Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physics Reports 402 (2004) 103–266
PHYSICS REPORTS
www.elsevier.com/locate/physrep

Non-Gaussianity from inflation: theory and observations

N. Bartolo[a], E. Komatsu[b], S. Matarrese[c, d, *], A. Riotto[d]

[a] *Astronomy Centre, University of Sussex Falmer, Brighton, BN1 9QH, UK*
[b] *Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[c] *Dipartimento di Fisica "G. Galilei", Università di Padova, via Marzolo 8, I-35131 Padova, Italy*
[d] *INFN, Sezione di Padova, via Marzolo 8, I-35131 Padova, Italy*

Accepted 9 August 2004
editor: M.P. Kamionkowski

Abstract

This is a review of models of inflation and of their predictions for the primordial non-Gaussianity in the density perturbations which are thought to be at the origin of structures in the Universe. Non-Gaussianity emerges as a key observable to discriminate among competing scenarios for the generation of cosmological perturbations and is one of the primary targets of present and future Cosmic Microwave Background satellite missions. We give a detailed presentation of the state-of-the-art of the subject of non-Gaussianity, both from the theoretical and the observational point of view, and provide all the tools necessary to compute at second order in perturbation theory the level of non-Gaussianity in any model of cosmological perturbations. We discuss the new wave of models of inflation, which are firmly rooted in modern particle physics theory and predict a significant amount of non-Gaussianity. The review is addressed to both astrophysicists and particle physicists and contains useful tables which summarize the theoretical and observational results regarding non-Gaussianity.

PACS: 98.80.Cq; 98.70.Vc

* Corresponding author.
E-mail address: sabino.matarrese@pd.infn.it (S. Matarrese).

0370-1573/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.physrep.2004.08.022

Contents

"*... the linear perturbations are so surprisingly simple that a perturbation analysis to second order may be feasible...*"
(Sachs and Wolfe, 1967)

1. Introduction

One of the relevant ideas in modern cosmology is represented by the inflationary paradigm. It is widely believed that there was an early epoch in the history of the Universe—before the epoch of primordial nucleosynthesis—when the Universe expansion was accelerated. Such a period of *cosmological inflation* can be attained if the energy density of the Universe is dominated by the vacuum energy density associated with the potential of a scalar field φ, called the inflaton field. Through its kinematic properties, namely the acceleration of the Universe, the inflationary paradigm can elegantly solve the flatness, the horizon and the monopole problems of the standard Big–Bang cosmology, and in fact the first model of inflation by Guth in 1981 [101] was introduced to address such problems. However all over the years inflation has become so popular also because of another compelling feature. It can explain the production of the first density perturbations in the early Universe which are the seeds for the large-scale structure (LSS) in the distribution of galaxies and the underlying dark matter and for the cosmic microwave background (CMB) temperature anisotropies that we observe today. In fact inflation has become the dominant paradigm to understand the initial conditions for structure formation and CMB anisotropies. In the inflationary picture, primordial density and gravity-wave fluctuations are created from quantum fluctuations "redshifted" out of the horizon during an early period of superluminal expansion of the Universe, where they are "frozen" [206,102,107,163,273,22]. Perturbations at the surface of last scattering are observable as temperature anisotropy in the CMB, which was first detected by the Cosmic Background Explorer (COBE) satellite [266,36,94]. The last and most impressive confirmation of the inflationary paradigm has been recently provided by the data of the Wilkinson Microwave Anisotropy Probe (WMAP) mission [38]. The WMAP collaboration has produced a full-sky map of the angular variations of the CMB, with unprecedented accuracy. WMAP data confirm the inflationary mechanism as responsible for the generation of curvature (adiabatic) superhorizon fluctuations [225].

Since the primordial cosmological perturbations are tiny, the generation and evolution of fluctuations during inflation has been studied within linear perturbation theory. Within this approach, the primordial density perturbation is Gaussian; in other words, its Fourier components are uncorrelated and have random phases. Despite the simplicity of the inflationary paradigm, the mechanism by which cosmological adiabatic perturbations are generated is not yet established. In the standard slow-roll scenario associated to one-single field models of inflation, the observed density perturbations are due to fluctuations of the inflaton field itself when it slowly rolls down along its potential. When inflation ends, the inflaton φ oscillates about the minimum of its potential $V(\varphi)$ and decays, thereby reheating the Universe. As a result of the fluctuations each region of the Universe goes through the same history but at slightly different times. The final temperature anisotropies are caused by inflation lasting for different amounts of time in different regions of the Universe leading to adiabatic perturbations. Under this hypothesis, the WMAP data set already allows to extract the parameters relevant for distinguishing among single-field inflation models [225,123].

An alternative to the standard scenario is represented by the curvaton mechanism [195,79,176,201,175] where the final curvature perturbations are produced from an initial isocurvature perturbation associated with the quantum fluctuations of a light scalar field (other than the inflaton), the curvaton, whose energy density is negligible during inflation. The curvaton isocurvature perturbations are transformed into adiabatic ones when the curvaton decays into radiation much after the end of inflation.

Recently, other mechanisms for the generation of cosmological perturbations have been proposed, the inhomogeneous reheating scenario [76,130,77,192,10], the ghost inflationary scenario [14] and the D-cceleration scenario [264], just to mention a few. For instance, the inhomogeneous reheating scenario acts during the reheating stage after inflation if superhorizon spatial fluctuations in the decay rate of the inflaton field are induced during inflation, causing adiabatic perturbations in the final reheating temperature in different regions of the Universe.

The generation of gravity-wave fluctuations is a generic prediction of an accelerated de Sitter expansion of the Universe whatever mechanism for the generation of cosmological perturbations is operative. Gravitational waves, whose possible observation might come from the detection of the *B*-mode of polarization in the CMB anisotropy [121,261], may be viewed as ripples of space–time around the background metric.

Since curvature fluctuations are (nearly) frozen on superhorizon scales, a way of characterizing them is to compute their spectrum on scales larger than the horizon. In the standard slow-roll inflationary models where the fluctuations of the inflaton field φ are responsible for the curvature perturbations, the power-spectrum $\mathscr{P}_{\mathscr{R}}$ of the comoving curvature perturbation $\mathscr{R}$ (which is a measure of the spatial curvature as seen by comoving observers) is given by

$$\mathscr{P}_{\mathscr{R}}(k) = \frac{1}{2M_{\mathrm{P}}^2\epsilon}\left(\frac{H_*}{2\pi}\right)^2\left(\frac{k}{aH_*}\right)^{n_{\mathscr{R}}-1}, \tag{1}$$

where $n_{\mathscr{R}} = 1 - 6\epsilon + 2\eta \simeq 1$ is the spectral index, $M_{\mathrm{P}} \equiv (8\pi G_{\mathrm{N}})^{-1/2} \simeq 2.4 \times 10^{18}$ GeV is the reduced Planck scale. Here

$$\epsilon = \frac{M_{\mathrm{P}}^2}{2}\left(\frac{V'}{V}\right)^2, \qquad \eta = M_{\mathrm{P}}^2\left(\frac{V''}{V}\right) \tag{2}$$

are the so-called slow-roll parameters ($\epsilon, \eta \ll 1$ during inflation), $H_* = \dot{a}/a$ indicates the Hubble rate during inflation and primes here denote derivatives with respect to φ. The WMAP has determined the amplitude of the power-spectrum as $\mathscr{P}_{\mathscr{R}}(k) \simeq 2.95 \times 10^{-9} A$, where $A = 0.6 - 1$ depending on the model under consideration [225,269], which implies that

$$\frac{1}{2M_{\rm P}^2\epsilon}\left(\frac{H_*}{2\pi}\right)^2 \simeq (2-3)\times 10^{-9}\ , \tag{3}$$

