Non-Gaussianity from large-scale structure surveys

Licia Verde

ICREA (Institució Catalana de Recerca i Estudis Avançat) &
ICC-UB (Instituto de Ciencias del Cosmos, Universitat de Barcelona,
Marti i Franques 1, 08028 Barcelona, ES

Paper submitted to the special issue “Testing the Gaussianity and Statistical Isotropy of the Universe” of Advances in Astronomy

Abstract.
With the advent of galaxy surveys which provide large samples of galaxies or galaxy clusters over a volume comparable to the horizon size (SDSS-III, HETDEX, Euclid, JDEM, LSST, Pan-STARRS, CIP etc.) or mass-selected large cluster samples over a large fraction of the extra-galactic sky (Planck, SPT, ACT, CMBPol, B-Pol), it is timely to investigate what constraints these surveys can impose on primordial non-Gaussianity. I illustrate here three different approaches: higher-order correlations of the three dimensional galaxy distribution, abundance of rare objects (extrema of the density distribution), and the large-scale clustering of halos (peaks of the density distribution). Each of these avenues has its own advantages, but, more importantly, these approaches are highly complementary under many respects.

1 email: liciaverde@icc.ub.edu
1 Introduction

The recent advances in the understanding of the origin and evolution of the Universe have been driven by the advent of high-quality data, in unprecedented amount (just think of WMAP and SDSS for example). Despite this, most of the information about cosmological parameters come from the analysis of a massive compression of the data: the power spectrum of their statistical fluctuations over the mean. The power spectrum is a complete statistical description of a random field only if it is Gaussian.

Even the simplest inflationary models predict deviations from Gaussian initial conditions. These deviations are expected to be small, although “small” in some models may be “detectable”. For a thorough review of inflationary non-Gaussianity see [11], for our purpose it will be sufficient to say that to describe inflation-motivated departures from gaussian initial conditions many write ([2-5] (see [1] and references therein). On the other hand, many inflationary models have an equilateral-type non-Gaussianity, i.e. the bispectrum is dominated by equilateral triangles [6]. Ref. [7,8] have proposed a functional form which closely approximates the behavior of the inflationary bispectrum and which is useful for efficient data-analysis:

\[
\frac{\Phi}{\phi} = \frac{f_{NL}}{f_{NL}} \left( \phi^2 - \langle \phi^2 \rangle \right). \tag{1}
\]

Here \( \phi \) denotes a gaussian field and \( \Phi \) denotes Bardeen’s gauge-invariant potential, which, on sub-Hubble scales reduces to the usual Newtonian peculiar gravitational potential, up to a minus sign. In the literature, there are two conventions for Eq. (1): the large-scale structure (LSS) and the Cosmic Microwave Background (CMB) one. In the LSS convention \( \Phi \) is linearly extrapolated at \( z = 0 \); in the CMB convention \( \Phi \) is instead primordial: thus \( f_{NL}^{LSS} = \frac{g(z=\infty)}{g(0)} f_{NL}^{CMB} \sim 1.3 f_{NL}^{CMB} \), where \( g(z) \) denotes the linear growth suppression factor relative to an Einstein-de-Sitter Universes. In the past few years it has become customary to always report \( f_{NL}^{CMB} \) values even if, for simplicity as it will be clear below, one carries out the calculations with \( f_{NL}^{LSS} \).

While for simplicity one may just assume \( f_{NL} \) in eq. 1 to be a constant (yielding the so-called local model or local-type) in reality the expression is more complicated and \( f_{NL} \) is scale and configuration dependent. In general, the non-Gaussianity is specified by writing down the bispectrum of \( \Phi \). For example, one can see that for the local model the bispectrum is:

\[
B_{\Phi}(k_1,k_2,k_3) = 2 f_{NL} P_\phi(k_1) P_\phi(k_2) + 2 \text{cyc.} \tag{2}
\]

Where \( P \) denotes the power spectrum and it is often assumed that \( P_\phi = P_\phi \); “cyc.” denotes two cyclic terms over \( k_1,k_2,k_3 \).

It has been shown [6] that for non-Gaussianity of the local type the bispectrum is dominated by the so-called squeezed configurations, triangles where one wavevector length is much smaller than the other two. Models such as the curvaton for example have a non-Gaussianity of the local type. Standard, single-field slow roll inflation also yields a local non-Gaussianity but with an unmeasurably small \( f_{NL} \) (see [1] and references therein). On the other hand, many inflationary models have an equilateral-type non Gaussianity, i.e. the bispectrum is dominated by equilateral triangles [6]. Ref. [7,8] have proposed a functional form which closely approximates the behavior of the inflationary bispectrum and which is useful for efficient data-analysis:

\[
B(k_1,k_2,k_3) = 6 f_{NL} \left[ -P(k_1)P(k_2) + 2 \text{cyc.} \right] - 2 [P(k_1)P(k_2)P(k_3)]^{2/3} + P^{1/3}(k_1)P^{2/3}(k_2)P(k_3) + 5 \text{cyc.}. \tag{3}
\]

Note that the same numerical value for \( f_{NL} \) gives rise to a larger skewness in the local case than in the equilateral case (e.g. see [2]), explaining why the CMB constraints on \( f_{NL} \) are weaker for the equilateral case [7,8,10].

Specific deviations from a single field, slow roll, canonical kinetic energy, Bunch-Davies vacuum, leave their specific signature on the bispectrum “shape” (i.e. the dependence of \( B \) on the shape of the triangle made by the three \( \vec{k} \) vectors), see discussion in [11] and references therein.

Non-Gaussianity therefore offers a probe of aspects of inflation (namely the interactions
of the inflaton) that are difficult to probe by other means (i.e., measuring the shape of the primordial power spectrum and properties of the stochastic background of gravity waves). So, how could primordial non-Gaussianity be tested?

One could look at the early Universe: by looking at CMB anisotropies we can probe cosmic fluctuations at a time when their statistical distribution should have been close to their original form but the signal is small. On the other hand, one could analyze the statistics of the large-scale structures, close to the present day, when the signal is larger, but this is a more complicated approach, since gravitational instability (for the dark matter distribution) and bias (for galaxies or clusters of galaxies) introduce non-Gaussian features in an initially Gaussian field and they mask the signal one is after. Finally, the abundance of rare events (such as galaxy clusters and high-redshift galaxies) probes the tails of the PDF of the density field, which are extremely sensitive to deviations from Gaussianity. Here, I will concentrate on the signature of non-Gaussianity on large-scale structure (i.e. at redshift \( z < 1 \)) as they can be traced by galaxy surveys (i.e., I will not consider wide field weak gravitational lensing surveys); other contributions to this review will focus on non-Gaussian signatures on the CMB, thus here it will be sufficient to give only a very brief and succinct introduction to the subject.

At recombination the density fluctuations are small (\( \delta_\Phi \sim 10^{-5} \)), and the CMB temperature fluctuations are directly related to \( \Phi \) making this a very clean probe. However, effectively only one redshift can be tested giving us only a 2-dimensional information.

The most widespread technique for testing Gaussianity in the CMB is to use the CMB bispectrum:

\[
\langle a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} \rangle = B_{l_1 l_2 l_3} (\ell_1 \ell_2 \ell_3)
\]

(4)

where the \( a_{l_n}^{m_n} \) are the coefficients of the spherical harmonic expansion of the CMB temperature fluctuation: \( \Delta T / T = \sum_{\ell m} a_{\ell m} Y_{\ell m} \) and the presence of the 3-J symbol ensures that the bispectrum is defined if \( \ell_1 + \ell_2 + \ell_3 = \text{even} \), \( \ell_j + \ell_k \geq l_1 \geq |l_j - \ell_k| \) (triangle rule) and that \( m_1 + m_2 + m_3 = 0 \). It should be however clear that secondary CMB anisotropies and foregrounds also induce a CMB bispectrum which can mask or partially mimic the signal see e.g., Refs. [12-5, 13, 14, 15, 16, 17] and references therein.

In the last few years, this area of research has received an impulse, motivated by the recent full sky CMB data from WMAP. In particular it has been shown that the constraints can be greatly improved by effectively “reconstructing” the potential \( \Phi \) from CMB temperature and polarization data rather than simply using the temperature bispectrum alone [18, 19]. This technique would yield constraints on non-Gaussianity of the local type of \( \Delta f_{NL} \sim 1 \) for an ideal experiment and \( \Delta f_{NL} \sim 3 \) for the Planck satellite. This is particularly promising as \( f_{NL} \) of order unity or larger is produced by broad classes of inflationary models (see e.g., [1] and references therein).

