power-spectrum-bispectrum covariance. As a consequence, it would be important to assess the impact of these terms on parameter constraints. As a test, in

\(^{11}\)Another check on the reliability of the inverse can be done by performing the so-called half-inverse test, first introduced in [36], which we outline in Appendix C. We thank an anonymous referee for suggesting this test.

\(^{12}\)We are thankful to Anže Slosar for suggesting this check.
Figure 11: Histogram of $\chi^2_{th}$ values using the theoretical covariance as defined in Eq. (4.2). Top Left: All triangles for the $Q_1$ grid. Top Right: Squeezed triangles such that the long mode is at least 3 times smaller than the short modes for the $Q_1$ grid. Bottom Left: Squeezed triangles such that the long mode is at least 10 times smaller than the short modes for the $Q_1$ grid. Bottom Right: Squeezed triangles such that the long mode is at least 30 times smaller than the short modes for the $Q_2$ grid. This test shows that our model is much closer to a $\chi^2$ distribution than considering a Gaussian covariance. As expected, the agreement of our model improves more and more as we restrict to squeezed triangles. For all configurations, the shift of the mean of the model with respect to the the expected $\chi^2$ is around 20%, which is what we find from comparing our model to simulations.
this section we estimate the Fisher information content with and without our corrections on local primordial non-Gaussianity. In this context, constraints on local non-Gaussianity are expected to be affected significantly by our findings, given that the bispectrum information is crucial to break degeneracies [55].

5.1 Local Primordial non-Gaussianity

Interactions among fields taking place during inflation produce a deviation from Gaussianity in the primordial curvature perturbations that seed the formation of structure. Knowing about these interactions is paramount for understanding fundamental physics at very high energies, hence the search for so called primordial non-Gaussianity is one of the major goals in cosmological searches. While primordial perturbations are nowadays constrained to be very close to Gaussian, there is still room for interesting primordial non-Gaussianity to be discovered, and large scale structure surveys are expected to be at the forefront of such an effort [103]. Here, we consider primordial non-Gaussianity of the local type, which is modeled as

\[ \phi(x) = \phi_G(x) + f_{\text{loc}}^{\text{NL}} (\phi_G^2(x) - \langle \phi_G^2 \rangle) + O(\phi_G^3), \]  

(5.1)

being \( \phi \) the curvature perturbation, \( \phi_G \) a Gaussian random field, and \( f_{\text{loc}}^{\text{NL}} \) parametrizes the amplitude of non-Gaussianity. The primordial bispectrum generated by \( \phi \) peaks for squeezed configurations. That is why we expect our findings to be particularly important for this parameter. The modeling of how the primordial signature of \( \phi \) is imprinted in galaxy clustering has been extensively studied in the literature, both for the power spectrum ([104–106], see [107] for a recent review and references therein) and for the bispectrum [55, 65, 72, 108–123].

5.2 N-body simulations with non-Gaussian initial conditions

We are only interested in the qualitative, rather than quantitative, impact of our terms on \( f_{\text{loc}}^{\text{NL}} \) constraints. Thus, we use simulations with and without primordial non-Gaussianity to infer the change in the power spectrum and bispectrum as a function of \( f_{\text{loc}}^{\text{NL}} \), rather than modeling it using perturbation theory. We take advantage of the Eos\(^{13}\) simulations, run using the GADGET-2 [98]. The suite consists of several realizations run with Gaussian and non-Gaussian initial conditions of the local and equilateral type. For our analysis, we use halo catalogs from 10 realizations in the G85L, NG250L and NG250mL sets, which are initialized with local type primordial non-Gaussianity \( f_{\text{loc}}^{\text{NL}} = 0, +250 \) and \( -250 \), respectively. These simulations are following the evolution of 1536\(^3\) dark matter particles in a periodic cubic box with a side length of 2000 Mpc/h. The initial conditions are set at \( z_i = 99 \) by displacing each particle using second order Lagrangian perturbation theory with the publicly available code 2LPTic [99] for Gaussian initial conditions, and its modification 2LPTicNG [125] for non-Gaussian initial conditions. The linear transfer function is obtained using the Boltzmann code CLASS [126]. The halo catalogs are generated using the public code Rockstar [127], identifying candidate halos with a Friends-of-Friends (FoF) algorithm [128] with a linking length \( \lambda = 0.28 \). We require that halos are constituted by a minimum of 50 particles. We compute the power spectrum and bispectrum in an analogous grid to Q1 (see Table 2), where both

\(^{13}\)Information on the Eos suite is available at https://mbiagetti.gitlab.io/cosmos/nbody/eos/. For a recent publication using the same subset, see [124].
\( k_{\text{min}} \) and \( k_{\text{max}} \) are half the size, since the EOS box size is twice the one of QUIJOTE simulations. We call these measurements \( E1 \).

### 5.3 Fisher matrix forecasts

We want to determine the Fisher matrix for the parameter \( f_{\text{loc}}^{\text{NL}} \) using our predictions for the theoretical covariance matrix of the power spectrum and bispectrum. We use the simulations with primordial non-Gaussianity to infer the numerical derivative of the power spectrum and bispectrum with respect to \( f_{\text{loc}}^{\text{NL}} \). The Fisher matrix is defined as

\[
F_{f_{\text{loc}}^{\text{NL}}}^{f_{\text{loc}}^{\text{NL}}} = \frac{\partial D}{\partial f_{\text{loc}}^{\text{NL}}} \cdot C^{-1} \cdot \frac{\partial D^T}{\partial f_{\text{loc}}^{\text{NL}}}
\]

being \( D = (P, B) \) the vector of power spectrum and bispectrum values as a function of \( k \) and \( C \) the full covariance matrix as defined in Equation (2.22).\(^\text{14}\) The derivative is performed numerically as

\[
\frac{\partial D}{\partial f_{\text{loc}}^{\text{NL}}} = \frac{D(NG250L) - D(NGm250L)}{2f_{\text{NL}}},
\]

where \( f_{\text{NL}} = 250 \) and \( D(X) \) is the vector of \( P \) and \( B \) values as measured for the \( X \) simulation using the \( E1 \) grid. We show the derivatives of the power spectrum and bispectrum as a function of wavenumber in Figure 12.