or

$$H_* \simeq (0.9-1.2)\times 10^{15}\epsilon^{1/2}\,\text{GeV}. \tag{4}$$

The Friedmann equation in the slow-roll limit, $H^2 = V/(3M_{\rm P}^2)$, then gives "the energy scale of inflation",

$$V^{1/4} \simeq (6.3-7.1)\times 10^{16}\epsilon^{1/4}\,\text{GeV}\ . \tag{5}$$

On the other hand, the power-spectrum of gravity-wave modes h_{ij} is given by

$$\mathscr{P}_T(k) = \frac{k^3}{2\pi^2}\langle h_{ij}^* h^{ij}\rangle = \frac{8}{M_{\rm P}^2}\left(\frac{H_*}{2\pi}\right)^2\left(\frac{k}{aH_*}\right)^{n_T}\ , \tag{6}$$

where $n_T = -2\epsilon$ is the tensor spectral index. Since the fractional change of the power-spectra with scale is much smaller than unity, one can safely consider the power-spectra as being roughly constant on the scales relevant for the CMB anisotropy and define a tensor-to-scalar amplitude ratio

$$r = \frac{\mathscr{P}_T}{\mathscr{P}_{\mathscr{R}}} = 16\epsilon\ . \tag{7}$$

The spectra $\mathscr{P}_{\mathscr{R}}(k)$ and $\mathscr{P}_T(k)$ provide the contact between theory and observation. The present WMAP data set allows to extract an upper bound, $r < 1.28$ (95%) [225,123], or $\epsilon < 0.08$. This limit together with Eq. (5) provides an upper bound on the energy scale of inflation,

$$V^{1/4} < 3.8\times 10^{16}\,\text{GeV}\ . \tag{8}$$

The corresponding upper bound on the Hubble rate during inflation is $H_* < 3.4\times 10^{14}$ GeV. A positive detection of the *B*-mode in CMB polarization, and therefore an indirect evidence of gravitational waves from inflation, once foregrounds due to gravitational lensing from local sources have been properly treated, requires $\epsilon > 10^{-5}$ corresponding to $V^{1/4} > 3.5\times 10^{15}$ GeV and $H_* > 3\times 10^{12}$ GeV [122,126,259].[1]

However, *what if* the curvature perturbation is generated through the quantum fluctuations of a scalar field other than the inflaton? Then, what is the expected amplitude of gravity-wave fluctuations in such scenarios? Consider, for instance, the curvaton scenario and the inhomogeneous reheating scenario. They liberate the inflaton from the responsibility of generating the cosmological curvature perturbation and therefore avoid slow-roll conditions. Their basic assumption is that the initial curvature perturbation due to the inflaton field is negligible. The common lore to achieve such a condition is to assume that the

[1] If "cleaning" of the gravitational lensing effect can be achieved down to the level envisaged in Ref. [259], then another source of *B*-mode polarization will limit our ability to detect the signature of primordial gravitational waves. This comes from vector and tensor modes arising from the second-order evolution of scalar perturbations [197] and represents the ultimate barrier to gravitational-wave detection if $\epsilon < 10^{-7}$.

energy scale of the inflaton potential is too small to match the observed amplitude of CMB anisotropy, that is $V^{1/4} \ll 10^{16}$ GeV. Therefore—while certainly useful to construct low-scale models of inflation—it is usually thought that these mechanisms predict an amplitude of gravitational waves which is far too small to be detectable by future satellite experiments aimed at observing the *B*-mode of the CMB polarization (see however Ref. [229]). This implies that a future detection of the *B*-mode of the CMB polarization would favour the slow-roll models of inflation as generators of the cosmological perturbations. On the other hand, the lack of signal of gravity waves in the CMB anisotropies will not give us any information about the mechanism by which cosmological perturbations are created.

A precise measurement of the spectral index of comoving curvature perturbations will be a powerful tool to constrain inflationary models. Slow-roll inflation predicts $|n_{\mathscr{R}} - 1|$ significantly below 1. Deviations of $n_{\mathscr{R}}$ from unity are generically (but not always) proportional to $1/N$, where N is the number of e-folds till the end of inflation. The predictions of different models for the spectral index $n_{\mathscr{R}}$, and for its scale-dependence, are well summarised in the review [174] within slow-roll inflationary models. Remarkably, the eventual accuracy $\Delta n_{\mathscr{R}} \sim 0.01$ offered by the *Planck* satellite[2] is just what one might have specified in order to distinguish between various slow-roll models of inflation. Observation will discriminate strongly between slow-roll models of inflation in next 10 or 15 years. If cosmological perturbations are due to the inflaton field, then in 10 or 15 years there may be a consensus about the form of the inflationary potential, and at a deeper level we may have learned something valuable about the nature of the fundamental interactions beyond the Standard Model. However, *what if* Nature has chosen the other mechanisms for the creation of the cosmological perturbations, which generically predict a value of $n_{\mathscr{R}}$ very close to unity with a negligible scale dependence? Then, it implies that a precise measurement of the spectral index will not allow us to efficiently discriminate among different scenarios.

These "*what if*" options would be discouraging if they turn out to be true. They would imply that all future efforts for measuring tensor modes in the CMB anisotropy and the spectral index of adiabatic perturbations are of no use to disentangle the various scenarios for the creation of the cosmological perturbations.

There is, however, a third observable which will prove fundamental in providing information about the mechanism chosen by Nature to produce the structures we see today. It is the deviation from a pure Gaussian statistics, i.e., the presence of higher-order connected correlation functions of CMB anisotropies. The angular n-point correlation function

$$\langle f(\hat{\mathbf{n}}_1) f(\hat{\mathbf{n}}_2) \dots f(\hat{\mathbf{n}}_n) \rangle \ , \tag{9}$$

is a simple statistic characterizing a clustering pattern of fluctuations on the sky, $f(\hat{\mathbf{n}})$. The bracket denotes the ensemble average, and Fig. 1 sketches its meaning. If the fluctuation is Gaussian, then the two-point correlation function specifies all the statistical properties of $f(\hat{\mathbf{n}})$, for the two-point correlation function is the only parameter in a Gaussian distribution. If it is not Gaussian, then we need higher-order correlation functions to determine the statistical properties.

For instance, a non-vanishing three-point function of scalar perturbations, or its Fourier transform, the bispectrum, is an indicator of a non-Gaussian feature in the cosmological perturbations. The importance of the bispectrum comes from the fact that it represents the lowest order statistics able to distinguish non-Gaussian from Gaussian perturbations. An accurate calculation of the primordial bispectrum of cosmological perturbations has become an extremely important issue, as a number of present and future

[2] See, for instance, http://www.rssd.esa.int/index.php?project = PLANCK.

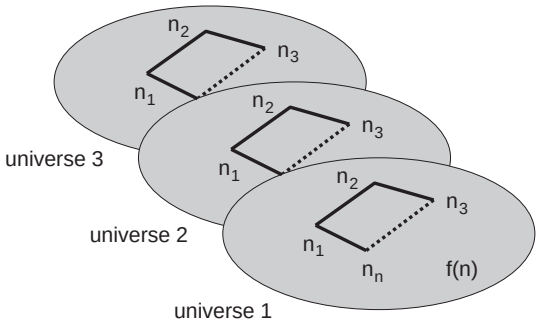

Fig. 1. Ensemble average of angular correlation function. A schematic view of the ensemble average of the *n*-point angular correlation function, $f(\hat{\mathbf{n}}_1)f(\hat{\mathbf{n}}_2)f(\hat{\mathbf{n}}_3)\ldots f(\hat{\mathbf{n}}_n)$. We measure it on each Universe, and then average it over many Universes.

experiments, such as WMAP and Planck, will allow to constrain or detect non-Gaussianity of CMB anisotropy with high precision. A phenomenological way of parametrizing the level of non-Gaussianity in the cosmological perturbations is to introduce a non-linearity parameter $f_{\rm NL}$ through Bardeen's gravitational potential[3]

$$\Phi = \Phi_{\rm L} + f_{\rm NL} \star (\Phi_{\rm L})^2 \ , \tag{10}$$

where $\Phi_{\rm L}$ represents the gravitational potential at linear order and the $\star$—products reminds the fact that the non-linearity parameter might have a non-trivial scale dependence. As Eq. (10) shows, in order to compute and keep track of the non-Gaussianity of the cosmological perturbations throughout the different stages of the evolution of the Universe, one has to perform a perturbation around the homogeneous background *up to second order*.