Currently, the most stringent constraints for the local type are \( 27 < f_{NL} < 147 \) at the 95% confidence (central value 87) from WMAP 3 yr data [20]; and from the WMAP 5 years data, \( -9 < f_{NL} < 111 \) at the 95% confidence level (central value 55) [10] and \( -4 < f_{NL} < 80 \) [21]. Despite the heated debate on whether \( f_{NL} = 0 \) is ruled out or not, the two measurement are not necessarily in conflict: the two central values differ by only about 1σ; different, although not independent, data sets were used with different galactic cuts, and the maximum multipole considered in the analyses is also different. What makes the subject very interesting, is that, if the central value for \( f_{NL} \) is truly around 60, forthcoming data will yield a highly-significant detection.
2 Higher-order correlations

Theoretical considerations (see discussion in e.g., [11] and references therein) lead us to define primordial non-Gaussianity by its bispectrum. While in principle there may be types of non-Gaussianity which would be more directly related to higher-order correlations (e.g., [22] and references therein), and while a full description of a non-Gaussian distribution would require the specification of all the higher-order correlations, it is clear that quantities such as the bispectrum enclose information about the phase correlation between k-modes. In the Gaussian case, different Fourier mode are uncorrelated (by definition of Gaussian random phases) and a statistic like the power spectrum does not carry information about phases. The bispectrum is the lowest-order correlation with zero expectation value in a Gaussian random field. But, even if the initial conditions were Gaussian, non-linear evolution due to gravitational instability generates a non-zero bispectrum. In particular, gravitational instability has its own “signature” bispectrum, at least in the next-to-leading order in cosmological perturbation theory [23]:

\[ B(\vec{k}_1, \vec{k}_2, \vec{k}_3) = 2P(k_1)P(k_2)J(\vec{k}_1, \vec{k}_2) + 2\text{cyc}. \tag{5} \]

where \( J(\vec{k}_1, \vec{k}_2) \) is the gravitational instability “kernel” which depends very weakly on cosmology and for an Einstein-de-Sitter Universe is:

\[ J(\vec{k}_1, \vec{k}_2) = \frac{5}{7} + \frac{\vec{k}_1 \cdot \vec{k}_2}{2k_1k_2} \left( \frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \left( \frac{\vec{k}_1 \cdot \vec{k}_2}{k_1k_2} \right)^2. \tag{6} \]

In the highly non-linear regime the detailed form of the kernel changes, but it is something that could be computed and calibrated by extending perturbation theory beyond the next-to-leading order and by comparing with numerical N-body simulations (see other contributions in this issue). It was recognized a decade ago [4] that this signal is quite large compared to any expected primordial non-Gaussianity and that the primordial signal “redshifts away” compared to the gravitational signal. In fact, a primordial signal given by a local type of non-Gaussianity parameterized by a given \( f_{NL} \), would affect the late-time dark matter density bispectrum with a contribution of the form:

\[ B_{NL, \text{local}}(\vec{k}_1, \vec{k}_2, \vec{k}_3, z) = 2f_{NL}P(k_1)P(k_2) \frac{\mathcal{F}(\vec{k}_1, \vec{k}_2)}{D(z)/D(z = 0)} + 2\text{cyc}. \tag{7} \]

where \( D(z) \) is the linear growth function which in an Einstein-de Sitter universe goes like \((1 + z)^{-1}\) and

\[ \mathcal{F} = \frac{M(k_3)}{M(k_1)M(k_2)}; \quad M(k) = \frac{2}{3} k^2 T(k), \tag{8} \]

\( T(k) \) denoting the transfer function, \( H_0 \) the Hubble parameter and \( \Omega_m0 \) the matter density parameter. Clearly the two contributions have different scale and redshift dependence and the two kernel shapes in configuration space are different, thus, making the two component, at least in principle and for high signal-to-noise, separable.

Unfortunately, with galaxy surveys, one does not observe the dark matter distribution directly. Dark matter halos are believed to be hosts for galaxy formation, and different galaxies at different redshifts populate halos following different prescriptions. In large-scale structure studies, often the assumption of linear, scale independent bias is made. A linear bias will not introduce a non-zero bispectrum in a Gaussian field and its effect on a field with a non-zero bispectrum is only to rescale its bispectrum amplitude. This is, however, an approximation, possibly roughly valid at large scales for dark matter halos, and when looking at the power spectrum, but unlikely to be true in detail. To go beyond the linear bias assumption, often the assumption of quadratic bias is made, where the relation between dark matter overdensity field and galaxy
field is specified by two parameters: \( b_1 \) and \( b_2 \): \( \delta_\text{\textbf{f}}(x) = b_1\delta_\text{DM}(x) + b_2(\delta_\text{DM}^2 - \langle \delta_\text{DM}^2 \rangle) \); \( b_1 \) and \( b_2 \) are assumed to be scale-independent (although this assumption must break down at some point) but they can vary with redshift. Clearly, a quadratic bias will introduce non-Gaussianity even on an initially Gaussian field. In summary, for local non-Gaussianity and scale-independent quadratic bias we have \[24\]:

\[
B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, z) = 2P(k_1)P(k_2)b_1(z)^3 \times 
\left[ f_{\text{NL}} \frac{\mathcal{F}(\mathbf{k}_1, \mathbf{k}_2)}{D(z)} + J(\mathbf{k}_1, \mathbf{k}_2) \frac{b_2(z)}{2b_1(z)} \right] + \text{cyc.}
\]

Before the above expression can be compared to observations it needs to be further complicated by redshift space distortions (and shot noise). Realistic surveys use the redshift as a proxy for distance, but gravitationally-induced peculiar velocities distort the redshift-space galaxy distribution. We will not go into these details here as including redshift space distortions (and shot noise) will not change the gist of the message.

From a practical point of view, it is important to note that photometric surveys, although in general can cover larger volumes than spectroscopic ones, are not suited for this analysis: the projection effects due to the photo-z smearing along the line-of-sight is expected to suppress significantly the sensitivity of the measured bispectrum to the shape of the primordial one (see e.g., \[25\] and references therein; other contributions to this issue); the non-linear bias signature is not uncontroversial, and there have been so far no detection of any extra (primordial) bispectrum contributions.

Of course one could also consider higher-order correlations. One of the advantages of considering e.g., the trispectrum is that, contrary to the bispectrum, it has very weak non-linear growth \[29\], but has the disadvantage that the signal is de-localized: the number of possible configurations grows fast with the dimensionality \( n \) of the \( n \)-point function!

In summary, higher-order correlations as observed in the CMB or in the evolved Universe, can be used to determine the bispectrum shape. The two approaches should be seen as complementary as they are affected by different systematic effects and probe different scales. The next two probes we consider have a less rich sensitivity to the bispectrum shape, but their own peculiar advantages.

### 3 The Mass Function

The abundance of collapsed objects (dark matter halos as traced e.g., by galaxies and galaxy clusters) contains important information about the properties of initial conditions on galaxy and clusters scales. The Gaussian assumption plays a central role in analytical predictions for the abundance and statistical properties of the first objects to collapse in the Universe. In this context, the formalism proposed by Press & Schechter \[30\], with its later extensions and improvements, has become the standard lore for predicting the number of collapsed dark matter halos as a function of redshift. However, even a small deviation from Gaussianity can have a deep impact on those statistics which probe the tails of the distribution. This is indeed the case...
for the abundance of high-redshift objects like galaxies and clusters at \( z \gtrsim 1 \) which correspond to high peaks, i.e. rare events, in the underlying dark matter density field. Therefore, even small deviations from Gaussianity might be potentially detectable by looking at the statistics of high-redshift systems. Before proceeding let us introduce some definitions.

We are interested in predictions for rare objects, that is the collapsed objects that form in extreme peaks of the density field \( \delta(x) = \delta \rho / \rho \). The statistics of collapsed objects can be described by the statistics of the density perturbation smoothed on some length scale \( R \) (or equivalently a mass scale \( M = 4/3 \pi R^3 \rho \), \( \delta_R \)).

To incorporate non-Gaussian initial conditions into predictions for the smoothed density field, we need an expression for the probability distribution function (PDF) for \( \delta_R \). For a particular real-space expansion like Eq. (1), one may make a formal change of variable in the Gaussian PDF to generate a normalized distribution [31]. However, this may not be possible in general and the change of variables does not work for the smoothed cumulants of the density field. In general, the PDF for a generic non-Gaussian distribution can be written exactly as a function of the cumulant’s generating function \( W_R \) for the smoothed density field:

\[
\mathcal{P}(\delta_R) d\delta_R = \int \frac{d\lambda}{2\pi} \exp[-i\lambda \delta_R + W_R(\lambda)] d\delta_R
\]

with

\[
W_R(\lambda) = \sum_{n=2}^{\infty} \frac{(i\lambda)^n}{n!} \mu_{n,R}
\]

where \( \mu_{n,R} \) denote the cumulants and for example \( \mu_{2,R} = \sigma_R^2 = \langle \delta_R^2 \rangle \) and the skewness \( \mu_{3,R} \) is related to the normalized skewness of the smoothed density field \( S_{3,R} = \mu_{3,R}/\mu_{2,R}^2 \). It is useful to define a “skewness per f_{NL} unit” \( S_{3,R}^{f_{NL}=1} \) so that \( S_{3,R} = f_{NL} S_{3,R}^{f_{NL}=1} \). The skewness

\[
\mu_{3,R} = \int \frac{d^3k_1 d^3k_2 d^3k_3}{(2\pi)^7} B_{\delta R}(\vec{k}_1, \vec{k}_2, \vec{k}_3) \delta^D(\vec{k}_1 + \vec{k}_2 + \vec{k}_3)
\]

where \( \delta^D \) denotes the Dirac delta function and \( B_{\delta R} \) denotes the bispectrum of the \( \delta \) overdensity field smoothed on scale \( R \). It is related to the potential one trivially by remembering the Poisson equation: \( \delta_R(\vec{k}) = M(k) W_R(k) \Phi(\vec{k}) \). Here, \( W_R(k) \) denotes the smoothing kernel: usually taken to be the Fourier transform of the top hat window. In any practical application the dimensionality of the integration can be reduced by collapsing the expressing \( \vec{k}_3 \) as a function of \( \vec{k}_1 \) and \( \vec{k}_2 \).