The predicted uncertainty on \( f_{\text{loc}}^{\text{NL}} \) is given by \( \Delta f_{\text{loc}}^{f_{\text{loc}}^{\text{NL}}} = 1/F_{f_{\text{loc}}^{\text{NL}}}^{f_{\text{loc}}^{\text{NL}}}^{1/2} \). We compute \( \Delta f_{\text{loc}}^{f_{\text{loc}}^{\text{NL}}} \) for four different combinations of terms in the covariance matrix

- **Gaussian**: Numerical diagonal covariance matrix
- **Model**: The full covariance matrix.
- **Model Squeezed (S=N)**: The full covariance matrix, where we include only squeezed triangles with a squeezing \( Nk_L < k_S \) for two values of \( N = 3 \) and 10.
- **Model No-Cross**: Covariance matrix setting the P-B cross-covariance to zero, but including off-diagonal terms in the bispectrum covariance

We show the results of the Fisher matrix analysis in Figure 13 and we quote all uncertainties in a Table 3. As expected, including the bispectrum off-diagonal terms has a strong impact on constraints, degrading the constraint on \( f_{\text{loc}}^{\text{NL}} \) by more than a factor of 2 when considering the bispectrum alone and similarly for the joint power spectrum-bispectrum constraints. This result is in agreement with what found in [123] using a perturbative model. The cross power spectrum-bispectrum covariance does not have a sizable impact on the joint constraint, improving it only by \( \sim 10\% \).

\(^\text{14}\)One might be worried that by neglecting non-Gaussian covariance terms in the \( P \) covariance we are underestimating the gain due to the bispectrum. However, for the specific case of local-type non-Gaussianity, most of the constraining power in the power spectrum comes from the large scales, because of the scale dependent bias effect [104–106]. Therefore we do not expect a strong impact in our estimations due to neglecting non-Gaussianity in the power spectrum covariance. A similar discussion applies for the cross power spectrum-bispectrum covariance: we do include terms where the power spectrum is evaluated at large scales, since these are the sizable terms, while the small scale power spectrum terms have a lower impact on the uncertainty.
Figure 12: Numerical derivative of power spectrum and bispectrum. Left: Power spectrum numerical derivative with respect to $f_{\text{NL}}^\text{loc}$ as a function of wavenumber $k$ for the E1 measurements. Error bars indicate standard deviation from the mean of 10 realizations. Right: Bispectrum numerical derivative with respect to $f_{\text{NL}}^\text{loc}$ for all the triangles in the E1 measurements. Error bars indicate standard deviation from the mean of 10 realizations.

Figure 13: Predicted uncertainties (1-$\sigma$) on $f_{\text{NL}}^\text{loc}$. Power spectrum (Orange), bispectrum (Red) and joint power spectrum-bispectrum (Blue) uncertainties for different prescriptions we took for the covariance matrix.
|                | P  | B  | P+B |
|----------------|----|----|-----|
| Gaussian       | 22.8 | 7.6 | 7.2 |
| Model No-Cross | 22.8 | 16.7 | 13.45 |
| Model          | 22.8 | 16.7 | 12.5 |
| Model Squeezed (S=3) | 22.8 | 35.7 | 17.4 |
| Model Squeezed (S=10) | 22.8 | 57.7 | 20.5 |

Table 3: 1-σ uncertainties on $f_{\text{NL}}^{\text{loc}}$ from the four different covariance matrices considered. Each column represents constraints from the power spectrum alone, the bispectrum alone and the joint power spectrum-bispectrum, respectively. The covariance matrix is on the set G85L, which has $f_{\text{NL}}^{\text{loc}} = 0$. The table (and related figure 13) shows that the Gaussian covariance underestimates the uncertainty on $f_{\text{NL}}^{\text{loc}}$ by a factor of $\approx 2$ with respect to our model covariance, eq. (3.16). Interestingly, including the cross bispectrum-power spectrum covariance in the model slightly improves the constraint over a model where this cross-covariance is neglected.

6 Conclusions

We studied the halo bispectrum covariance with a robust numerical estimate obtained from a large set of over 2,000 Quijote simulations. We compared such estimate with an approximate theoretical model using response function arguments. We thus identified the regime where non-Gaussian terms are most important in the bispectrum covariance: squeezed configurations. We pay particular attention to the accurate evaluation of numerical factors related to mode counts in the power spectrum and bispectrum estimators. In addition to the complete measurement of all triangular configurations at large scales, we provide measurements for squeezed configurations where short-wavelength mode extends deep into the nonlinear regime, $k \sim 1 \, h \, \text{Mpc}^{-1}$, for the first time.

We can summarize our findings as follows:

- The Gaussian approximation for the bispectrum variance does not work when squeezed triangular configurations are involved (see e.g. Figures 6 and 7). Non-Gaussian contributions can exceed by an order of magnitude the Gaussian one, as already pointed out by [72] for a different (though related) observable, and [63] for the matter bispectrum case.

- For the same reason, off-diagonal contributions to the bispectrum covariance are particularly large, with correlation coefficients of $O(1)$, for squeezed configurations. Even mildly squeezed triangles (where $k_L \lesssim 3k_s$) are highly correlated when they share the same long mode.

- Off-diagonal contributions to the power spectrum-bispectrum cross-covariance can also be large for squeezed triangles correlated to the power spectrum of the long mode. This leads to correlation coefficients of order one for the full $P + B$ covariance matrix. This was pointed out by [63, 72], who, however, did not compare with simulations.

- We roughly estimate the size of each contribution to the covariance. We then simplify the largest terms using response function arguments. We thus find a simple prescription, Eq. (3.16). This is similar to that adopted by [28, 123], who included the factor of 2 in front
of the $BB$ term, only as a way to estimate the size of the $PT$ term. We also give a simple prescription to invert the full covariance matrix.

- Our implementation of Eq. (3.16), taking advantage of direct measurements of the power spectrum and bispectrum from simulations shows an accuracy better than 20%, not restricted to squeezed triangles. For triangles closer to equilateral, and at small scales, this degrades but does not exceed 40% (see Figure 4).

- It is crucial, in order to obtain an accurate theoretical description of the full $P+B$ covariance, to carefully evaluate the mode-counting factors that enter its expression. We provide an analytic approximation to the exact sums over the grid wave-numbers.

- We explicitly quantify how a fully non-Gaussian prescription for the bispectrum covariance can affect constraints on parameters such as the amplitude of local primordial non-Gaussianity $f_{NL}^{loc}$. This case is particularly interesting, since most of the information on $f_{NL}^{loc}$ is encoded in squeezed configurations [110]. We find that the uncertainty on $f_{NL}^{loc}$ almost doubles with respect to the Gaussian diagonal covariance case. The power spectrum-bispectrum cross-covariance, on the other hand, can reduce it by $\sim 10\%$.