The non-Gaussianity in the primordial cosmological perturbations and, in particular, theoretical and observational determinations of the non-linearity parameter, $f_{\rm NL}$, are the subject of this review. Our goals are to present a thorough, detailed and updated review of the state-of-the-art on the subject of non-Gaussianity, both from the theoretical and observational point of view, and to provide the reader with all the tools necessary to compute the level of non-Gaussianity in any model of cosmological perturbations.

Surprisingly, despite the importance of the subject of non-Gaussianity in the cosmological perturbations and despite the fact that its detection is one of the primary goals of the present and future satellite missions such as WMAP and Planck, not much attention has been devoted to this issue on the theoretical side. For instance, no firm theoretical predictions were available till about two years ago for the case of slow-roll models of inflation. Spurred by the large amount of data available in the next future, a new wave of models, which are firmly rooted in modern particle theory, have been recently proposed to generate a large and detectable amount of non-Gaussianity from inflation.

Probably, one of the reasons why the theoretical investigations of the primordial non-Gaussianity were so limited was because no direct observational constraints on $f_{\rm NL}$ were available until the year 2001. Many authors have shown that CMB temperature anisotropy is consistent with Gaussianity since the very first detection of anisotropy in the COBE DMR data (see Ref. [137] and references therein); however, very

[3] Non-Gaussian models containing quadratic non-linearities, as in Eq. (10), were introduced in the study of inflationary perturbations in Refs. [112,131,82], and have become a sort of "standard lore" for the comparison of theoretical predictions on primordial non-Gaussianity to CMB and LSS observational data [64,204,292,254,203,291,285,193,284,143,17,256].

little attention has been paid to put *quantitative* constraints on degrees to which the data are consistent with Gaussianity. Since non-Gaussian fluctuations have infinite degrees of freedom as opposed to Gaussian fluctuations, testing a Gaussian hypothesis is a very difficult task; one statistical method showing CMB *consistent* with Gaussianity does not mean that CMB is *really* Gaussian. If one does not have any specific, physically motivated non-Gaussian models to constrain (such as those described above), then one cannot learn anything about the nature of temperature fluctuations from the statement that just says, "the CMB is consistent with Gaussianity". Rather, "*How Gaussian is it?*" is a more relevant question, when we try to constrain (and exclude) certain non-Gaussian models.

The first direct comparison between the inflationary non-Gaussianity and observational data was attempted for the COBE DMR data in 2001, using the angular bispectrum, the harmonic counterpart of the three-point correlation function [145]. A very weak constraint, $|f_{\rm NL}| < 1500$ (68%) was found. Although this constraint is still too weak to be useful, it explicitly demonstrated that measurements of non-Gaussianity can put quantitative constraints on inflationary models. The angular bispectrum of the CMB is particularly useful in finding a limit on $f_{\rm NL}$, as the exact analytical calculation of the bispectrum from inflationary non-Gaussianity is possible [143]. While the COBE DMR data constrain non-Gaussianity on large scales ($\sim 7^\circ$), in Ref. [252] a constraint is obtained on small scales ($\sim 10'$), $|f_{\rm NL}| < 950$ (68%), using the MAXIMA data. A recent analysis of the bispectrum of the VSA data [265] gives an upper bound of 5400 on the value of $|f_{\rm NL}|$ (95%). The WMAP team has measured the bispectrum to obtain the tightest limit to date, $-58 < f_{\rm NL} < 134$ (95%) [139].

What about other statistical tools? Currently, analytical predictions exist only for the bispectrum [143] and the trispectrum (the harmonic counterpart of the four-point function) [215]. Predictions for other tools, e.g., Minkowski functionals, are usually much more difficult; however, one can still use other statistical tools to constrain $f_{\rm NL}$ by using a Monte Carlo method: direct comparison between measurements on the observed sky maps and those on simulated non-Gaussian sky maps. The authors of Ref. [59] have measured the spherical Mexican-hat wavelets on the COBE DMR data, and compared them to simulated measurements on non-Gaussian maps (which include only the Sachs–Wolfe [247] effect), finding $|f_{\rm NL}| < 1100$ (68%). This methodology can be applied to any other statistics, if we have accurate simulations taking into account not only the Sachs–Wolfe effect but also the full effect of the Integrated Sachs–Wolfe effect and baryon–photon fluid dynamics. The WMAP team has simulated such full-sky non-Gaussian maps that include all the relevant effects. By comparing the Minkowski functionals measured on the WMAP maps and those on the simulated maps, they obtain $f_{\rm NL} < 139$ (95%) [139]. Using the same simulations and WMAP map, the authors of Ref. [209] find $f_{\rm NL} < 220$ (95%) with the spherical Mexican-hat wavelets. Also, in Ref. [58] the local curvature of the CMB on the WMAP map has been measured and compared to non-Gaussian CMB map simulations [160], finding $-180 < f_{\rm NL} < 240$ (95%). Despite these statistical tools being very different and complementary to some extent, they give similar constraints on $f_{\rm NL}$. (Although Ref. [91] finds a much tighter limit on $f_{\rm NL}$, a direct comparison is not straightforward as their definition of $f_{\rm NL}$ differs from ours.) A theoretical study suggests that the inflationary non-Gaussianity can be detected with the bispectrum, if $f_{\rm NL} > 20$ and 5 for the WMAP and Planck data, respectively [143]. The current limits from the WMAP data are weaker than the theoretical expectation, probably because of the current measurements treating the effects of inhomogeneous noise and Galaxy cut sub-optimally. An optimal method for measuring the bispectrum is still very time consuming [252], while other statistics may have a better chance to overcome this issue. For this study, having accurate non-Gaussian simulations is crucial. In Fig. 2, we show some examples of non-Gaussian sky maps at the Planck resolution, simulated by the spherical-coordinates method of Ref. [160].