It is important at this point to make a small digression to specify definitions of key quantities. Even in linear theory, the normalized skewness of the density field depends on redshift; however in the Press & Schechter framework one should always use linearly extrapolated quantities at \( z = 0 \). In this context therefore, when writing \( S_{3,R} = f_{NL} S_{3,R}^{f_{NL}=1} \), if \( S_{3,R}^{f_{NL}=1} \) is that if the density field extrapolated linearly at \( z = 0 \) then \( f_{NL} \) must be the LSS one and not the CMB one.

To compute the abundance of collapsed objects fro the PDF one will then follow the Press
& Schechter swindle: first compute

\[ P(\delta_c | z, R) = \int_{\delta_c(z)}^{\infty} d\delta_c P(\delta_c) \] (13)

where \( \delta_c \) denotes the critical threshold for collapse; then the number of collapsed objects is:

\[ n(M, z) dM = \frac{3H_0^2 \Omega_m}{8\pi G M^2} \left| \frac{dP(> \delta_c | z, R)}{dM} \right|. \] (14)

Note that the redshift dependence is usually enclosed only in \( \delta_c : \sigma_M \) is computed on the field linearly extrapolated at \( z = 0 \), and \( \delta_c(z) = \Delta_c(z)D(z = 0)/D(z) \) and \( \Delta_c(z) \) depends very weakly on redshift and \( \Delta_c(z = 0) \approx 1.68 \).

Eq. (14) however cannot be computed analytically and exactly starting from Eq. (11): some approximations need to be done in order to obtain an analytically manageable expression. Two approaches have been taken so far in the literature which we will briefly review below.

3.1 MVJ approach. The authors of Ref. [31] proceed by first performing the integration over \( \delta_R \) to obtain an exact expression for \( P(> \delta_c | z, M) \). At this point they expand the generating functional to the desired order, e.g., keeping only terms up to the skewness, then perform a Wick rotation to change variables and finally a saddle-point approximation to evaluate the remaining integral. The saddle point approximation is very good for large thresholds \( \delta_c/\sigma_M \gg 1 \), thus for rare and massive peaks. For the final expression for the mass function they obtain:

\[
\hat{n}(M, z) = \frac{3H_0^2 \Omega_m}{8\pi G M^2} \frac{1}{\sqrt{2\pi \sigma_M}} \exp \left[ -\frac{\delta_c^2}{2\sigma_M^2} \right] \times \\
\exp \left[ \frac{dS_{3, M}}{6 d\ln M} (\frac{\delta_c}{\sigma_M} - \frac{S_{3, M} \sigma_M}{6} (\frac{\delta_c^4}{\sigma_M^4} - 2 \frac{\delta_c^2}{\sigma_M^2} - 1)) \\
+ \frac{1}{6} \frac{dS_{3, M}}{d\ln M} \frac{\delta_c^2}{\sigma_M^2} \right].
\] (15)

where \( \sigma_M \) denotes the rms value of the density field, the subscript \( M \) denotes that the density field has been smoothed on a scale \( R(M) \) corresponding to \( R(M) = [M^3/(4\pi)]^{1/3} \), and \( \delta_c = \delta_c \sqrt{1 - \delta_c S_{3, M}/3} \).

This derivation shows that the mass function in principle depends on all cumulants, but that if non-Gaussianity is described by a bispectrum (and all higher order connected correlations are assumed to be zero or at least negligible), it depends only on the skewness. The mass function does not carry explicit information about the shape of non-Gaussianity. Nevertheless for a given numerical value of \( f_{NL} \) the skewness can have different amplitude and scale dependence for different models of non-Gaussianity, as illustrated in Fig. (1).

3.2 LMSV approach The authors of Ref. [9] (LMSV) instead proceed by using the saddle point approximation in the expression for \( P(\delta_R) \) and then using the Edgeworth expansion truncated at the desired order. The resulting simplified expression for the PDF can then be integrated to obtain \( P(> \delta_c | z, R) \) and derived to obtain the mass function. The final mass function in this approximation is given by:

\[
\hat{n}(M, z) = \frac{3H_0^2 \Omega_m}{8\pi G M^2} \frac{1}{\sqrt{2\pi \sigma_M}} \exp \left[ -\frac{\delta_c^2}{2\sigma_M^2} \right] \times \\
\exp \left[ \frac{dS_{3, M}}{6 d\ln M} (\frac{\delta_c}{\sigma_M} - \frac{S_{3, M} \sigma_M}{6} (\frac{\delta_c^4}{\sigma_M^4} - 2 \frac{\delta_c^2}{\sigma_M^2} - 1)) \\
+ \frac{1}{6} \frac{dS_{3, M}}{d\ln M} \frac{\delta_c^2}{\sigma_M^2} \right].
\] (16)

Without knowledge of all higher cumulants one is forced to use approximate expressions for the mass function. By truncating the Edgeworth expansion at the linear order in the skewness, the resulting PDF is no longer positive-definite. The number of terms that should be kept in the expansion depends on \( \delta_c(z)/\sigma_M \). The truncation used is a good approximation of the true PDF if \( \delta_c(z)/\sigma_M \) is small (and non-Gaussianity is small) but for rare events (the tails of the distribution) \( \delta_c(z)/\sigma_M \) is large. One thus expect this approximation to break down for large masses, high redshift and high \( f_{NL} \).
Ref. [9] quantified the range of validity of their approximation by assuming that when terms proportional to $S_3^2$ become important is no longer valid to neglect terms proportional to higher-order cumulants. Then they define the validity regime of their mass function to be where corrections from the $S_3^2$ are unimportant. They find, as expected, that for very massive objects the approximation breaks down and that the upper mass limit for applicability of the mass function decreases with redshift and $f_{NL}$. But for low masses, redshifts and $f_{NL}$ their formula is better than the MVJ. On the other hand the MVJ range of validity extends to higher masses, redshifts and $f_{NL}$ values, as expected, as MVJ applied the saddle point approximation to $P(>\delta_c|M,z)$ which is an increasingly good approximation for rare objects.

Of course, the natural observable to apply this method to are not only galaxy surveys (and the clusters found there), but, especially suited, are the mass-selected large clusters surveys offered by on-going Sunyaev-Zeldovich experiments (e.g., Planck, ACT, SPT).

3.3 Comparison with N-body simulations

Before we proceed we should consider that the Press & Shechter formulation of the mass function even in the Gaussian initial conditions case, can be significantly improved see e.g., [34, 35, 36]. Much improved expressions have been extensively calibrated on Gaussian initial conditions N-body simulations. The major limitations in both the MVJ and LMSV derivations (since they follow the classic Press & Shechter formulation) are the assumption of spherical collapse and the sharp $k$-space filtering. In addition, the excursion set improvement on the original Press & Schechter swindle relies on the random-phase hypothesis, which is not satisfied for non-Gaussian initial conditions. Since these improvements of the mass function have not yet been generalized to generic non-Gaussian initial conditions (but work is on-going, see other contributions in this issue) the analytical results above should be used to model fractional corrections to the Gaussian case.

Thus the non-Gaussian mass function, $n_{NG}(M,z)$ can be written as a function of a Gaussian one, $n_G(M,z)$ (accurately calibrated on N-body simulations) with a non-Gaussian
correction factor $\mathcal{R}$ (see e.g., [37, 9]):

$$n_{NG}(M, z) = n_{G}(M, z)\mathcal{R}(S_3, M, z)$$

(17)

where

$$\mathcal{R}(S_3, M, z) = \frac{\hat{n}(M, z, f_{NL})}{\hat{n}(M, z, f_{NL} = 0)}$$

(18)

and $\hat{n}$ is given by the MVJ or LMSV approximation. The correction $\mathcal{R}$ can then be calibrated on N-body simulations.