Clearly, an alternative implementation of our theoretical prescription can be given in terms of perturbation theory (e.g. [25]) or some fitting function (e.g. [8, 90, 91]), since it should capture non-linear corrections at small scales. We stress that Eq. (3.16) constitutes a relatively simple extension of the Gaussian variance. It provides a much better approximation to the full covariance that can find application to likelihood and Fisher analyses (e.g. [28, 32, 54, 56]) or to tests of consistency relations [72, 82–84]. These results could be affected by additional non-Gaussian contributions to the covariance, although in the realistic setting of an actual redshift survey the difference might be less relevant than in our ideal case (e.g. [66] for the power spectrum case).

Of course, several improvements can be considered. For instance, it would be straightforward to include shot-noise contribution to the 5- and 6-point correlation functions [49]. We leave this for future work, along with the extension to redshift-space, preferring to highlight here the range of validity of the simplest non-Gaussian model. Furthermore, we can use this covariance to assess the power of very squeezed configurations to constrain $f_{NL}^{loc}$, see e.g. [72, 82].

Acknowledgments

The authors would like to thank Anže Slosar for useful comments. We would also like to thank Francisco Villaescusa-Navarro and the whole QUIJOTE team for making the simulation suite available. M.B acknowledges support from the Netherlands Organization for Scientific Research (NWO), which is funded by the Dutch Ministry of Education, Culture and Science (OCW) under VENI grant 016.Veni.192.210. M.B. also acknowledges support from the NWO project "Cosmic origins from simulated universes" for the computing time allocated to run a subset of the EOS simulations on CARTESIUS, a supercomputer that is part of the Dutch National Computing Facilities. L.C. was supported by ANID scholarship No.21190484 and is supported by "Beca posgrado PUCV, termino de tesis, 2021". J.N. is supported by FONDECYT grant 1211545, “Measuring the Field Spectrum of the Early Universe”. E.S. is partially supported by the INFN INDARK PD51 grant.
A Integrals for open and flat triangles

We list here the coefficients $\Sigma_{ij}^{\alpha \beta}$ appearing in the expression for the covariance (2.15) (for the terms satisfying $k_1 = k_1^j$). In section 2.3 we outlined how they can be computed. In all these expressions we take $k_3 > k_2 > k_1$, and we used $k_3 = k_1 + k_2$ for flattened configurations, and $k_3 = k_1 + k_2 + \Delta k$ for open configurations.

The coefficient for the covariance between two flattened triangles is

$$
N_i^i N_j^j \Sigma_{ij}^{\alpha \beta} = \frac{\pi^3 \Delta k^5}{2903040k_f^5} \left[ 144 \Delta k^2 \left( 867k_1 \left( k_2^j + k_2^i \right) + 867 \left( (k_2^j)^2 + (k_2^i)^2 \right) + 1532k_1^2 \right) \\
+ 1572480k_1 \Delta k \left( k_1 \left( k_2^j + k_2^i \right) + (k_2^j)^2 + (k_2^i)^2 \right) \\
+ 14176512k_2^j \left( k_1 + k_2^j \right) k_2^i \left( k_1 + k_2^i \right) + 6347\Delta k^4 \right]. \quad (A.1)
$$

The coefficient for the covariance between two open triangles is

$$
N_i^i N_j^j \Sigma_{ij}^{\alpha \beta} = \frac{\pi^3 \Delta k^5}{10080k_f^5} \left[ \frac{3}{32} \Delta k^8 \left( 632 \left( k_2^j + k_2^i \right) + 457k_1^2 \right) \\
+ \frac{1}{4} \Delta k^7 \left( 601k_1 \left( k_2^j + k_2^i \right) + 3 \left( 336k_2^j k_2^i + 79(k_2^j)^2 + 79(k_2^i)^2 \right) + 132k_1^2 \right) \\
+ 7\Delta k^6 \left( 13k_1^2 \left( k_2^j + k_2^i \right) + k_1 \left( 72k_2^j k_2^i + 13(k_2^j)^2 + 13(k_2^i)^2 \right) + 36k_2^j k_2^i \left( k_2^j + k_2^i \right) \right) \\
+ 252k_2^j \left( k_1 + k_2^j \right) k_2^i \Delta k^5 \left( k_1 + k_2^i \right) + \frac{8027\Delta k^6}{576} \right]. \quad (A.2)
$$

The coefficient for the covariance between a flattened triangle (i) and an open triangle (j) is

$$
N_i^i N_j^j \Sigma_{ij}^{\alpha \beta} = \frac{\pi^3 \Delta k^5}{5806080k_f^5} \left[ 54 \Delta k^3 \left( 663k_1 + 488k_2^j \right) \\
+ 144\Delta k^2 \left( 939k_1 k_2^j + 288k_2^i + 2549k_2^j k_1 + 2549(k_2^j)^2 + 183(k_2^i)^2 \right) \\
+ 4032\Delta k \left( 27k_1 (k_2^j)^2 + 3 \left( 9k_1^2 + 128k_2^i k_1 + 128(k_2^i)^2 \right) k_2^j + 143k_1 k_2^j \left( k_1 + k_2^j \right) \right) \\
+ 1548288k_2^j \left( k_1 + k_2^j \right) k_2^i \left( k_1 + k_2^i \right) + 6253\Delta k^4 \right]. \quad (A.3)
$$

The coefficient for the covariance between a closed triangle (i) and an open triangle (j) is

$$
N_i^i N_j^j \Sigma_{ij}^{\alpha \beta} = \frac{1}{240k_f^5} \pi^3 k_1^3 k_2^3 \Delta k^5 \left( 10\Delta k \left( 3k_1 + 8k_2^j \right) + 80k_2^j \left( k_1 + k_2^j \right) + 19\Delta k^2 \right). \quad (A.4)
$$

The coefficient for the covariance between a closed triangle (i) and a flattened triangle (j) is

$$
N_i^i N_j^j \Sigma_{ij}^{\alpha \beta} = \frac{1}{12k_f^5} \pi^3 k_1^3 k_2^3 \Delta k^5 \left( 96k_2^j \left( k_1 + k_2^j \right) + 13k_1 \Delta k \right). \quad (A.5)
$$
**B Cross-correlation matrix with correlated noise**

Measurements of the covariance matrix need a large number of simulations. Using few simulations makes the measurement noisy and unstable, with noise that is *highly correlated*. We illustrate this in Figure 14. We show the correlation matrix measured from simulations using an increasingly large number of realizations. Note that for a few realizations there are large regions far from the diagonal where the matrix seems to be systematically negative (red) or systematically positive (blue) rather than randomly distributed around zero (as expected from uncorrelated noise).\(^{15}\) This structure gradually disappears as the number of realizations increases.