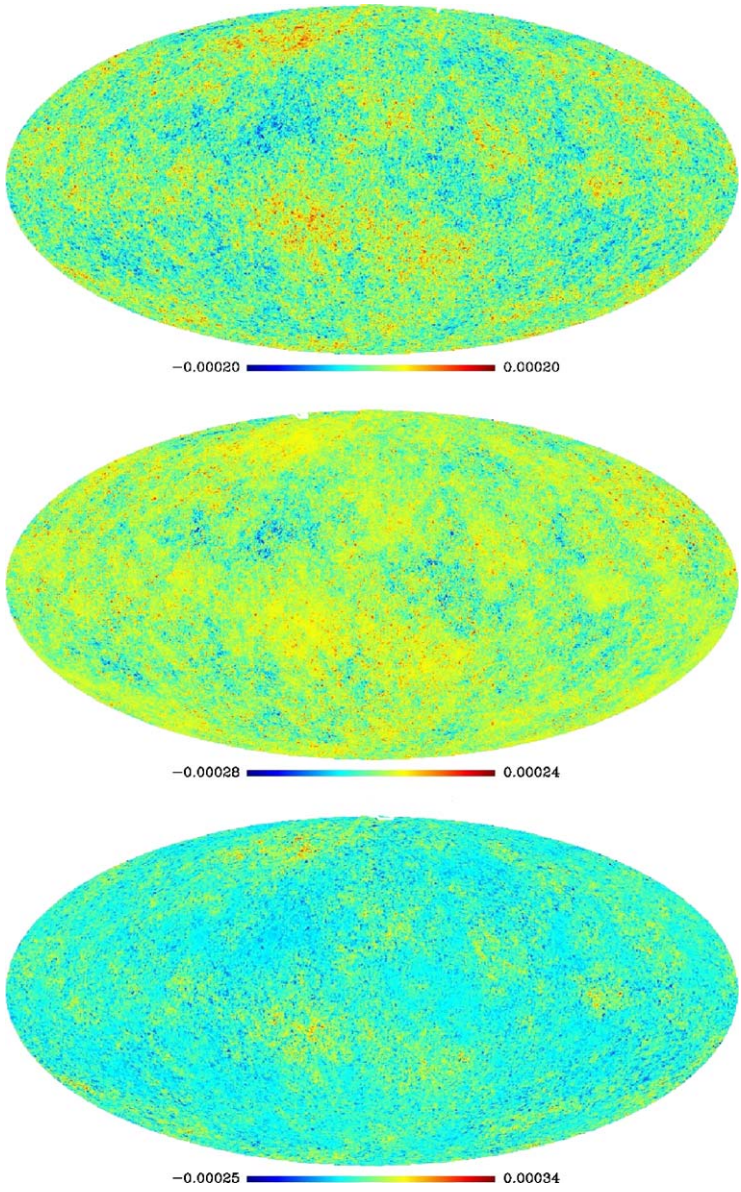

Fig. 2. Planck-resolution simulations of non-Gaussian CMB maps. From top to bottom, the three panels show a Gaussian and two non-Gaussian simulations of CMB maps, at the Planck resolution (FWHM $= 5'$). The non-Gaussian maps are obtained from the model of Eq. (10), using the spherical-coordinates algorithm of Ref. [160], with initial power-spectrum and radiation transfer function of a standard "concordance" (ΛCDM) model. The values of the non-linearity parameter are $f_{\rm NL} = 0$ (top), $f_{\rm NL} = 3000$ (middle) and $f_{\rm NL} = -3000$ (bottom); the high values of $|f_{\rm NL}|$ are chosen to make the non-Gaussian effects visible by eye (note that color-scales are calibrated to the temperature interval of each map).

So far we have talked only about measuring non-Gaussianity from temperature maps. On the other hand, adding polarization information will help to improve our sensitivity to $f_{\rm NL}$, as polarization probes a part of the spectrum of primordial fluctuations that cannot be measured by temperature alone. More specifically,

the polarization radiation transfer function is non-zero at wave numbers k for which the temperature transfer function is zero; thus, polarization contains information which is maximally complementary to temperature [144]. Therefore, one could measure $f_{\rm NL}$ as small as ~ 3 by combining the temperature and polarization bispectra. In addition, if we combine the bispectrum with other statistics, then sensitivity would further improve, depending on the extent to which those statistics are complementary [5]. It is important to keep improving our sensitivity until we reach a critical sensitivity, $f_{\rm NL} \sim 1$, which is set by non-Gaussian contributions from ubiquitous second-order perturbations.

Before concluding this Introduction, let us mention another important source of primordial non-Gaussianity which will not be covered in this review. The topological defects, cosmic strings in particular, are potential sources of strong non-Gaussianity. Although the current observations have ruled out the topological defects as being the primary source of cosmological perturbations, it is still quite possible that topological defects do exist and contribute to a part of the perturbations (see, e.g., Refs. [49,230] for the latest results from the WMAP data). Since the topological defects are intrinsically very non-Gaussian, even a modest energy density of defects may give rise to a detectable level of non-Gaussianity. In particular, the small-scale CMB experiments at an angular scale of $\sim 1'$ have good chance to test (or detect) non-Gaussianity from the cosmic strings, via the so-called Kaiser–Stebbins effect (see Refs. [120,95] for the temperature and Ref. [42] for the polarization). Accurate numerical simulations are needed to search for signatures of topological defects through non-Gaussianity. Only recently, improved simulations of CMB sky maps from cosmic strings have become available by solving the full Boltzmann [149,150]. It is very important to improve the dynamical range of the cosmic strings simulations and make accurate predictions for the CMB sky maps (both in terms of temperature and polarization), which can be compared with future small-scale CMB experiments (recent progress on making a map on small scales has been reported in Ref. [151]).

We end this Introduction with an overview of the present article. The article is addressed to a wide audience, including both cosmologists and particle physicists. To cope with this problem, we have tried to make each section reasonably homogeneous regarding the background knowledge that is taken for granted, while at the same time allowing considerable variation from one section to another.

Section 2 contains a brief review of the inflationary paradigm and an introduction to the theory of quantum fluctuations for a generic scalar field evolving in a fixed de Sitter background. Correlation functions up to order three are evaluated for an interacting scalar field.

Sections 3 and 4 are devoted to the theory of cosmological perturbations at first and second order including gravity. This treatment is done in a gauge-invariant way and the equations up to second order necessary to follow the evolution of non-linearities are provided.

Section 5 deals with the standard slow-roll scenario where cosmological fluctuations are due to the inflaton field. The goal of this section is to show that the main contribution to the non-Gaussian signal comes from the post-inflationary evolution.

Sections 6 and 7 are devoted to the non-Gaussianity predicted in the curvaton and the inhomogeneous reheating scenarios, respectively.

Section 8 contains all the necessary tools to relate the level of non-Gaussianity parametrized by $f_{\rm NL}$ deduced from the true measurements to the one predicted theoretically within a given model.

Section 9 contains a mini-review of alternative models of inflation with the respective predictions of non-Gaussianity.

All the results of the previous sections are summarized in Table 1.

Table 1
Predictions of the non-linearity parameter $f_{\rm NL}$ from different scenarios for the generation of cosmological perturbations

	$f_{\rm NL}(\mathbf{k}_1,\mathbf{k}_2)$	Comments
Single-field inflation	$\frac{7}{3}-g(\mathbf{k}_1,\mathbf{k}_2)$	$g(\mathbf{k}_1,\mathbf{k}_2)=4\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k^2}-3\frac{(\mathbf{k}_1\cdot\mathbf{k}_2)^2}{k^4}+\frac{3}{2}\frac{k_1^4+k_2^4}{k^4}$
Curvaton scenario	$-[-\frac{2}{3}+\frac{5}{6}r-\frac{5}{4r}]-g(\mathbf{k}_1,\mathbf{k}_2)$	$r\approx(\frac{\rho_\sigma}{\rho})_{\rm decay}$
Inhomogeneous reheating	$\frac{13}{12}-I-g(\mathbf{k}_1,\mathbf{k}_2)$	$I=-\frac{5}{2}+\frac{5}{12}\frac{\bar{\Gamma}}{\alpha\Gamma_1}$ "Minimal case" $I=0$ ($\alpha=\frac{1}{6}$, $\Gamma_1=\bar{\Gamma}$)
Multiple scalar fields	$\lesssim\frac{\mathcal{P}_\mathcal{S}}{\mathcal{P}_\mathcal{R}}\cos^2\Delta(4\cdot10^3\cdot\frac{V_{\chi\chi}}{3H^2})\cdot60\frac{H}{\chi}$	Order of magnitude estimate of the absolute value
"Unconventional" inflation set-ups		
Warm inflation	$-\frac{5}{6}(\frac{\dot{\varphi}_0}{H^2})[\ln(\frac{\Gamma}{H})\frac{V'''}{\Gamma}]$	Second-order corrections not included
Ghost inflation	$-140\cdot\beta\cdot\alpha^{-8/5}$	Post-inflationary corrections not included
D-cceleration	$-10^{-1}\gamma^2$	Post-inflationary corrections not included

In the inhomogeneous reheating scenario $\Gamma_1/\bar{\Gamma}\leqslant1$ and $0<\alpha\leqslant\frac{1}{6}$. In the multiple-field case $-1\leqslant\cos\Delta\leqslant1$ is defined in Eq. (373) measuring the correlation between the adiabatic and entropy perturbations, while $\mathcal{P}_\mathcal{S}/\mathcal{P}_\mathcal{R}$ is the isocurvature fraction. In ghost inflation the coefficients α and β are typically $\sim\mathcal{O}(1)$. In the D-cceleration mechanism of inflation the coefficient γ is expected to be $\gamma>1$. For the multiple-field case and the unconventional inflation set-ups, the estimates can receive relevant corrections in the range $f_{\rm NL}\sim1$ from the post-inflationary evolution of the perturbations.