Ref. [33] argue that the same correction that in the Gaussian case modifies the collapse threshold, $\delta_c$, to improve over the original Press & Shechter formulation, may apply to the non-Gaussian correction. The detailed physical interpretation of this is still matter of debate in the literature [38, 39, 40]. In summary, Ref. [33] proposes to write the non-Gaussian correction factor for the MVJ [31] case as:

$$\mathcal{R}_{NG}(M, z, f_{NL}) = \exp\left[\delta^3_{ec} \frac{S_{3, M}}{6\sigma^2_M}\right] \times$$

$$\left|\frac{1}{6} \frac{\delta_{ec}}{1 - \frac{\delta_{ec} S_{3, M}}{3}} \frac{d S_{3, M}}{d \ln \sigma_M} + \sqrt{1 - \frac{\delta_{ec} S_{3, M}}{3}}\right|,$$

(19)

and for the LMSV [9] case:

$$\mathcal{R}_{NG}(M, z, f_{NL}) = 1 + \frac{1}{6} \frac{\sigma^2_M}{\delta_{ec}} \times$$

$$\left[S_{3, M}\left(\frac{\delta^4_{ec}}{\sigma^4_M} - 2\frac{\delta^2_{ec}}{\sigma^2_M} - 1\right) + \frac{d S_{3, M}}{d \ln \sigma_M}\left(\frac{\delta^2_{ec}}{\sigma^2_M} - 1\right)\right],$$

(20)

where $\delta_{ec}$ denotes the modified critical density for collapse, which for high peaks is $\delta_{ec} \sim \delta_c \sqrt{q}$. Ref. [33] calibrated these expressions on N-body simulations to find $q = 0.75$. We anticipate here that the validity of this extrapolation (i.e. in terms of a correction to the critical collapse threshold) can be tested independently on the large-scale non-Gaussian halo bias as described in §4. Note that, in both cases, in the limit of small non-Gaussianity the correction factors reduce to

$$\mathcal{R} = 1 + S_{3, M}\frac{\delta^3_{ec}}{6\sigma^2_M}.$$

(21)

Fig. 3 The points show the non-Gaussian correction to the mass function as measured in the N-body simulations of [33]. Blue corresponds to $f_{NL} = 200$ and red to $f_{NL} = -200$. The dashed lines correspond to the MVJ formulation and the dot-dashed line to the LMSV formulation. In both cases the substitution $\delta_c \rightarrow \delta_{ec}$ has been performed. The y axis should be interpreted as $\mathcal{R}(M, z, f_{NL} = 200)$. Reproduced from fig. 7 Ref. [33].

Non-Gaussian mass functions have been computed from simulations and compared with different theoretical predictions in several works [41, 42, 43, 44, 45]. In the past, conflicting results were reported, but the issue seems to have been settled: there is agreement among mass function measured from different non-Gaussian simulations performed by different groups as shown for example in Fig. (2). As expected both MVJ and LMSV prescriptions for the non-Gaussian correction to the mass function agree with the simulation results, provided one make the substitution $\delta_c \rightarrow \delta_{ec}$, with some tentative indication that MVJ may be better for very massive objects while LMSV performs better for less rare events. This is shown in Fig. (3) where the points represent measurements from N-body simulations presented in [33].

3.4 Voids While galaxy clusters form at the highest overdensities of the primordial density
field and probe the high density tail of the PDF, voids form in the low density regions and thus probe the low-density tail of the PDF. Most of the volume of the evolved universe is underdense, so it seems interesting to pay attention to the distribution of underdense regions. A void distribution function can be derived in an analogous way to the Press Schechter mass function by realizing that negative density fluctuations grow into voids [32]: a critical underdensity $\delta_v$ is necessary for producing a void and this plays the role of the critical overdensity $\delta_c$ for producing bound objects (halos). The more underdense a void is the more negative $\delta_v$ becomes. The precise value of $\delta_v(z)$ depends on the precise definition of a void (and may depend on the observables used to find voids); realistic values of $\delta_v(z = 0)$ are expected to be $\gtrsim -1$. In the absence of a better prescription, here, following [32], $\delta_v$ is treated as a phenomenological parameter and results are shown for a range of $\delta_v$ values. To derive the non-Gaussian void probability function one proceeds as above with the only subtlety that $\delta_v$ is negative and that $P(> \delta) = 1 - P(\leq \delta)$ thus $|dP(> \delta)/dM| = |dP(\leq \delta)/dM|$. Thus the void PDF as a function of $|\delta_v|$ can be obtained from the PDF of MVJ [31] or LMSV [9], provided one keeps track of the sign of each term. For example in the LMSV approximation the void distribution function becomes [32]:

$$
\hat{n}(R, z, f_{NL}) = \frac{9}{2\pi^2} \sqrt{\frac{\pi}{2}} R \left\{ \frac{1}{2} e^{-R^2/2\sigma^2} \left\{ \frac{d \ln \sigma_M}{d \ln M} \right\} + \frac{1}{6} \frac{dS_3}{dM} \sigma_M \left( \frac{\delta^2_v}{\sigma^2_M} - 1 \right) \right\}.
$$

(22)

where the expression is reported as a function of the smoothing radius rather than the mass, since a void Lagrangian radius is probably easier to determine than its mass.

Note that while a positive skewness ($f_{NL} > 0$) boosts the number of halos at the high mass end (and slightly suppress the number of low-mass halos), it is a negative skewness that will increase the voids size distribution at the largest voids end (and slightly decrease it for small void sizes). Ref. [32] concluded that the abundance of voids is sensitive to non-Gaussianity: $|\delta_v|$ is expected to be smaller than $\delta_c$ by a factor 2 to 3. If voids probe the same scales as halos then they should provide constraints on $f_{NL}$ 2 to 3 times worse. However voids may probe slightly larger scales than halos: in many non-Gaussian models, $S_{NL}^2$ increases with scales (see e.g., Fig. [1]), compensating for the threshold.

The approach reviewed here provides a rough estimate of the fractional change in abundance due to primordial non-Gaussianity but will not provide reliably the abundance itself. It is important to stress here that rigorously quantitative results will need to be calibrated on cosmological simulations and mock survey catalogs.

4 Effects on the Halo power spectrum

Recently, Refs. [43, 46] showed that primordial non-Gaussianity affects the clustering of dark matter halos (i.e., density extrema) inducing a scale-dependent bias for halos on large scales. This can be seen for example by considering halos as regions where the (smoothed) linear density field exceeds a suitable threshold. All correlations and peaks considered in the section are those of the initial density field (linearly extrapolated to the present time). Thus for example in the Gaussian case [47, 48, 49] for high peaks we would have the following relation between the correlation function of halos of mass $M$, $\xi_M(r)$ and that of the dark matter distribution smoothed on scale $R$, corresponding to mass $M$, $\xi_R(r)$:

$$
\xi_M(r) \approx b_L^2 \xi_R(r)
$$

(23)

where $b_L \approx \delta_c/\sigma^2_R$ denotes the Lagrangian bias, although more refined expressions can be found in e.g., [50] and [51].
The Lagrangian bias appears here because correlations and peaks are those of the initial density field (linearly extrapolated). Making the standard assumptions that halos move coherently with the underlying dark matter, one can obtain the final Eulerian bias as \( b_E = 1 + b_L \), using the techniques outlined in [52], [53], and [50] and [51].

The two-point correlation function of regions above a high threshold has been obtained, for the general non-Gaussian case, in [54], [55] and [56]:

\[
\xi_{h,M}(x_1 - x_2) = -1 + \exp[X]
\]

where

\[ X = \sum_{N=2}^{\infty} \sum_{j=1}^{N-1} \frac{\nu^N \sigma^N_R}{j!(N-1)!} \xi^{(N)}(x_1, x_2, \ldots, x_{j+N}) \],

(25)

where \( \nu = \delta_c \sigma_R \). For large separations the exponential can be expanded to first order. This is what we will do in what follows but we will comment on this choice below.

For small non-Gaussianities\(^1\) we can keep terms up to the three-point correlatin function \( \xi^{(3)} \), obtaining that the correction to the halo correlation function, \( \Delta \xi_h \) due to a non-zero three-point function is given by:

\[
\Delta \xi_h = \frac{\nu^3}{2\sigma_R^3} \left[ \xi^{(3)}_R(x_1, x_2, x_2) + \xi^{(2)}_R(x_1, x_1, x_2) \right]
\]

(26)

For a general bispectrum \( B(k_1, k_2, k_3) \) this yields a correction to the power spectrum (see [46] for steps in the derivation):

\[
\frac{\Delta P}{P} = \frac{\delta_c(z)}{M_R(k)} \frac{1}{4\pi^2\sigma_R^2} \int dk_1 k_1^2 M_R(k_1) \times \int_{-1}^{1} d\mu M_R(\sqrt{\alpha}) B_\phi(k_1, \sqrt{\alpha}, k) \frac{P_\phi(k_1)}{P_\phi(k)} \]  

(27)

Fig. 4 The scale-dependence of the large-scale halo bias induced by a non-zero bispectrum for different types of non-Gaussianity. The dashed line corresponds to the local type and the dot-dot-dot-dashed to equilateral type. Figure reproduced from [57].

where we have made the substitution \( \alpha = k_1^2 + k_2^2 + 2k_1k_2) \mu \). Here \( M_R = W_R \cdot M \). The effect on the halo bias is \( \Delta b_h / b_h = \frac{1}{2} \frac{\Delta b}{b} \) and thus

\[
b_h^{NL} = 1 + \frac{\Delta_c(z)}{\sigma_R^2 D^2(z)} \left[ 1 + \delta_c(z) \beta_R(k) \right],
\]

(28)

where the expression for \( \beta \) can be obtained by comparing to Eq. [27]. The term \( \Delta_c(z)/[\sigma_R^2 D^2(z)] \approx b^G - 1 \) can be recognized as the Gaussian Lagrangian halo bias.