---

**Figure 14:** Model cross correlation matrix for Q2 set, as in Figure 8, considering different numbers of realizations. Upper left: 100 realizations. Upper right: 300 realizations. Bottom left: 500 realizations. Bottom right: 1000 realizations.

\(^{15}\)One could also check whether the noise is Gaussian by performing a Kolmogorov-Smirnov test. We deserve a more thorough check of the noise for future projects.
C Half-inverse test

As a further check of the inverse, we can compute the relation \( H = C_{th}^{-1/2} C_{N-body} C_{th}^{-1/2} - I \), where \( C_{th} \) is our model covariance and \( C_{N-body} \) is the covariance computed from the simulations [36]. If the model and covariance computed from the simulations were identical, this relation should be zero. However, we expect that our model is accurate within \( \approx 20\% \) of the covariance made out of a very large number of realizations. Following [129], we can check whether the noise scales as \( 1/\sqrt{N_{\text{sims}}} \). Since \( H \) is symmetric by construction, we can plot only the lower half-triangle of \( H \) and compare it to a noise scaling exactly as \( 1/\sqrt{N_{\text{sims}}} \), which we put on the upper half-triangle of the same matrix. We show this check for the Q2 set of simulations for \( N_{\text{sims}} = 100, 250, 500 \) and 1000 realizations in Figure 15\textsuperscript{16}. It shows that increasing the number of simulations, (and therefore lowering the noise which scales as \( 1/\sqrt{N_{\text{sims}}} \)), an additional structure emerges, which is particularly evident on the diagonal (lower right panel). There is also further structure on off-diagonal cells parallel to the diagonal. We reserve a better study of these features to future projects.

Figure 15: Half-inverse test for the Q2 set, considering different numbers of realizations. Upper left: 100 realizations. upper right: 250 realizations. Bottom left: 500 realizations. Bottom right: 1000 realizations.

\textsuperscript{16}Note that, for ease of visualization, in this case the triangle index is growing from bottom to top in the y-axis, differently than previous plots.
References

[1] EUCLID Collaboration, R. Laureijs et al., “Euclid Definition Study Report,” arXiv:1110.3193 [astro-ph.CO].

[2] DESI Collaboration, A. Aghamousa et al., “The DESI Experiment Part I: Science, Targeting, and Survey Design,” arXiv:1611.00036 [astro-ph.IM].

[3] LSST Dark Energy Science Collaboration, A. Abate et al., “Large Synoptic Survey Telescope: Dark Energy Science Collaboration,” arXiv:1211.0310 [astro-ph.CO].

[4] O. Doré et al., “Cosmology with the SPHEREX All-Sky Spectral Survey,” arXiv:1412.4872 [astro-ph.CO].

[5] J. N. Fry, “Gravity, bias, and the galaxy three-point correlation function,” Phys. Rev. Lett. 73 (Jul, 1994) 215–219. https://link.aps.org/doi/10.1103/PhysRevLett.73.215.

[6] S. Matarrese, L. Verde, and A. F. Heavens, “Large scale bias in the universe: Bispectrum method,” Mon. Not. Roy. Astron. Soc. 290 (1997) 651–662, arXiv:astro-ph/9706059.

[7] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, “The Bispectrum of IRAS redshift catalogs,” Astrophys. J. 546 (2001) 652, arXiv:astro-ph/0004087.

[8] R. Scoccimarro and H. M. P. Couchman, “A fitting formula for the nonlinear evolution of the bispectrum,” Mon. Not. Roy. Astron. Soc. 325 (2001) 1312, arXiv:astro-ph/0009427 [astro-ph].

[9] M. Takada and B. Jain, “The Three - point correlation function in cosmology,” Mon. Not. Roy. Astron. Soc. 340 (2003) 580–608, arXiv:astro-ph/0209167.

[10] I. Szapudi, “Three - point statistics from a new perspective,” Astrophys. J. Lett. 605 (2004) L89, arXiv:astro-ph/0404476.

[11] E. Gaztanaga, P. Norberg, C. M. Baugh, and D. J. Croton, “Statistical analysis of galaxy surveys. 2. The 3-point galaxy correlation function measured from the 2dFGRS,” Mon. Not. Roy. Astron. Soc. 364 (2005) 620–634, arXiv:astro-ph/0506249.

[12] R. C. Nichol et al., “The effect of large-scale structure on the sdss galaxy three-point correlation function,” Mon. Not. Roy. Astron. Soc. 368 (2006) 1507–1514, arXiv:astro-ph/0602548.

[13] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, “Cosmology and the Bispectrum,” Phys. Rev. D 74 (2006) 023522, arXiv:astro-ph/0604505.

[14] E. Sefusatti and E. Komatsu, “The bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and non-linear galaxy bias,” Phys. Rev. D76 (2007) 083004, arXiv:0705.0343 [astro-ph].

[15] R. E. Smith, R. K. Sheth, and R. Scoccimarro, “An analytic model for the bispectrum of galaxies in redshift space,” Phys. Rev. D 78 (2008) 023523, arXiv:0712.0017 [astro-ph].

[16] M. Liguori, E. Sefusatti, J. R. Fergusson, and E. P. S. Shellard, “Primordial non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure,” Adv. Astron. 2010 (2010) 980523, arXiv:1001.4707 [astro-ph.CO].

[17] J. E. Pollack, R. E. Smith, and C. Porciani, “Modelling large-scale halo bias using the bispectrum,” Mon. Not. Roy. Astron. Soc. 420 (2012) 3469, arXiv:1109.3458 [astro-ph.CO].

[18] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, “Renormalized Halo Bias,” JCAP 1408 (2014) 056, arXiv:1402.5916 [astro-ph.CO].
[19] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer, “The Bispectrum in the Effective Field Theory of Large Scale Structure,” *JCAP* 05 (2015) 007, arXiv:1406.4135 [astro-ph.CO].