Section 10 contains a mini-review of the present observational constraints on non-linearities in the cosmological perturbations and a thorough discussion of the future prospects on the detectability of non-Gaussianity. Our conclusions are drawn in Section 11. Finally we provide the reader with three appendices where she/he could find the full derivation of second-order geometric quantities and Einstein and Klein–Gordon equations as well as the Wigner-3j symbols.

2. The inflationary paradigm

As we have mentioned in the Introduction, one of the relevant ideas of modern cosmology is represented by the inflationary paradigm. Here we just summarize some of the basics of inflation. For more details the reader is referred to some reviews on the subject [167,156,159,174,157,207,243].

As far as the dynamics of Inflation is concerned one can consider a homogeneous and isotropic Universe described by the Friedmann–Robertson–Walker (FRW) metric

$$\mathrm{d}s^2=-\mathrm{d}t^2+a^2(t)\left[\frac{\mathrm{d}r^2}{1-Kr^2}+r^2(\mathrm{d}\theta^2+\sin^2\theta\,\mathrm{d}\phi^2)\right], \tag{11}$$

where t is the cosmic time, r, θ, ϕ are the comoving (polar) coordinates, $a(t)$ is the scale-factor of the Universe, and K is the curvature constant of three-dimensional hypersurfaces. If the Universe is filled

with matter described by the energy–momentum tensor $T_{\mu\nu}$ of a perfect fluid with energy density ρ and pressure P, the Einstein equations

$$G_{\mu\nu} = 8\pi G_{\rm N} T_{\mu\nu} \ , \tag{12}$$

with $G_{\mu\nu}$ the Einstein tensor and $G_{\rm N}$ the Newtonian gravitational constant give the Friedmann equations

$$H^2 = \frac{8\pi G_{\rm N}}{3}\,\rho - \frac{K}{a^2} \ , \tag{13}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_{\rm N}}{3}(\rho + 3P) \ , \tag{14}$$

where $H = \dot{a}/a$ is the Hubble expansion parameter and dots denote differentiation with respect to cosmic time t. Eq. (14) shows that a period of inflation is possible if the pressure P is negative with

$$P < -\frac{\rho}{3} \ . \tag{15}$$

In particular a period of the history of Universe during which $P = -\rho$ is called a *de Sitter stage*. From the energy continuity equation $\dot{\rho} + 3H(\rho + P) = 0$ and Eq. (13) (neglecting the curvature K which is soon redshifted away as a^{-2}) we see that in a de Sitter phase $\rho =$ constant and

$$H = H_I = \text{constant} \ . \tag{16}$$

Solving Eq. (14) we also see the scale-factor grows exponentially

$$a(t) = a_i \mathrm{e}^{H_I(t-t_i)} \ , \tag{17}$$

where t_i is the time inflation starts. In fact condition (15) can be satisfied by a scalar field, the inflaton φ.

The action for a minimally coupled scalar field φ is given by

$$S = \int \mathrm{d}^4x\sqrt{-g}\,\mathscr{L} = \int \mathrm{d}^4x\sqrt{-g}\left[-\frac{1}{2}\,g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi)\right] \ , \tag{18}$$

where g is the determinant of the metric tensor $g_{\mu\nu}$, $g_{\mu\nu}$ is the contravariant metric tensor, such that $g_{\mu\nu}g^{\nu\lambda} = \delta^\lambda_\mu$; finally $V(\varphi)$ specifies the scalar field potential. By varying the action with respect to φ one obtains the Klein–Gordon equation

$$\Box\varphi = \frac{\partial V}{\partial\varphi} \ , \tag{19}$$

where $\Box$ is the covariant D'Alembert operator

$$\Box\varphi = \frac{1}{\sqrt{-g}}\,\partial_\nu(\sqrt{-g}\,g^{\mu\nu}\partial_\mu\varphi) \ . \tag{20}$$

In a FRW Universe described by metric (11), the evolution equation for φ becomes

$$\ddot{\varphi} + 3H\dot{\varphi} - \frac{\nabla^2\varphi}{a^2} + V'(\varphi) = 0 \ , \tag{21}$$

where $V'(\varphi) = (\mathrm{d}V(\varphi)/\mathrm{d}\varphi)$. Note, in particular, the appearance of the friction term $3H\dot{\varphi}$: a scalar field rolling down its potential suffers a friction due to the expansion of the Universe. The energy–momentum tensor for a minimally-coupled scalar field φ is given by

$$T_{\mu\nu} = -2\frac{\partial\mathscr{L}}{\partial g^{\mu\nu}} + g_{\mu\nu}\mathscr{L} = \partial_\mu\varphi\partial_\nu\varphi + g_{\mu\nu}\left[-\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V(\varphi)\right] . \tag{22}$$

We can now split the inflaton field as

$$\varphi(t,\mathbf{x}) = \varphi_0(t) + \delta\varphi(t,\mathbf{x}),$$

where φ_0 is the 'classical' (infinite wavelength) field, that is the expectation value of the inflaton field on the initial isotropic and homogeneous state, while $\delta\varphi(t,\mathbf{x})$ represents the quantum fluctuations around φ_0. In this section, we will be only concerned with the evolution of the classical field φ_0. The next section will be devoted to the crucial issue of the evolution of quantum perturbations during inflation. This separation is justified by the fact that quantum fluctuations are much smaller than the classical value and therefore negligible when looking at the classical evolution. To not be overwhelmed by the notation, we will keep indicating from now on the classical value of the inflaton field by φ. A homogeneous scalar field $\varphi(t)$ behaves like a perfect fluid with background energy density and pressure given by

$$\rho_\varphi = \frac{\dot{\varphi}^2}{2} + V(\varphi) , \tag{23}$$

$$P_\varphi = \frac{\dot{\varphi}^2}{2} - V(\varphi) . \tag{24}$$

Therefore if

$$V(\varphi) \gg \dot{\varphi}^2 , \tag{25}$$

we obtain the following condition:

$$P_\varphi \simeq -\rho_\varphi . \tag{26}$$

From this simple calculation, we realize that a scalar field whose energy is dominant in the Universe and whose potential energy dominates over the kinetic term gives inflation. Inflation is driven by the vacuum energy of the inflaton field. Notice that the ordinary matter fields, in the form of a radiation fluid, and the spatial curvature K are usually neglected during inflation because their contribution to the energy density is redshifted away during the accelerated expansion.[4] Moreover the basic picture we have discussed here refers only to the simplest models of inflation, where only a single scalar field is present. We will consider later on also some non-standard models of inflation involving more than one scalar field (multiple-field inflation).

[4] For the very same reason also any small inhomogeneities are wiped out as soon as inflation sets in, thus justifying the use of the background FRW metric.