So far the derivation is generic for all types of non-Gaussianity specified by a given bispectrum. We can then consider specific cases. In particular for local non-Gaussianity we obtain:

\[
\beta_R(k) = \frac{2f_{NL}}{8\pi^2\sigma_R^2 M_R(k)} \int dk_1 k_1^2 M_R(k_1) P_\phi(k_1) \times \int_{-1}^{1} d\mu M_R(\sqrt{\alpha}) \frac{P_\phi(\sqrt{\alpha})}{P_\phi(k)} + 2). \]  

(29)

Thus \( \Delta b_h / b_h \) is \( 2f_{NL} \) times a redshift-dependent factor \( \Delta_c(z)/D(z) = \delta_c(z) \), times a \( k \)-and mass-dependent factor. The function \( \beta_R(k) \) is shown as the dashed line in Fig. (4). This result for the local non-Gaussianity has been derived in at least three other ways. Ref. [43].

\(^1\) Effectively that is for values of \( f_{NL} \) consistent with observations
generalize the Kaiser [47] argument of high peak bias for the local non-Gaussianity. Starting from \( \nabla^2 \Phi = \nabla^2 \phi + 2 f_{NL}[\phi \nabla^2 \phi + \nabla \phi]^2 \) where near peaks \( |\nabla \phi|^2 \) is negligible they obtain \( \delta = \delta_{NL} = 1 + 2 f_{NL} \phi \). The Poisson equation to convert \( \phi \) in \( \delta \) then gives the scale-dependence. More details are presented elsewhere in this issue.

Ref. [58] work in the peak-background split. This approach is especially useful to understand that it is the coupling between very large and small scales introduced by local (squeezed-configurations) non-Gaussianity to boost (or suppress) the peaks clustering. In this approach, the density field can be written as \( \rho(x) = \bar{\rho}(1 + \delta_l + \delta_s) \) where \( \delta_l \) denotes long wavelength fluctuations and \( \delta_s \) short wavelength fluctuations. \( \delta_l \) is the one responsible for modulating halo formation (i.e., to boost peaks above the threshold for collapse), so the halo number density is \( n = \bar{n}(1 + b_L \delta_l) \) and \( b_L = \bar{n}^{-1} \partial n / \partial \delta_l \).

In the local non-Gaussian case they decompose the Gaussian field \( \phi \) as a combination of long and short wavelength fluctuations \( \phi = \phi_l + \phi_s \) thus \( \Phi = \phi_l + f_{NL} \phi_l^2 + (1 + 2 f_{NL} \phi_l) \phi_s + f_{NL} \phi_s^2 + \text{const.} \). Also in this non-Gaussian case one can split the density field \( \delta \) in \( \delta_l \) and \( \delta_s \) and relate this to \( f_{NL} \) (it is easier to work in Fourier space): \( \delta_l(k) = \alpha(k) \phi_l(k) \) and \( \delta_s = \alpha(k)(1 + 2 f_{NL} \phi_l) \phi_s + f_{NL} \phi_s^2 \) \( \equiv \alpha(k) \{ X_1 \phi_s + X_2 \phi_s^2 \} \) the last equality giving the definition of \( X_1 \) and \( X_2 \). Note that \( \delta_s \) cannot be ignored here because \( \phi_l \) enters in \( X_1 \), in other words, local non-Gaussianity couples long and short wavelength modes. The local halo number density is now function of \( \delta_l, X_1 \) and \( X_2 \) yielding the following result for the Lagrangian halo bias:

\[
\begin{align*}
\Delta b = f_{NL} \delta_c (b_{NL}^{2 \text{G}} + 1)(\partial \delta_c / \partial \delta_l) & = b_{NL}^{2 \text{G}} \left[ \frac{\partial n}{\partial \delta_l} + 2 f_{NL} \frac{\partial \phi_l}{\partial \delta_l} \frac{\partial n}{\partial X_1} \right] \\
2 f_{NL} \frac{\partial \phi_l}{\partial \delta_l} \frac{\partial \ln n}{\partial \ln \sigma_8} & = b_{NL}^{2 \text{G}} (1 + 2 f_{NL} \alpha(k) \delta_c) \tag{30}
\end{align*}
\]

where \( \alpha(k) \) encloses the scale-dependence of the effect. Ref. [59] rederives the ellipsoidal collapse for small deviations from Gaussianity of the local type. They find that a non-zero \( f_{NL} \) modifies the threshold for collapse, the modification is proportional to \( f_{NL} \). This should sound familiar from §3. They then use the definition \( b_L = \bar{n}^{-1} \partial n / \partial \delta_l \) keeping track of the fact that \( \delta \) is “modulated” by \( f_{NL} \).

The effect of the non-Gaussian halo bias on the power spectrum is shown in Fig. (5) where the points are measurements from an N-body simulations of [33] (see figure caption for more details).

The above result

\[
\Delta b = f_{NL} \delta_c (b_{NL}^{2 \text{G}} + 1) \left[ \frac{\delta_c(z_f)}{\sigma_M^2} - \frac{1}{\delta_c(z_f)} \right] \tag{31}
\]

where \( \beta_G(k) = f_{NL} \beta_{NL}^{2 \text{G}}(k) \), can be improved in several ways.

First of all, we have not made any distinction between the redshift at which the object is being observed \( (z_o) \) and that at which is being formed \( (z_f) \). Except for the rarest events this should be accounted for. The Gaussian Lagrangian bias expression used so far is an approximation, a more accurate expression is [52, 53, 50]

\[
b_{NL}^{G} = \frac{1}{D(z_o)} \left[ \frac{\delta_c(z_f)}{\sigma_M^2} - \frac{1}{\delta_c(z_f)} \right] \tag{32}
\]

Then, the halo bias expressions are derived within the “classical” Press & Schechter theory, as we have seen in §3, subsequent improvements on the mass function can be seen as a correction to the collapse threshold. In the expression for the Gaussian halo bias \( b_{NL}^{G} = \bar{n}^{-1} \partial n^G / \partial \delta_c \) one can consider mass functions that are better fit to simulations than the standard Press & Schechter one obtaining:

\[
b_{NL}^{G} = \frac{1}{D(z_o)} \left[ \frac{q \delta_c(z_f)}{\sigma_M^2} - \frac{1}{\delta_c(z_f)} \right] + \frac{2 q \delta_c^2(z_f)}{\delta_c(z_f) D(z_o)} \left[ 1 + \left( \frac{q \delta_c^2(z_f)}{\sigma_M^2} \right)^{-1} \right]. \tag{33}
\]
The parameters \( q \) and \( p \) account for non-spherical collapse and fit to numerical simulations yield \( q \sim 0.75, p = 0.3 \) e.g., [34]. In this expression the term in the second line is usually sub-dominant. The term “\(-1/\delta_c\)” in the first line is known as “anti-bias”, and it becomes negligible for old halos \( z_f \ll z_o \). Note that by including the anti-bias correction in \( b_h \) of Eq. [31] one recovers the “recent mergers” approximation of Ref. [58].

The same correction should also apply to the non-Gaussian correction to the halo bias:

\[
\Delta b = f_{NL} q' \delta_c (b_h^G - 1) \beta_{R}^{\text{fNL}=1}(k) .
\]

(34)

where \( q' \) should coincide with \( q \) above; it can be calibrated to N-body simulations and is found indeed to be \( q' = 0.75 \) [33]. The non-Gaussian halo bias prediction and results from N-body simulations with local non-Gaussianity are shown in Fig. (6).

Finally one may note that for \( f_{NL} \) large and negative, Eq. [27, 28] would formally yield \( b_h^{\text{fNL}} \) and \( P_h(k) \) negative on large enough scales. This is a manifestation of the breakdown of the approximations made: a) all correlations of higher order than the bispectrum were neglected: for large NG this truncation may not hold; b) The exponential in Eq. (24) was expanded to linear order. This however could be easily corrected for, remembering that the \( P(k) \) obtained from Eq. (27) is in reality the Fourier transform of \( X \), the argument of the exponential. One would then compute the halo correlation function using Eq. (24) and Fourier transforming back to obtain the halo power-spectrum.