[20] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore, “The One-Loop Matter Bispectrum in the Effective Field Theory of Large Scale Structures,” *JCAP* 10 (2015) 039, arXiv:1406.4143 [astro-ph.CO].

[21] N. McCullagh, D. Jeong, and A. S. Szalay, “Toward accurate modelling of the non-linear matter bispectrum: standard perturbation theory and transients from initial conditions,” *Mon. Not. Roy. Astron. Soc.* 455 no. 3, (2016) 2945–2958, arXiv:1507.07824 [astro-ph.CO].

[22] A. Lazanu, T. Giannantonio, M. Schmittfull, and E. P. S. Shellard, “The matter bispectrum of large-scale structure: three-dimensional comparison between theoretical models and numerical simulations,” arXiv:1510.04075 [astro-ph.CO].

[23] R. de Belsunce and L. Senatore, “Tree-Level Bispectrum in the Effective Field Theory of Large-Scale Structure extended to Massive Neutrinos,” *JCAP* 02 (2019) 038, arXiv:1804.06849 [astro-ph.CO].

[24] 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 no. 1, (2019) 364–384, arXiv:1803.02132 [astro-ph.CO].

[25] A. Eggemeier, R. Scoccimarro, and R. E. Smith, “Bias Loop Corrections to the Galaxy Bispectrum,” *Phys. Rev. D* 99 no. 12, (2019) 123514, arXiv:1812.03208 [astro-ph.CO].

[26] S. Floerchinger, M. Garny, A. Katsis, N. Tetradis, and U. A. Wiedemann, “The dark matter bispectrum from effective viscosity and one-particle irreducible vertices,” *JCAP* 09 (2019) 047, arXiv:1907.10729 [astro-ph.CO].

[27] T. Steele and T. Baldauf, “Precise Calibration of the One-Loop Bispectrum in the Effective Field Theory of Large Scale Structure,” *Phys. Rev. D* 103 no. 2, (2021) 023520, arXiv:2009.01200 [astro-ph.CO].

[28] D. Alkhanishvili, C. Porciani, E. Sefusatti, M. Biagetti, A. Lazanu, A. Oddo, and V. Yankelevich, “The reach of next-to-leading-order perturbation theory for the matter bispectrum,” arXiv:2107.08054 [astro-ph.CO].

[29] M. Schmittfull, T. Baldauf, and U. Seljak, “Near optimal bispectrum estimators for large-scale structure,” *Phys. Rev. D* 91 no. 4, (2015) 043530, arXiv:1411.6595 [astro-ph.CO].

[30] 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 no. 1, (2018) 1341–1376, arXiv:1801.09280 [astro-ph.CO].

[31] V. Yankelevich and C. Porciani, “Cosmological information in the redshift-space bispectrum,” *Mon. Not. Roy. Astron. Soc.* 483 no. 2, (2019) 2078–2099, arXiv:1807.07076 [astro-ph.CO].

[32] D. Gualdi, H. Gil-Marin, and L. Verde, “Joint analysis of anisotropic power spectrum, bispectrum and trispectrum: application to N-body simulations,” arXiv:2104.03976 [astro-ph.CO].

[33] L. Verde et al., “The 2dF Galaxy Redshift Survey: The Bias of galaxies and the density of the Universe,” *Mon. Not. Roy. Astron. Soc.* 335 (2002) 432, arXiv:astro-ph/0112161.

[34] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, “The power spectrum and bispectrum of SDSS DR11 BOSS galaxies – I. Bias and gravity,” *Mon. Not. Roy. Astron. Soc.* 451 no. 1, (2015) 539–580, arXiv:1407.5668 [astro-ph.CO].
[35] H. Gil-Marín, L. Verde, J. Noreña, A. J. Cuesta, L. Samushia, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, “The power spectrum and bispectrum of SDSS DR11 BOSS galaxies – II. Cosmological interpretation,” *Mon. Not. Roy. Astron. Soc.* 452 no. 2, (2015) 1914–1921, arXiv:1408.0027 [astro-ph.CO].

[36] Z. Slepian *et al.*, “The large-scale 3-point correlation function of the SDSS BOSS DR12 CMASS galaxies,” arXiv:1512.02231 [astro-ph.CO].

[37] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres, and M. D. Olmstead, “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies,” *Mon. Not. Roy. Astron. Soc.* 465 no. 2, (2017) 1757–1788, arXiv:1606.00439 [astro-ph.CO].

[38] Z. Slepian *et al.*, “Detection of baryon acoustic oscillation features in the large-scale three-point correlation function of SDSS BOSS DR12 CMASS galaxies,” *Mon. Not. Roy. Astron. Soc.* 469 no. 2, (2017) 1738–1751, arXiv:1607.06097 [astro-ph.CO].

[39] P. Monaco, “Approximate methods for the generation of dark matter halo catalogs in the age of precision cosmology,” *Galaxies* 4 no. 4, (2016) 53, arXiv:1605.07752 [astro-ph.CO].

[40] M. Colavincenzo *et al.*, “Comparing approximate methods for mock catalogues and covariance matrices – III: bispectrum,” *Mon. Not. Roy. Astron. Soc.* 482 no. 4, (2019) 4883–4905, arXiv:1806.09499 [astro-ph.CO].

[41] R. Scoccimarro, “The bispectrum: from theory to observations,” *Astrophys. J.* 544 (2000) 597, arXiv:astro-ph/0004086.

[42] E. Sefusatti and R. Scoccimarro, “Galaxy bias and halo-occupation numbers from large-scale clustering,” *Phys. Rev. D* 71 (2005) 063001, arXiv:astro-ph/0412626.

[43] K. C. Chan and L. Blot, “Assessment of the Information Content of the Power Spectrum and Bispectrum,” *Phys. Rev. D* 96 no. 2, (2017) 023528, arXiv:1610.06585 [astro-ph.CO].

[44] J. Byun, A. Eggemeier, D. Regan, D. Seery, and R. E. Smith, “Towards optimal cosmological parameter recovery from compressed bispectrum statistics,” *Mon. Not. Roy. Astron. Soc.* 471 no. 2, (2017) 1581–1618, arXiv:1705.04392 [astro-ph.CO].

[45] K. C. Chan, A. Moradinezhad Dizgah, and J. Noreña, “Bispectrum Supersample Covariance,” *Phys. Rev. D* 97 no. 4, (2018) 043532, arXiv:1709.02473 [astro-ph.CO].