2.1. The slow-rolling inflaton field

Let us now better quantify under which circumstances a scalar field may give rise to a period of inflation. The equation of motion of the homogeneous scalar field is

$$\ddot{\varphi} + 3H\dot{\varphi} + V'(\varphi) = 0 \ . \tag{27}$$

If we require that $\dot{\varphi}^2 \ll V(\varphi)$, the scalar field slowly rolls down its potential. Such a *slow-roll* period can be achieved if the inflaton field φ is in a region where the potential is sufficiently flat. We may also expect that—being the potential flat—$\ddot{\varphi}$ is negligible as well. We will assume that this is true and we will quantify this condition soon. The Friedmann equation (13) becomes

$$H^2 \simeq \frac{8\pi G_{\rm N}}{3} V(\varphi) \ , \tag{28}$$

where we have assumed that the inflaton field dominates the energy density of the Universe. The new equation of motion becomes

$$3H\dot{\varphi} = -V'(\varphi) \ , \tag{29}$$

which gives $\dot{\varphi}$ as a function of $V'(\varphi)$. Using Eq. (29) the slow-roll conditions then require

$$\dot{\varphi}^2 \ll V(\varphi) \Longrightarrow \frac{(V')^2}{V} \ll H^2 \tag{30}$$

and

$$\ddot{\varphi} \ll 3H\dot{\varphi} \Longrightarrow V'' \ll H^2. \tag{31}$$

Eqs. (30) and (31) represent the flatness conditions on the potential which are conveniently parametrized in terms of the so-called *slow-roll parameters*, which are built from V and its derivatives V', V'', V''', $V^{(n)}$, with respect to φ [159,158,174]. In particular, one can define the two slow-roll parameters [159,158,174] in Eq. (2).

Achieving a successful period of inflation requires the slow-roll parameters to be $\epsilon, |\eta| \ll 1$. Indeed there exists a hierarchy of slow-roll parameters [158]. For example one can define the slow-roll parameter related to the third-derivative of the potential $\xi^2 = 1/(8\pi G_{\rm N})\,(V^{(1)}V^{(3)}/V^2)$ which is a second-order slow-roll parameter (note that ξ^2 can be negative). The parameter ϵ can also be written as $\epsilon = -\dot{H}/H^2$, thus it quantifies how much the Hubble rate H changes with time during inflation. In particular notice that, since

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 = (1-\epsilon)H^2 \ ,$$

inflation can be attained only if $\epsilon < 1$. As soon as this condition fails, inflation ends. At first-order in the slow-roll parameters ϵ and η can be considered constant, since the potential is very flat. In fact it is easy to see that that $\dot{\epsilon}, \dot{\eta} = \mathcal{O}(\epsilon^2, \eta^2)$.[5]

[5] With $\mathcal{O}(\epsilon, \eta)$ and $\mathcal{O}(\epsilon^2, \eta^2)$ we indicate general combinations of the slow-roll parameters of lowest order and next order, respectively.

Despite the simplicity of the inflationary paradigm, the number of inflationary models that have been proposed so far is enormous, differing for the kind of potential and for the underlying particle physics theory. In that respect the reader is referred to the review [174]. We just want to mention here that a useful classification in connection with the observations may be the one in which the single-field inflationary models are divided into three broad groups as "small field", "large field" (or chaotic) and "hybrid" type, according to the region occupied in the $(\epsilon$–$\eta)$ space by a given inflationary potential [74]. Typical examples of the large-field models $(0 < \eta < 2\epsilon)$ are polynomial potentials $V(\varphi) = \Lambda^4(\varphi/\mu)^p$, and exponential potentials, $V(\varphi) = \Lambda^4 \exp(\varphi/\mu)$. The small-field potentials $(\eta < -\epsilon)$ are typically of the form $V(\varphi) = \Lambda^4[1-(\varphi/\mu)^p]$, while generic hybrid potentials $(0 < 2\epsilon < \eta)$ are of the form $V(\varphi) = \Lambda^4[1+(\varphi/\mu)^p]$. In fact according to such a scheme, the WMAP data set already allows to extract the parameters relevant for distinguishing among single-field inflation models [225,23,123,154].

Here we want to make an important comment. A crucial quantity for the inflationary dynamics and for understanding the generation of the primordial perturbations during inflation is the Hubble radius (also called the Hubble horizon size) $R_{\rm H} = H^{-1}$. The Hubble radius represents a characteristic length scale beyond which causal processes cannot operate. A key point is that during inflation the comoving Hubble horizon, $(aH)^{-1}$, decreases in time as the scale-factor, a, grows quasi-exponentially, and the Hubble radius remains almost constant (indeed the decrease of $(aH)^{-1}$ is a consequence of the accelerated expansion, $\ddot{a} > 0$, characterizing inflation). Therefore, a given comoving length scale, L, will become larger than the Hubble radius and *leaves the Hubble horizon*. On the other hand, the comoving Hubble radius increases as $(aH)^{-1} \propto a^{1/2}$ and a during radiation and matter dominated era, respectively.

Previously, we have defined inflation as a period of accelerated expansion of the Universe; however, this is actually not sufficient. A successful inflation must last for a long enough period in order to solve the horizon and flatness problems. By "a long enough period" we mean a period of accelerated expansion of the Universe long enough that a small, smooth patch of size that is smaller than the Hubble radius can grow to encompass *at least* the entire observable Universe. Typically the amount of inflation is measured in terms of the number of e-foldings, defined as

$$N_{\rm TOT} = \int_{t_i}^{t_f} H \, {\rm d}t \ , \tag{32}$$

where t_i and t_f are the time inflation starts and ends, respectively. To explain smoothness of the observable Universe, we impose that the largest scale we observe today, the present horizon H_0^{-1} (~ 4200 Mpc), was reduced during inflation to a value λ_{H_0} at t_i, which is smaller than $H_{\rm I}^{-1}$ during inflation. Then, it follows that we must have $N_{\rm TOT} > N_{\rm min}$, where $N_{\rm min} \approx 60$ is the number of e-foldings before the end of inflation when the present Hubble radius leaves the horizon. A very useful quantity is the number of e-foldings from the time when a given wavelength λ leaves the horizon during inflation to the end of inflation,

$$N_\lambda = \int_{t(\lambda)}^{t_f} H \, {\rm d}t = \ln\left(\frac{a_f}{a_\lambda}\right) \ , \tag{33}$$

where $t(\lambda)$ is the time when λ leaves the horizon during inflation and $a_\lambda = a(t(\lambda))$. The cosmologically interesting scales probed by the CMB anisotropies correspond to $N_\lambda \simeq 40$–60.

Inflation ends when the inflaton field starts to roll fast along its potential. During this regime $V'' > H^2$ (or $\eta > 1$). The scalar field will reach the minimum of its potential and will start to oscillate around it. By this time any other contribution to the energy density and entropy of the Universe has been redshifted

away by the inflationary expansion. However we know that the Universe must be repopulated by a hot radiation fluid in order for the standard Big–Bang cosmology to set in. This is achieved through a process, called *reheating*, by which the energy of the inflaton field is transferred to radiation during the oscillating phase. In the ordinary scenario of reheating [8,162,1,75] such a transfer corresponds to the decay of the inflaton field into other lighter particles to which it couples through a decay rate Γ_φ. Such a decay damps the inflaton oscillations and when the decay products thermalize and form a thermal background the Universe is finally reheated. Alternatively, reheating may occur through preheating [133].

2.2. Inflation and cosmological perturbations

Besides the background inflationary dynamics, it is of crucial importance to discuss the issue of the evolution of the quantum fluctuations of the inflaton field $\delta\varphi(t, \mathbf{x})$. In the inflationary paradigm associated with these vacuum fluctuations there are primordial energy density perturbations, which survive after inflation and are the origin of all the structures in the Universe. Our current understanding of the origin of structure in the Universe is that once the Universe became matter dominated ($z \sim 3200$) primeval density inhomogeneities ($\delta\rho/\rho \sim 10^{-5}$) were amplified by gravity and grew into the structure we see today [221,66]. The existence of these inhomogeneities was in fact confirmed by the COBE discovery of CMB anisotropies. In this section we just want to summarize in a qualitative way the process by which such "seed" perturbations are generated during inflation. This would also help the reader to better appreciate the alternative mechanisms that have been proposed recently to the inflationary scenario in order to explain the primordial density perturbations.