So far we have concentrated on local non-Gaussianity, but the expression of Eqs. (27, 28) is more general. Using this formulation, Ref. [57] computed the quantity \( \beta_{R}^{\text{fNL}=1}(k) \) for several types of non-Gaussianity (equilateral, local and enfolded); this is shown in Fig. (4). It is clear that the non-Gaussian halo-bias effect has some sensitivity to the bispectrum shape, for example the effect for the equilateral
Fig. 6 The quantity $\Delta b/b$ as function of $k$, for simulation snapshots at $z = 0.44$, 1.02 and 1.54. Simulation outputs and theory lines are shown for $f_{NL} = \pm 100$ and $f_{NL} = \pm 200$. Figure reproduced from [33]. The $f_{NL}$ values reported in this figure legend should be interpreted as $f^{LSS}_{NL}$.

type of non-Gaussianity is suppressed by orders of magnitude compared to the local-type and the flattened case is somewhere in the middle. Fig. (4) also shows a type of non-Gaussianity arising from General-relativistic (RG) corrections on scales comparable to the Hubble radius. Note that perturbations on super-Hubble scales are initially needed in order to “feed” the GR correction terms. In this respect the significance of this contribution is analogous to the well-known large-scale anti-correlation between CMB temperature and E-mode polarization: it is a consequence of the properties of the inflationary mechanism to lay down the primordial perturbations. This effect has the same magnitude as a local non-Gaussianity with $f_{NL} \gtrsim 1$.

The next logical step is then to ask how well present or forthcoming data could constrain non-Gaussianity using the halo-bias effect. It is interesting to note that surveys that aim at measuring Baryon Acoustic Oscillations (BAO) in the galaxy distribution to constrain dark energy are well suited to also probe non-Gaussianity: they cover large volumes and their galaxy number density is well suited so that on the scale of interest (both for BAO and non-Gaussianity) shot noise does not dominate the signal. Photometric surveys are also well suited: as the non-Gaussian signal is localized at very large scales and is a smooth function of $k$, the photo-z smearing effects are unimportant.

The theory developed so far describes the clustering of halos while we observe galaxies. Different galaxy populations occupy dark matter halos following different prescriptions. If we think in the halo-model framework (e.g., [60] and references therein) at very large scales only the “two halo” contribution matters and the details of the halo occupation distribution (the “one-halo” term) become unimportant.

What is important to keep in mind is that the effect of the non-Gaussianity parameter one wants to measure, $f_{NL}$, is fully degenerate with the value of the Gaussian (small scales) halo bias. Fig. (7) shows the dependence of the non-Gaussian correction on the Gaussian bias.

Thus highly biased tracers will show a a larger non-Gaussian effect for the same $f_{NL}$ value. Of course for a given cosmological model the Gaussian bias can be measured accurately by comparing the predicted dark matter power spectrum with the observed one. Alternatively, two differently biased tracers can be used in tandem to disentangle the two effects [61, 62].
Since clustering amplitude may depend on the entire halo history, it becomes then interesting to model in details the dependence of the effect on the halo merger tree (Reid et al., in preparation).

4.1 Outlook for the future How well can this method do to constrain primordial non-Gaussianity compared with the other techniques presented here? The Integrated Sachs Wolfe (ISW) effect offers a window to probe clustering on the largest scales (where the signal is large); on the other hand, a measurement of clustering of tracers of dark matter halos is a very direct window into this effect. A Fisher matrix approach [63, 59, 58] shows that the ISW signal is weighted at relatively low redshift (where dark energy starts dominating) while the non-Gaussian signal grows with redshift, thus making the shape of the halo power spectrum a more promising tool. An overview of current constraints from different approaches can be found in table 1 and future forecasts in table 2 for non-Gaussianity of the local type. Large, mass-selected cluster samples as produced by SZ-based experiments will provide a optimally suited data-set for this technique (see e.g. [64]).

While for a given $f_{NL}$ model such as the local one, methods that exploit the non-Gaussian bias seem to yield the smallest error-bars for large-scale structure, it should be kept in mind that the bispectrum can be used to investigate the full configuration dependence of $f_{NL}$ and thus is a very powerful tool to discriminate between different type of non-Gaussianity. In addition CMB-bispectrum and halo bias test non-Gaussianity on very large scales while the large scale structure bispectrum mostly probes mildly non-linear scales. As primordial non-Gaussianity may be scale-dependent, all these techniques are highly complementary.

The above estimates assume that the underlying cosmological model is known. The large-scale shape of the power spectrum can be affected by cosmology. Carbone et al. (in prep.) explore possible degeneracies between $f_{NL}$ and cosmological parameters. They find that the parameters that are most strongly correlated with $f_{NL}$ are parameters describing dark energy clustering, neutrino mass and running of the primordial power spectrum spectral slope. For surveys that cover a broad redshift range the error on $f_{NL}$ degrade little when marginalizing over these extra parameters: the peculiar redshift dependence of the non-Gaussian signal lifts the degeneracy.

5 Conclusions

A natural question to ask at this point may be “what observable will have better chances to constrain primordial non-Gaussianity?”

In principle the abundance of rare events is a very powerful probe of non-Gaussianity; however, in practice, it is limited by the practical difficulty of determining the mass of the observed objects and its corresponding large uncertainty in the determination. This point is stressed e.g., in [9]. With the advent of high-precision measurements of gravitational lensing by massive clusters, the mass uncertainty, at least for small to moderate size clusters samples...
Table 1 Current recent 2 − σ constraints on local $f_{\text{NL}}$

| Data/method | $f_{\text{NL}}$ | reference |
|-------------|---------------|-----------|
| Photometric LRG - bias | $63^{+54+101}_{-85-131}$ | Slosar et al. 2008 |
| Spectroscopic LRG- bias | $70^{+74+139}_{-83-191}$ | Slosar et al. 2008 |
| QSO - bias | $8_{-37-77}^{+26+47}$ | Slosar et al. 2008 |
| combined | $28_{-24-57}^{+23+42}$ | Slosar et al. 2008 |
| NVSS–ISW | $105^{+647+755}_{-337-1155}$ | Slosar et al. 2008 |
| NVSS–ISW | $236 \pm 127(2 - \sigma)$ | Afshordi&Tolley 2008 |
| WMAP3-Bispectrum | $30 \pm 84$ | Spergel et al. (WMAP) 2007 |
| WMAP3-Bispectrum | $32 \pm 68$ | Creminelli et al. 2007 |
| WMAP3-Bispectrum | $87 \pm 60$ | Yadav & Wandelt 2008 |
| WMAP-Bispectrum | $38 \pm 42$ | Smith et al. 2009 |
| WMAP5-Bispectrum | $51 \pm 60$ | Komatsu et al. (WMAP) 2008 |
| WMAP5-Minkowski | $-57 \pm 121$ | Komatsu et al. (WMAP) 2008 |

Table 2 Forecasts 1 − σ constraints on local $f_{\text{NL}}$

| Data/method | $\Delta f_{\text{NL}} (1 - \sigma)$ | reference |
|-------------|-------------------------------|-----------|
| BOSS–bias | 18 | Carbone et al. 2008 |
| ADEPT/Euclid–bias | 1.5 | Carbone et al. 2008 |
| PANNStarrs –bias | 3.5 | Carbone et al. 2008 |
| LSST–bias | 0.7 | Carbone et al. 2008 |
| LSST–ISW | 7 | Afshordi&Tolley 2008 |
| BOSS–bisperectrum | 35 | Sefusatti & Komatsu 2008 |
| ADEPT/Euclid –bispectrum | 3.6 | Sefusatti & Komatsu 2008 |
| Planck-Bispectrum | 3 | Yadav et al. 2007 |
| BPOL-Bispectrum | 2 | Yadav et al. 2007 |

Can be greatly reduced. Forthcoming Sunyaev-Zeldovich experiments will provide large samples of mass-selected clusters which could then be followed up by lensing mass measurements (see e.g., [65, 66]). So far there is only one very high redshift ($z = 1.4$) very massive $M \approx 8 \times 10^{14} M_\odot$ with high-precision mass determination via gravitational lensing [67]. Ref. [68] pointed out that this object is extremely rare, for Gaussian initial conditions there should be 0.002 such objects or less in the surveyed area, which is uncomfortably low probability. But the cluster mass is very well determined: a non-Gaussianity still compatible with CMB constraints could bring the probability of observing of the object to more comfortable values. This result should be interpreted as a “proof of principle” showing that this a potentially powerful avenue to pursue.

The measurement of the three-point correlation function allows one to map directly the shape-dependence of the bispectrum. For large-scale structures the limiting factors are the large non-Gaussian contribution induced by gravitational evolution and the uncertainly of the non-linear behavior of galaxy bias.