[46] C. Hahn, F. Villaescusa-Navarro, E. Castorina, and R. Scoccimarro, “Constraining $M_\nu$ with the bispectrum. Part I. Breaking parameter degeneracies,” *JCAP* 03 (2020) 040, arXiv:1909.11107 [astro-ph.CO].

[47] A. Oddo, E. Sefusatti, C. Porciani, P. Monaco, and A. G. Sánchez, “Toward a robust inference method for the galaxy bispectrum: likelihood function and model selection,” *JCAP* 03 (2020) 056, arXiv:1908.01774 [astro-ph.CO].

[48] D. Gualdi and L. Verde, “Galaxy redshift-space bispectrum: the Importance of Being Anisotropic,” *JCAP* 06 (2020) 041, arXiv:2003.12075 [astro-ph.CO].

[49] N. S. Sugiyama, S. Saito, F. Beutler, and H.-J. Seo, “Perturbation theory approach to predict the covariance matrices of the galaxy power spectrum and bispectrum in redshift space,” *Mon. Not. Roy. Astron. Soc.* 497 no. 2, (2020) 1684–1711, arXiv:1908.06234 [astro-ph.CO].

[50] J. Byun, A. Oddo, C. Porciani, and E. Sefusatti, “Towards cosmological constraints from the
compressed modal bispectrum: a robust comparison of real-space bispectrum estimators,” *JCAP* **03** (2021) 105, [arXiv:2010.09579](https://arxiv.org/abs/2010.09579) [astro-ph.CO].

[51] C. Hahn and F. Villaescusa-Navarro, “Constraining $M_\nu$ with the bispectrum. Part II. The information content of the galaxy bispectrum monopole,” *JCAP* **04** (2021) 029, [arXiv:2012.02200](https://arxiv.org/abs/2012.02200) [astro-ph.CO].

[52] A. Oddo, F. Rizzo, E. Sefusatti, C. Porciani, and P. Monaco, “Cosmological parameters from the likelihood analysis of the galaxy power spectrum and bispectrum in real space,” [arXiv:2108.03204](https://arxiv.org/abs/2108.03204) [astro-ph.CO].

[53] F.-S. Kitaura *et al.,* “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: mock galaxy catalogues for the BOSS Final Data Release,” *Mon. Not. Roy. Astron. Soc.* **456** no. 4, (2016) 4156–4173, [arXiv:1509.06400](https://arxiv.org/abs/1509.06400) [astro-ph.CO].

[54] A. Eggemeier, R. Scoccimarro, R. E. Smith, M. Crocce, A. Pezzotta, and A. G. Sánchez, “Testing one-loop galaxy bias: Joint analysis of power spectrum and bispectrum,” *Phys. Rev. D* **103** no. 12, (2021) 123550, [arXiv:2102.06902](https://arxiv.org/abs/2102.06902) [astro-ph.CO].

[55] 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](https://arxiv.org/abs/2010.14523) [astro-ph.CO].

[56] M. M. Ivanov, O. H. E. Philcox, T. Nishimichi, M. Simonović, M. Takada, and M. Zaldarriaga, “Precision analysis of the redshift-space galaxy bispectrum,” [arXiv:2110.10161](https://arxiv.org/abs/2110.10161) [astro-ph.CO].

[57] D. Gualdi, M. Manera, B. Joachimi, and O. Lahav, “Maximal compression of the redshift space galaxy power spectrum and bispectrum,” *Mon. Not. Roy. Astron. Soc.* **476** no. 3, (2018) 4045–4070, [arXiv:1709.03600](https://arxiv.org/abs/1709.03600) [astro-ph.CO].

[58] H. L. Child, Z. Slepian, and M. Takada, “A Physical Picture of Bispectrum Baryon Acoustic Oscillations in the Interferometric Basis,” [arXiv:1811.12396](https://arxiv.org/abs/1811.12396) [astro-ph.CO].

[59] H. L. Child, M. Takada, T. Nishimichi, T. Sunayama, Z. Slepian, S. Habib, and K. Heitmann, “Bispectrum as Baryon Acoustic Oscillation Interferometer,” *Phys. Rev. D* **98** no. 12, (2018) 123521, [arXiv:1806.11147](https://arxiv.org/abs/1806.11147) [astro-ph.CO].

[60] D. Gualdi, H. Gil-Marín, R. L. Schulmann, M. Manera, B. Joachimi, and O. Lahav, “Enhancing BOSS bispectrum cosmological constraints with maximal compression,” *Mon. Not. Roy. Astron. Soc.* **484** no. 3, (2019) 3713–3730, [arXiv:1806.02853](https://arxiv.org/abs/1806.02853) [astro-ph.CO].

[61] D. Gualdi, H. Gil-Marín, M. Manera, B. Joachimi, and O. Lahav, “GEOMAX: beyond linear compression for three-point galaxy clustering statistics,” *Mon. Not. Roy. Astron. Soc.* **497** no. 1, (2020) 776–792, [arXiv:1912.01011](https://arxiv.org/abs/1912.01011) [astro-ph.CO].

[62] D. Gualdi, H. Gil-Marín, M. Manera, B. Joachimi, and O. Lahav, “Geometrical compression: a new method to enhance the BOSS galaxy bispectrum monopole constraints,” *Mon. Not. Roy. Astron. Soc.* **484** no. 1, (2019) L29–L34, [arXiv:1901.00987](https://arxiv.org/abs/1901.00987) [astro-ph.CO].

[63] A. Barreira, “The squeezed matter bispectrum covariance with responses,” *JCAP* **03** (2019) 008, [arXiv:1901.01243](https://arxiv.org/abs/1901.01243) [astro-ph.CO].

[64] M. Shirasaki, N. S. Sugiyama, R. Takahashi, and F.-S. Kitaura, “Constraining primordial non-Gaussianity with postreconstructed galaxy bispectrum in redshift space,” *Phys. Rev. D* **103** no. 2, (2021) 023506, [arXiv:2010.04567](https://arxiv.org/abs/2010.04567) [astro-ph.CO].

[65] A. Barreira, “Predictions for local PNG bias in the galaxy power spectrum and bispectrum and the consequences for $f_{NL}$ constraints,” [arXiv:2107.06887](https://arxiv.org/abs/2107.06887) [astro-ph.CO].
[66] D. Wadekar and R. Scoccimarro, “Galaxy power spectrum multipoles covariance in perturbation theory,” *Phys. Rev. D* **102** no. 12, (2020) 123517, arXiv:1910.02914 [astro-ph.CO].