First of all, in order for structure formation to occur via gravitational instability, there must have been small preexisting fluctuations on relevant physical length scales (say, a galaxy scale $\sim 1\,\mathrm{Mpc}$) which left the Hubble radius in the radiation-dominated and matter-dominated eras. However in the standard Big–Bang model these small perturbations have to be put in by hand, because it is impossible to produce fluctuations on any length scales larger than the horizon size. Inflation is able to provide a mechanism to generate both density perturbations and gravitational waves. As we mentioned in the previous section, a key ingredient of this mechanism is the fact that during inflation the comoving Hubble horizon $(aH)^{-1}$ decreases with time. Consequently, the wavelength of a quantum fluctuation in the scalar field whose potential energy drives inflation soon exceeds the Hubble radius. The quantum fluctuations arise on scales which are much smaller than the comoving Hubble radius $(aH)^{-1}$, which is the scale beyond which causal processes cannot operate. On such small scales one can use the usual flat space–time quantum field theory to describe the scalar field vacuum fluctuations. The inflationary expansion then stretches the wavelength of quantum fluctuations to outside the horizon; thus, gravitational effects become more and more important and amplify the quantum fluctuations, the result being that a net number of scalar field particles are created by the changing cosmological background [206,102,107,163,273]. On large scales the perturbations just follow a classical evolution. Since microscopic physics does not affect the evolution of fluctuations when its wavelength is outside the horizon, the amplitude of fluctuations is "frozen-in" and fixed at some non-zero value $\delta\varphi$ at the horizon crossing, because of a large friction term $3H\dot{\varphi}$ in the equation of motion of the field φ. The amplitude of the fluctuations on super-horizon scales then remains almost unchanged for a very long time, whereas its wavelength grows exponentially. Therefore, the appearance of such frozen fluctuations is equivalent to the appearance of a classical field $\delta\varphi$ that does not vanish after having averaged over some macroscopic interval of time. Moreover, the same mechanism also generates stochastic gravitational waves [272,2].

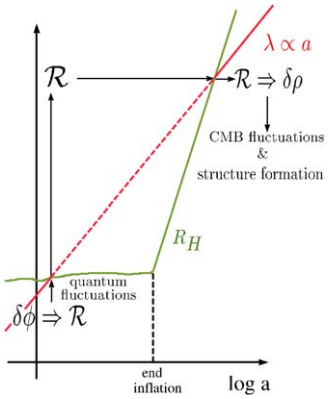

Fig. 3. Stretching of cosmological perturbations during inflation. Quantum perturbations in the curvature, $\mathscr{R}$, are created during inflation and their wavelengths, λ, are stretched from microscopic scales to astronomical scales during inflation.

The fluctuations of the scalar field produce primordial perturbations in the energy density, ρ_φ, which are then inherited by the radiation and matter to which the inflaton field decays during reheating after inflation. Once inflation has ended, however, the Hubble radius increases faster than the scale-factor, so the fluctuations eventually reenter the Hubble radius during the radiation or matter-dominated eras. The fluctuations that exit around 60 e-foldings or so before reheating reenter with physical wavelengths in the range accessible to cosmological observations. These spectra, therefore, preserve signature of inflation, giving us a direct observational connection to physics of inflation. We can measure inflationary fluctuations by different ways, including the analysis of CMB anisotropies. The WMAP collaboration has produced a full-sky map of the angular variations of the CMB, with unprecedented accuracy. The WMAP data confirm the detection of adiabatic *super-horizon fluctuations* which are a distinctive signature of an early epoch of acceleration [225].

The physical inflationary processes which give rise to the structures we observe today are illustrated in Fig. 3.

Quantum fluctuations of the inflaton field are generated during inflation. Since gravity acts on any component of the Universe, small fluctuations of the inflaton field are intimately related to fluctuations of the space–time metric, giving rise to perturbations of the curvature $\mathscr{R}$ (or ζ, which will be defined in the following; the reader may loosely think of them as a gravitational potential). The physical wavelengths λ of these perturbations grow exponentially and leave the horizon when $\lambda > H^{-1}$. On superhorizon scales, curvature fluctuations are frozen in and considered as classical. Finally, when the wavelength of these fluctuations reenters the horizon, at some radiation or matter-dominated epoch, the curvature (gravitational potential) perturbations of the space–time give rise to matter (and temperature) perturbations $\delta\rho$ via the Poisson equation. These fluctuations will then start growing, thus giving rise to the structures we observe today.

In summary, two are the key ingredients for understanding the observed structures in the Universe within the inflationary scenario:

- Quantum fluctuations of the inflaton field are excited during inflation and stretched to cosmological scales. At the same time, as the inflaton fluctuations couple to the metric perturbations through Einstein's equations, ripples on the metric are also excited and stretched to cosmological scales.

- The metric perturbations perturb baryons and photons, and they form acoustic oscillations once the wavelength of the perturbations becomes smaller than the horizon size.

Let us now see how quantum fluctuations are generated during inflation. In fact the mechanism by which the quantum fluctuations of the inflaton field are produced during an inflationary epoch is not peculiar to the inflaton field itself, rather it is generic to any scalar field evolving in an accelerated background. As we shall see, the inflaton field is peculiar in that it dominates the energy density of the Universe, thus possibly producing also metric perturbations.

In the following, we shall describe in a quantitative way how the quantum fluctuations of a generic scalar field evolve during an inflationary stage. For more details we refer the reader to the classical works [165,132,167,159,207] and to some recent reviews on the subject [157,243].

Let us first consider the case of a scalar field χ with an effective potential $V(\chi)$ in a pure de Sitter stage, during which H is constant. Notice that χ is a scalar field different from the inflaton—or the inflatons—that are driving the accelerated expansion.

2.3. Quantum fluctuations of a generic scalar field during a de Sitter stage

Let us first consider the case of a scalar field χ with an effective potential $V(\chi)$ in a pure de Sitter stage, during which H is constant. Notice that χ is a scalar field different from the inflaton—or the inflatons—that are driving the accelerated expansion.

We first split the scalar field $\chi(\tau, \mathbf{x})$ as

$$\chi(\tau, \mathbf{x}) = \chi(\tau) + \delta\chi(\tau, \mathbf{x}) \ , \tag{34}$$

where $\chi(\tau)$ is the homogeneous classical value of the scalar field and $\delta\chi$ are its fluctuations and τ is the conformal time, related to the cosmic time t through $\mathrm{d}\tau = \mathrm{d}t/a(t)$. The scalar field χ is quantized by implementing the standard technique of second quantization. To proceed we first make the following field redefinition:

$$\widetilde{\delta\chi} = a\delta\chi \ . \tag{35}$$

Introducing the creation and annihilation operators $a_{\mathbf{k}}$ and $a^{\dagger}_{\mathbf{k}}$ we promote $\widetilde{\delta\chi}$ to an operator which can be decomposed as

$$\widetilde{\delta\chi}(\tau, \mathbf{x}) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}} [u_k(\tau) a_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}} + u^*_k(\tau) a^{\dagger}_{\mathbf{k}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}] \ . \tag{36}$$

The creation and annihilation operators for $\widetilde{\delta\chi}$ (not for $\delta\chi$) satisfy the standard commutation relations

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}] = 0, \quad [a_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}'}] = \delta^{(3)}(\mathbf{k} - \mathbf{k}') \ , \tag{37}$$

and the modes $u_k(\tau)$ are normalized so that they satisfy the condition

$$u^*_k u'_k - u_k u^{*\prime}_k = -i \ , \tag{38}$$

deriving from the usual canonical commutation relations between the operators $\widetilde{\delta\chi}$ and its conjugate momentum $\Pi = \widetilde{\delta\chi}'$. Here primes denote derivatives with respect to the conformal time τ (not t).