The halo-bias approach can yield highly competitive constraints, but it is less sensitive to the bispectrum shape. Still, the big difference in the magnitude and shape of the scale-dependent biasing factor between different non-Gaussian models implies that the halo bias can become a
Table 3: Forecasted non-Gaussianity constraints: A) [20] B) [63] C) [69, 70] E) [57])
e.g., [15]

| CMB Bispectrum | Halo bias |
|----------------|----------|
| type NG        | Planck   | BPol | Euclid | LSST |
| Local          | $3^A$    | $2^B$ | $1.5^B$ | $0.7^B$ |
| Equilateral    | $25^C$   | $14^C$ | −       | −     |
| Enfolded       | $O10$    | $O10$ | $39^E$  | $18^E$ |
| 1-$\sigma$ errors |          |       |         |       |
| $\#\sigma$ detection |          |       |         |       |
| GR secondaries | N/A      | $1^E$ | $2^E$   | N/A   |

useful tool to study shapes when combined with e.g. measurements of the CMB bispectrum. Table 3 highlights this complementarity. For example, one could envision different scenarios.

If non-Gaussianity is local with negative $f_{NL}$ and CMB obtains a detection, then the halo bias approach should also give a high-significance detection (GR correction and primordial contributions add up), while if it is local but with positive $f_{NL}$, the halo-bias approach could give a lower statistical significance for small $f_{NL}$ as the GR correction contribution has the opposite sign.

If CMB detects $f_{NL}$ at the level of $\sim 10$ and of a form that is close to local, but halo bias does not detect it, then the CMB bispectrum is given by secondary effects.

If CMB detects non-Gaussianity but is not of the local type, then halo bias can help discriminate between equilateral and enfolded shapes: if halo bias sees a signal it indicates enfolded type.; if halo bias does not see a signal it indicates equilateral type. Thus even a non-detection of the halo-bias effect, in combination with CMB constraints can have an important discriminatory power.

In any case, if the simplest inflationary scenario holds, for surveys like Euclid and LSST, the halo-bias approach is expected to detect a non-Gaussian signal very similar to the local type signal with an amplitude of $f_{NL} \sim -1.5$ which is due to large-scales GR corrections to the Poisson equation. This effect should leave no imprint in the CMB: once again the combination of the two observable can help enormously to discriminate among models for the origin of cosmological structures.

In addition we should bear in mind that non-Gaussianity may be scale-dependent. In fact for models like DBI inflation it is expected to be scale-dependent. A proposed parameterization of the scale-dependence of non-Gaussianity is given by:

$$B_{\phi}(k_1, v k_2, k_3) = f_{NL} \left( \frac{K}{k_p} \right)^{n_g} F(k_1, k_2, k_3)$$ (35)

where $K = (k_1 k_2 k_3)^{1/3}$ [70], $k_p$ denotes the pivot and $n_g$ the slope or running of non-Gaussianity ,although other authors prefer to use $K = (k_1 + k_2 + k_3)/3$ [9,59] as for squeezed configurations $K \neq 0$. It is still an open issue which parameterization is better in practice.

In any case different observables probe different scales (see Fig 8) and their complementary means that “the combination is more than the sum of the parts”.

What is clear, however, is that the thorny systematic effects that enter in all these approaches will require that a variety of complementary avenues be taken to establish a robust detection of primordial non-Gaussianity.

Acknowledgements. This work is supported by MICCIN grant AYA2008-03531 and FP7-IDEAS-Phys.LSS 240117. I would like to thank my closest collaborators in many of the articles reviewed here: Carmelita Carbone, Klaus Dolag, Margherita Grossi, Alan Heavens, Raul Jimenez, Marc Kamionkowski, Marilena LoVerde, Sabino Matarrese, Lauro Moscardini, Sarah Shandera and my collaborators for the reported work-in-progress: Carmelita Carbone, Olga Mena, Beth Reid.
Fig. 8 Scale-dependent $f_{NL}$ and scales probed by different approaches mentioned here. The solid line has $n_g = 0$, the dotted line has $n_g = 0.2$ and the dashed one has $n_g = 0.6$. Hashed areas for CMB and halo-bias show allowed regions.

References

[1] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto. Non-Gaussianity from inflation: theory and observations. *Phys. Rep.*, 402:103–266, November 2004.
[2] D. S. Salopek and J. R. Bond. Nonlinear evolution of long-wavelength metric fluctuations in inflationary models. *Phys. Rev. D*, 42:3936–3962, December 1990.
[3] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach. The three-point correlation function of the cosmic microwave background in inflationary models. *ApJ*, 430:447–457, August 1994.
[4] L. Verde, L. Wang, A. F. Heavens, and M. Kamionkowski. Large-scale structure, the cosmic microwave background and primordial non-Gaussianity. *MNRAS*, 313:141–147, March 2000.
[5] E. Komatsu and D. N. Spergel. Acoustic signatures in the primary microwave background bispectrum. *Phys. Rev. D*, 63(6):063002–+, March 2001.
[6] D. Babich, P. Creminelli, and M. Zaldarriaga. The shape of non-Gaussianities. *Journal of Cosmology and Astro-Particle Physics*, 8:9–+, August 2004.
[7] P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark, and Z. Zaldarriaga M. Limits on non-Gaussianities from WMAP data. *Journal of Cosmology and Astro-Particle Physics*, 5:4–+, May 2006.
[8] P. Creminelli, L. Senatore, M. Zaldarriaga, and M. Tegmark. Limits on fNL parameters from Wilkinson Microwave Anisotropy Probe three-year data.
data. *Journal of Cosmology and Astro-Particle Physics*, 3:5–+, March 2007.

[9] M. LoVerde, A. Miller, S. Shandera, and L. Verde. Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures. *ArXiv e-prints*, 711, November 2007.

[10] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ArXiv e-prints*, 803, March 2008.

[11] E. Komatsu, N. Afshordi, N. Bartolo, D. Baumann, J. R. Bond, E. I. Buchbinder, C. T. Byrnes, X. Chen, D. J. H. Chung, A. Cooray, P. Creminelli, N. Dalal, O. Dore, R. Easther, A. V. Frolov, J. Khoury, W. H. Kinney, L. Kofman, K. Koyama, L. Leblond, J.-L. Lehners, J. E. Lidsey, M. Liguori, E. A. Lim, A. Linde, D. H. Lyth, J. Maldacena, S. Matarrese, L. McAllister, P. McDonald, S. Mukohyama, B. Ovrut, H. V. Peiris, A. Riotto, Y. Rodrigues, M. Sasaki, R. Scoccimarro, D. Seery, A. Sefusatti, K. M. Smith, A. A. Starobinsky, P. J. Steinhardt, F. Takahashi, M. Tegmark, A. J. Tolley, L. Verde, B. D. Wandelt, D. Wands, S. Weinberg, M. Wyman, A. P. S. Yadav, and M. Zaldarriaga. Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe. In *AGB Stars and Related Phenomena 2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010 of *Astronomy*, pages 158–+, 2009.

[12] D. M. Goldberg and D. N. Spergel. Microwave background bispectrum. II. A probe of the low redshift universe. *Physical Review D*, 59(10):103002–+, May 1999.

[13] L. Verde and D. N. Spergel. Dark energy and cosmic microwave background bispectrum. *Physical Review D*, 65(4):043007–+, February 2002.

[14] A. Cooray and W. Hu. Weak Gravitational Lensing Bispectrum. *ApJ*, 548:7–18, February 2001.

[15] A. Mangilli and L. Verde. Non-Gaussianity and the CMB bispectrum: Confusion between primordial and lensing-Rees-Sciama contribution? *Phys. Rev. D*, 80(12):123007–+, December 2009.

[16] P. Serra and A. Cooray. Impact of secondary non-Gaussianities on the search for primordial non-Gaussianity with CMB maps. *PRD*, 77(10):107305–+, May 2008.

[17] D. Hanson, K. M. Smith, A. Challinor, and M. Liguori. CMB lensing and primordial non-Gaussianity. *Phys. Rev. D*, 80(8):083004–+, October 2009.

[18] A. P. Yadav and B. D. Wandelt. CMB tomography: Reconstruction of adiabatic primordial scalar potential using temperature and polarization maps. *Phys. Rev. D*, 71(12):123004–+, June 2005.

[19] A. P. S. Yadav, E. Komatsu, and B. D. Wandelt. Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background. *ApJ*, 664:680–686, August 2007.

[20] A. P. S. Yadav and B. D. Wandelt. Evidence of Primordial Non-Gaussianity (fNL) in the Wilkinson Microwave Anisotropy Probe 3-Year Data at 2.8σ. *Physical Review Letters*, 100(18):181301–+, May 2008.

[21] K. M. Smith, L. Senatore, and M. Zaldarriaga. Optimal limits on f_{NL} from WMAP 5-year data. *Journal of Cosmology and Astro-Particle Physics*, 9:6–+, September 2009.

[22] P. Chingangbam and C. Park. Statistical nature of non-Gaussianity from cubic order primordial perturbations: CMB map simulations and genus statistic. *Journal of Cosmology and Astro-Particle Physics*, 12:19–+, December 2009.