[67] E. Sefusatti, C. Vale, K. Kadota, and J. Frieman, “Primordial non-Gaussianity and Dark Energy constraints from Cluster Surveys,” *Astrophys. J.* **658** (2007) 669–679, arXiv:astro-ph/0609124 [astro-ph].

[68] I. Kayo, M. Takada, and B. Jain, “Information content of weak lensing power spectrum and bispectrum: including the non-Gaussian error covariance matrix,” *Mon. Not. Roy. Astron. Soc.* **429** (2013) 344–371, arXiv:1207.6322 [astro-ph.CO].

[69] I. Kayo and M. Takada, “Cosmological parameters from weak lensing power spectrum and bispectrum tomography: including the non-Gaussian errors,” arXiv:1306.4684 [astro-ph.CO].

[70] M. Takada and W. Hu, “Power Spectrum Super-Sample Covariance,” *Phys. Rev. D* **87** no. 12, (2013) 123504, arXiv:1302.6994 [astro-ph.CO].

[71] A. Barreira and F. Schmidt, “Responses in Large-Scale Structure,” *JCAP* **06** (2017) 053, arXiv:1703.09212 [astro-ph.CO].

[72] R. de Putter, “Primordial physics from large-scale structure beyond the power spectrum,” arXiv:1802.06762 [astro-ph.CO].

[73] F. Villaescusa-Navarro *et al.*, “The Quijote simulations,” *Astrophys. J. Suppl.* **250** no. 1, (2020) 2, arXiv:1909.05273 [astro-ph.CO].

[74] E. Komatsu and D. N. Spergel, “Acoustic signatures in the primary microwave background bispectrum,” *Phys. Rev. D* **63** (2001) 063002, arXiv:astro-ph/0005036.

[75] Planck Collaboration, Y. Akrami *et al.*, “Planck 2018 results. IX. Constraints on primordial non-Gaussianity,” *Astron. Astrophys.* **641** (2020) A9, arXiv:1905.05697 [astro-ph.CO].

[76] M. Marinucci, T. Nishimichi, and M. Pietroni, “Measuring Bias via the Consistency Relations of the Large Scale Structure,” *Phys. Rev. D* **100** no. 12, (2019) 123537, arXiv:1907.09866 [astro-ph.CO].
[84] M. Marinucci, T. Nishimichi, and M. Pietroni, “Model independent measurement of the growth rate from the consistency relations of the LSS,” *JCAP* **07** (2020) 054, arXiv:2005.09574 [astro-ph.CO].

[85] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, “Nonlinear evolution of the bispectrum of cosmological perturbations,” *Astrophys. J.* **496** (1998) 586, arXiv:astro-ph/9704075.

[86] R. Scoccimarro, M. Zaldarriaga, and L. Hui, “Power spectrum correlations induced by nonlinear clustering,” *Astrophys. J.* **527** (1999) 1, arXiv:astro-ph/9901099.

[87] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga, “Probing primordial non-Gaussianity with large-scale structure,” *Phys. Rev.* **D69** (2004) 103513, arXiv:astro-ph/0312286 [astro-ph].

[88] C.-T. Chiang, *Position-dependent power spectrum: a new observable in the large-scale structure*. PhD thesis, Munich U., 2015. arXiv:1508.03256 [astro-ph.CO].

[89] A. Barreira, E. Krause, and F. Schmidt, “Complete super-sample lensing covariance in the response approach,” *JCAP* **06** (2018) 015, arXiv:1711.07467 [astro-ph.CO].

[90] H. Gil-Marín, C. Wagner, F. Fragkoudi, R. Jimenez, and L. Verde, “An improved fitting formula for the dark matter bispectrum,” *JCAP* **02** (2012) 047, arXiv:1111.4477 [astro-ph.CO].

[91] R. Takahashi, T. Nishimichi, T. Namikawa, A. Taruya, I. Kayo, K. Osato, Y. Kobayashi, and M. Shirasaki, “Fitting the nonlinear matter bispectrum by the Halofit approach,” *Astrophys. J.* **895** no. 2, (2020) 113, arXiv:1911.07886 [astro-ph.CO].

[92] D. Bertolini, K. Schutz, M. P. Solon, and K. M. Zurek, “The Trispectrum in the Effective Field Theory of Large Scale Structure,” *JCAP* **06** (2016) 052, arXiv:1604.01770 [astro-ph.CO].

[93] D. Gualdi, S. Novell, H. Gil-Marín, and L. Verde, “Matter trispectrum: theoretical modelling and comparison to N-body simulations,” *JCAP* **01** (2021) 015, arXiv:2009.02290 [astro-ph.CO].

[94] T. Steele and T. Baldauf, “Precise Calibration of the One-Loop Trispectrum in the Effective Field Theory of Large Scale Structure,” *Phys. Rev. D* **103** no. 10, (2021) 103518, arXiv:2101.10289 [astro-ph.CO].

[95] P. Creminelli, J. Noreña, M. Simonović, and F. Vernizzi, “Single-Field Consistency Relations of Large Scale Structure,” *JCAP* **1312** (2013) 025, arXiv:1309.3557 [astro-ph.CO].

[96] J. Sherman and W. J. Morrison, “Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix,” *The Annals of Mathematical Statistics* **21** no. 1, (1950) 124 – 127. https://doi.org/10.1214/aoms/1177729893.

[97] M. S. Bartlett, “An Inverse Matrix Adjustment Arising in Discriminant Analysis,” *The Annals of Mathematical Statistics* **22** no. 1, (1951) 107 – 111. https://doi.org/10.1214/aoms/1177729698.

[98] V. Springel, “The Cosmological simulation code GADGET-2,” *Mon. Not. Roy. Astron. Soc.* **364** (2005) 1105–1134, arXiv:astro-ph/0505010 [astro-ph].

[99] M. Crocce, S. Pueblas, and R. Scoccimarro, “Transients from Initial Conditions in Cosmological Simulations,” *Mon. Not. Roy. Astron. Soc.* **373** (2006) 369–381, arXiv:astro-ph/0606505 [astro-ph].

[100] A. Lewis, A. Challinor, and A. Lasenby, “Efficient computation of CMB anisotropies in closed FRW models,” *Astrophys. J.* **538** (2000) 473–476, arXiv:astro-ph/9911177 [astro-ph].
[101] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. Couchman, “Accurate Estimators of Correlation Functions in Fourier Space,” *Mon. Not. Roy. Astron. Soc.* **460** no. 4, (2016) 3624–3636, arXiv:1512.07295 [astro-ph.CO].

[102] R. E. Smith, “Covariance of cross-correlations: towards efficient measures for large-scale structure,” *Mon. Not. Roy. Astron. Soc.* **400** (2009) 851, arXiv:0810.1960 [astro-ph].

[103] P. D. Meerburg et al., “Primordial Non-Gaussianity,” arXiv:1903.04409 [astro-ph.CO].

[104] 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.* **D77** (2008) 123514, arXiv:0710.4560 [astro-ph].

[105] S. Matarrese and L. Verde, “The effect of primordial non-Gaussianity on halo bias,” *Astrophys. J.* **677** (2008) L77–L80, arXiv:0801.4826 [astro-ph].

[106] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, “Constraints on local primordial non-Gaussianity from large scale structure,” *JCAP* **0808** (2008) 031, arXiv:0805.3580 [astro-ph].

[107] M. Biagetti, “The Hunt for Primordial Interactions in the Large Scale Structures of the Universe,” *Galaxies* **7** no. 3, (2019) 71, arXiv:1906.12244 [astro-ph.CO].

[108] T. Baldauf, U. Seljak, and L. Senatore, “Primordial non-Gaussianity in the Bispectrum of the Halo Density Field,” *JCAP* **1104** (2011) 006, arXiv:1011.1513 [astro-ph.CO].

[109] E. Sefusatti, M. Crocce, and V. Desjacques, “The Matter Bispectrum in N-body Simulations with non-Gaussian Initial Conditions,” *Mon. Not. Roy. Astron. Soc.* **406** (2010) 1014–1028, arXiv:1003.0007 [astro-ph.CO].

[110] E. Sefusatti, M. Crocce, and V. Desjacques, “The Halo Bispectrum in N-body Simulations with non-Gaussian Initial Conditions,” *Mon. Not. Roy. Astron. Soc.* **425** (2012) 2903, arXiv:1111.6966 [astro-ph.CO].

[111] S. Yokoyama, T. Matsubara, and A. Taruya, “Halo/galaxy bispectrum with primordial non-Gaussianity from integrated perturbation theory,” *Phys. Rev.* **D89** no. 4, (2014) 043524, arXiv:1310.4925 [astro-ph.CO].

[112] G. Tasinato, M. Tellarini, A. J. Ross, and D. Wands, “Primordial non-Gaussianity in the bispectra of large-scale structure,” *JCAP* **1403** (2014) 032, arXiv:1310.7482 [astro-ph.CO].

[113] A. M. Dizgah, K. C. Chan, J. Noreña, M. Biagetti, and V. Desjacques, “Squeezing the halo bispectrum: a test of bias models,” arXiv:1512.06084 [astro-ph.CO].

[114] I. Hashimoto, A. Taruya, T. Matsubara, T. Namikawa, and S. Yokoyama, “Constraining higher-order parameters for primordial non-Gaussianities from power spectra and bispectra of imaging surveys,” *Phys. Rev. D* **93** no. 10, (2016) 103537, arXiv:1512.08352 [astro-ph.CO].

[115] 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 [astro-ph.CO].

[116] I. Hashimoto, S. Mizuno, and S. Yokoyama, “Constraining equilateral-type primordial non-Gaussianities from imaging surveys,” *Phys. Rev. D* **94** no. 4, (2016) 043532, arXiv:1605.07348 [astro-ph.CO].

[117] 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** no. 6, (2017) 063530, arXiv:1611.03590 [astro-ph.CO].
[118] E. Di Dio, H. Perrier, R. Durrer, G. Marozzi, A. Moradinezhad Dizgah, J. Noreña, and A. Riotto, “Non-Gaussianities due to Relativistic Corrections to the Observed Galaxy Bispectrum,” *JCAP* **03** (2017) 006, arXiv:1611.03720 [astro-ph.CO].

[119] C.-T. Chiang, A. M. Cieplak, F. Schmidt, and A. Slosar, “Response approach to the squeezed-limit bispectrum: application to the correlation of quasar and Lyman-α forest power spectrum,” *JCAP* **06** (2017) 022, arXiv:1701.03375 [astro-ph.CO].

[120] H. An, M. McAneny, A. K. Ridgway, and M. B. Wise, “Non-Gaussian Enhancements of Galactic Halo Correlations in Quasi-Single Field Inflation,” *Phys. Rev.* **D97** no. 12, (2018) 123528, arXiv:1711.02667 [hep-ph].

[121] A. Moradinezhad Dizgah, H. Lee, J. B. Muñoz, and C. Dvorkin, “Galaxy Bispectrum from Massive Spinning Particles,” *JCAP* **05** (2018) 013, arXiv:1801.07265 [astro-ph.CO].

[122] A. Moradinezhad Dizgah, G. Franciolini, A. Kehagias, and A. Riotto, “Constraints on long-lived, higher-spin particles from galaxy bispectrum,” *Phys. Rev. D* **98** no. 6, (2018) 063520, arXiv:1805.10247 [astro-ph.CO].

[123] A. Barreira, “On the impact of galaxy bias uncertainties on primordial non-Gaussianity constraints,” *JCAP* **12** (2020) 031, arXiv:2009.06622 [astro-ph.CO].

[124] M. Biagetti, A. Cole, and G. Shiu, “The Persistence of Large Scale Structures I: Primordial non-Gaussianity,” *JCAP* **04** (2021) 061, arXiv:2009.04819 [astro-ph.CO].

[125] 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 [astro-ph.CO].

[126] D. Blas, J. Lesgourgues, and T. Tram, “The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes,” *JCAP* **1107** (2011) 034, arXiv:1104.2933 [astro-ph.CO].

[127] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, “The Rockstar Phase-Space Temporal Halo Finder and the Velocity Offsets of Cluster Cores,” *Astrophys. J.* **762** (2013) 109, arXiv:1110.4372 [astro-ph.CO].

[128] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, “The Evolution of Large Scale Structure in a Universe Dominated by Cold Dark Matter,” *Astrophys. J.* **292** (1985) 371–394.

[129] J. Hou, Z. Slepian, and R. N. Cahn, “Measurement of Parity-Odd Modes in the Large-Scale 4-Point Correlation Function of SDSS BOSS DR12 CMASS and LOWZ Galaxies,” arXiv:2206.03625 [astro-ph.CO].