[23] P. Catelan, F. Lucchin, S. Matarrese, and L. Moscardini. Eulerian perturbation theory in non-flat universes: second-order approximation. *MNRAS*, 276:39–56, September 1995.

[24] S. Matarrese, L. Verde, and A. F. Heavens. Large-scale bias in the Universe: bispectrum method. *MNRAS*, 290:651–662, October 1997.

[25] L. Verde, A. F. Heavens, and S. Matarrese. Projected bispectrum in spherical harmonics and its application to angular galaxy catalogues. *MNRAS*, 318:584–598, October 2000.

[26] LSST Science Collaborations. LSST Science Book, Version 2.0. *ArXiv e-prints*, December 2009.
[27] E. Sefusatti and E. Komatsu. Bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and nonlinear galaxy bias. *Phys. Rev. D*, 76(8):083004–+, October 2007.

[28] L. Verde, A. F. Heavens, W. J. Percival, S. Matarrese, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, P. Norberg, J. A. Peacock, B. A. Peterson, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: the bias of galaxies and the density of the Universe. *MNRAS*, 335:432–440, September 2002.

[29] L. Verde and A. F. Heavens. On the Trispectrum as a Gaussian Test for Cosmology. *ApJ*, 553:14–24, May 2001.

[30] W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974.

[31] S. Matarrese, L. Verde, and R. Jimenez. The Abundance of High-Redshift Objects as a Probe of Non-Gaussian Initial Conditions. *ApJ*, 541:10–24, September 2000.

[32] M. Kamionkowski, L. Verde, and R. Jimenez. The void abundance with non-gaussian primordial perturbations. *Journal of Cosmology and Astro-Particle Physics*, 1:10–+, January 2009.

[33] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini. Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations. *MNRAS*, 398:321–332, September 2009.

[34] R. K. Sheth and G. Tormen. Large-scale bias and the peak background split. *MNRAS*, 308:119–126, September 1999.

[35] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The mass function of dark matter haloes. *MNRAS*, 321:372–384, February 2001.

[36] D. S. Reed, R. Bower, C. S. Frenk, A. Jenkins, and T. Theuns. The halo mass function from the dark ages through the present day. *MNRAS*, 374:2–15, January 2007.

[37] L. Verde, R. Jimenez, M. Kamionkowski, and S. Matarrese. Tests for primordial non-Gaussianity. *MNRAS*, 325:412–418, July 2001.

[38] J. Lee and S. F. Shandarin. Comparison of Analytical Mass Functions with Numerical Simulations. *ApJ*, 517:L5–L8, May 1999.

[39] B. E. Robertson, A. V. Kravtsov, J. Tinker, and A. R. Zentner. Collapse Barriers and Halo Abundance: Testing the Excursion Set Ansatz. *ApJ*, 696:636–652, May 2009.

[40] M. Maggiore and A. Riotto. The Halo Mass Function from Excursion Set Theory. III. Non-Gaussian Fluctuations. *ArXiv e-prints*, March 2009.

[41] X. Kang, P. Norberg, and J. Silk. Can a large-scale structure probe cosmic microwave background-constrained non-Gaussianity? *MNRAS*, 376:343–347, March 2007.

[42] M. Grossi, K. Dolag, E. Branchini, S. Matarrese, and L. Moscardini. Evolution of massive haloes in non-Gaussian scenarios. *MNRAS*, 382:1261–1267, December 2007.

[43] N. Dalal, O. Dore, D. Huterer, and A. Shirokov. The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects. *preprint*, 2007.

[44] V. Desjacques, U. Seljak, and I. T. Iliev. Scale-dependent bias induced by local non-Gaussianity: a comparison to N-body simulations. *MNRAS*, 396:85–96, June 2009.

[45] A. Pillepich, C. Porciani, and O. Hahn. Table of Contents. *MNRAS*, pages 1959–+, December 2009.

[46] S. Matarrese and L. Verde. The Effect of Primordial Non-Gaussianity on Halo Bias. *ApJ*, 677:L77–L80, April 2008.

[47] N. Kaiser. On the spatial correlations of Abell clusters. *ApJL*, 284:L9–L12, September 1984.

[48] H. D. Politzer and M. B. Wise. Relations between spatial correlations of rich clusters of galaxies. *ApJ*, 285:L1–L3, October 1984.

[49] L. G. Jensen and A. S. Szalay. N-point correlations for biased galaxy formation. *ApJ*, 305:L5–L9, June 1986.

[50] H. J. Mo and S. D. M. White. An analytic model for the spatial clustering of dark matter haloes. *MNRAS*, 282:347–361, September 1996.

[51] P. Catelan, F. Lucchin, S. Matarrese, and C. Porciani. The bias field of dark matter haloes.
[52] G. Efstathiou, C. S. Frenk, S. D. M. White, and M. Davis. Gravitational clustering from scale-free initial conditions. *MNRAS*, 235:715–748, December 1988.

[53] S. Cole and N. Kaiser. Biased clustering in the cold dark matter cosmogony. *MNRAS*, 237:1127–1146, April 1989.

[54] B. Grinstein and M. B. Wise. Non-Gaussian fluctuations and the correlations of galaxies or rich clusters of galaxies. *ApJ*, 310:19–22, November 1986.

[55] S. Matarrese, F. Lucchin, and S. A. Bonometto. A path-integral approach to large-scale matter distribution originated by non-Gaussian fluctuations. *ApJLett*, 310:L21–L26, November 1986.

[56] F. Lucchin, S. Matarrese, and N. Vittorio. Scale-invariant clustering and primordial biasing. *ApJLett*, 330:L21–L23, July 1988.

[57] L. Verde and S. Matarrese. Detectability of the Effect of Inflationary Non-Gaussianity on Halo Bias. *ApJ*, 706:L91–L95, November 2009.

[58] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan. Constraints on local primordial non-Gaussianity from large scale structure. *Journal of Cosmology and Astro-Particle Physics*, 8:31–+, August 2008.

[59] N. Afshordi and A. J. Tolley. Primordial non-Gaussianity, statistics of collapsed objects, and the integrated Sachs-Wolfe effect. *Phys. Rev. D*, 78(12):123507–+, December 2008.

[60] A. Cooray and R. Sheth. Halo models of large scale structure. *Phys. Rep.*, 372:1–129, December 2002.

[61] U. Seljak. Extracting Primordial Non-Gaussianity without Cosmic Variance. *Physical Review Letters*, 102(2):021302–+, January 2009.

[62] A. Slosar. Optimal weighting in $f_{NL}$ constraints from large scale structure in an idealised case. *Journal of Cosmology and Astro-Particle Physics*, 3:4–+, March 2009.

[63] C. Carbone, L. Verde, and S. Matarrese. Non-Gaussian Halo Bias and Future Galaxy Surveys. *ApJ*, 684:L1–L4, September 2008.

[64] C. Fedeli, L. Moscardini, and S. Matarrese. The clustering of galaxy clusters in cosmological models with non-Gaussian initial conditions: predictions for future surveys. *MNRAS*, 397:1125–1137, August 2009.

[65] C. Sealfon, L. Verde, and R. Jimenez. Stacking Weak-Lensing Signals of Sunyaev-Zel’dovich Clusters to Constrain Cluster Physics. *ApJ*, 649:118–128, September 2006.

[66] E. Rozo, E. S. Rykoff, A. Evrard, M. Becker, T. McKay, R. H. Wechsler, B. P. Koester, J. Hao, S. Hansen, E. Sheldon, D. Johnston, J. Annis, and J. Frieman. Constraining the Scatter in the Mass-richness Relation of maxBCG Clusters with Weak Lensing and X-ray Data. *ApJ*, 699:768–781, July 2009.

[67] M. J. Jee, P. Rosati, H. C. Ford, K. S. Dawson, C. Lidman, S. Perlmutter, R. Demarco, V. Strazzullo, C. Mullis, H. Böhringer, and R. Fassbender. Hubble Space Telescope Weak-lensing Study of the Galaxy Cluster XMMU J2235.3 - 2557 at z ≈ 1.4: A Surprisingly Massive Galaxy Cluster When the Universe is One-third of its Current Age. *ApJ*, 704:672–686, October 2009.

[68] R. Jimenez and L. Verde. Implications for primordial non-Gaussianity ($f_{NL}$) from weak lensing masses of high-z galaxy clusters. *Phys. Rev. D*, 80(12):127302–+, December 2009.

[69] Baumann et al. Probing Inflation with CMB Polarization. In S. Dodelson, D. Baumann, A. Cooray, J. Dunkley, A. Fraisse, et al., editor, *American Institute of Physics Conference Series*, volume 1141 of *American Institute of Physics Conference Series*, pages 10–120, June 2009.

[70] E. Sefusatti, M. Liguori, A. P. S. Yadav, M. G. Jackson, and E. Pajer. Constraining running non-gaussianity. *Journal of Cosmology and Astro-Particle Physics*, 12:22–+, December 2009.