The evolution equation for the scalar field $\chi(\tau, \mathbf{x})$ is given by the Klein–Gordon equation

$$\Box\chi = \frac{\partial V}{\partial \chi} \ , \tag{39}$$

where $\Box$ is the D'Alembert operator defined in Eq. (20). The Klein–Gordon equation gives in an unperturbed FRW Universe

$$\chi'' + 2\mathscr{H}\chi' = -a^2\frac{\partial V}{\partial \chi} \ , \tag{40}$$

where $\mathscr{H} \equiv a'/a$ is the Hubble expansion rate in conformal time. Now, we perturb the scalar field but neglect the metric perturbations in the Klein–Gordon equation (39), the eigenfunctions $u_k(\tau)$ obey the equation of motion

$$u_k'' + \left(k^2 - \frac{a''}{a} + m_\chi^2 a^2\right) u_k = 0 \ , \tag{41}$$

where $m_\chi^2 = \partial^2 V/\partial\chi^2$ is the effective mass of the scalar field. The modes $u_k(\tau)$ at very short distances must reproduce the form for the ordinary flat space–time quantum field theory. Thus, well within the horizon, in the limit $k/aH \to \infty$, the modes should approach plane waves of the form

$$u_k(\tau) \to \frac{1}{\sqrt{2k}}\,\mathrm{e}^{-\mathrm{i}k\tau}. \tag{42}$$

Eq. (41) has an exact solution in the case of a de Sitter stage. Before recovering it, let us study the limiting behavior of Eq. (41) on sub-horizon and superhorizon scales. On sub-horizon scales $k^2 \gg a''/a$, the mass term is negligible so that Eq. (41) reduces to

$$u_k'' + k^2 u_k = 0 \ , \tag{43}$$

whose solution is a plane wave

$$u_k \propto \mathrm{e}^{-\mathrm{i}k\tau} \ . \tag{44}$$

Thus fluctuations with wavelength within the cosmological horizon oscillate as in Eq. (42). This is what we expect since in the ultraviolet limit, i.e. for wavelengths much smaller than the horizon scales, we are approximating the space–time as flat. On superhorizon scales $k^2 \ll a''/a$, Eq. (41) reduces to

$$u_k'' - \left(\frac{a''}{a} - m_\chi^2 a^2\right) u_k = 0 \ . \tag{45}$$

Just for simplicity let us see what happens in the case of a massless scalar field ($m_\chi^2 = 0$). There are two solutions of Eq. (45), a growing and a decaying mode:

$$u_k = B_+(k)a + B_-(k)a^{-2} \ . \tag{46}$$

We can fix the amplitude of the growing mode, B_+, by matching the (absolute value of the) solution (46) to the plane wave solution (42) when the fluctuation with wave number k leaves the horizon ($k = aH$)

$$|B_+(k)| = \frac{1}{a\sqrt{2k}} = \frac{H}{\sqrt{2k^3}} \ , \tag{47}$$

so that the quantum fluctuations of the original scalar field χ on superhorizon scales are constant,

$$|\delta\chi_k| = \frac{|u_k|}{a} = \frac{H}{\sqrt{2k^3}} \ . \tag{48}$$

Now, let us derive the exact solution without any matching tricks. The exact solution to Eq. (41) introduces some corrections due to a non-vanishing mass of the scalar field. In a de Sitter stage, as $a = -(H\tau)^{-1}$

$$\frac{a''}{a} - m_\chi^2 a^2 = \frac{2}{\tau^2}\left(1 - \frac{1}{2}\frac{m_\chi^2}{H^2}\right) \ , \tag{49}$$

so that Eq. (41) can be recast in the form

$$u_k'' + \left(k^2 - \frac{\nu_\chi^2 - \frac{1}{4}}{\tau^2}\right) u_k = 0 \ , \tag{50}$$

where

$$\nu_\chi^2 = \left(\frac{9}{4} - \frac{m_\chi^2}{H^2}\right) \ . \tag{51}$$

When the mass m_χ^2 is constant in time, Eq. (50) is a Bessel equation whose general solution for *real* ν_χ reads

$$u_k(\tau) = \sqrt{-\tau}[c_1(k)\, H_{\nu_\chi}^{(1)}(-k\tau) + c_2(k)\, H_{\nu_\chi}^{(2)}(-k\tau)] \ , \tag{52}$$

where $H_{\nu_\chi}^{(1)}$ and $H_{\nu_\chi}^{(2)}$ are the Hankel functions of first and second kind, respectively. This result actually coincides with the solution found in the work by Bunch and Davies [55] for a free massive scalar field in de Sitter space–time. If we impose that in the ultraviolet regime $k \gg aH$ $(-k\tau \gg 1)$ the solution matches the plane-wave solution $\mathrm{e}^{-\mathrm{i}k\tau}/\sqrt{2k}$ that we expect in flat space–time, and knowing that

$$H_{\nu_\chi}^{(1)}(x \gg 1) \sim \sqrt{\frac{2}{\pi x}}\, \mathrm{e}^{\mathrm{i}(x - \pi/2\, \nu_\chi - \pi/4)}, \quad H_{\nu_\chi}^{(2)}(x \gg 1) \sim \sqrt{\frac{2}{\pi x}}\, \mathrm{e}^{-\mathrm{i}(x - \pi/2\, \nu_\chi - \pi/4)} \ ,$$

we set $c_2(k) = 0$ and $c_1(k) = \sqrt{\pi}/2\mathrm{e}^{\mathrm{i}(\nu_\chi + 1/2)\pi/2}$, which also satisfy the normalization condition (38). The exact solution becomes

$$u_k(\tau) = \frac{\sqrt{\pi}}{2}\, \mathrm{e}^{\mathrm{i}(\nu_\chi + 1/2)\pi/2} \sqrt{-\tau} H_{\nu_\chi}^{(1)}(-k\tau). \tag{53}$$

We are particularly interested in the asymptotic behavior of the solution when the mode is well outside the horizon. On superhorizon scales, since $H_{\nu_\chi}^{(1)}(x \ll 1) \sim \sqrt{2/\pi}\, \mathrm{e}^{-\mathrm{i}\pi/2} 2^{\nu_\chi - 3/2} (\Gamma(\nu_\chi)/\Gamma(3/2)) x^{-\nu_\chi}$, fluctuation (53) becomes

$$u_k(\tau) = \mathrm{e}^{\mathrm{i}(\nu_\chi - 1/2)\pi/2} 2^{(\nu_\chi - 3/2)}\, \frac{\Gamma(\nu_\chi)}{\Gamma(3/2)}\, \frac{1}{\sqrt{2k}} (-k\tau)^{1/2 - \nu_\chi} \ . \tag{54}$$

Thus we find that on superhorizon scales, the fluctuation of the scalar field $\delta\chi_k \equiv u_k/a$ with a non-vanishing mass is not exactly constant, but it acquires a dependence upon time

$$|\delta\chi_k| = 2^{(\nu_\chi-3/2)}\,\frac{\Gamma(\nu_\chi)}{\Gamma(3/2)}\,\frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu_\chi} \quad \text{(on superhorizon scales)}\ . \tag{55}$$

Notice that solution (55) is valid for values of the scalar field mass $m_\chi \leqslant 3/2H$. If the scalar field is very light, $m_\chi \ll 3/2H$, we can introduce the parameter $\eta_\chi = (m_\chi^2/3H^2)$ in analogy with the slow-roll parameters ϵ and η for the inflaton field, and make an expansion of the solution in Eq. (55) to lowest order in $\eta_\chi = (m_\chi^2/3H^2) \ll 1$ to find

$$|\delta\chi_k| = \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu_\chi}\ , \tag{56}$$

with

$$\frac{3}{2} - \nu_\chi \simeq \eta_\chi\ . \tag{57}$$

Eq. (56) shows a crucial result. When the scalar field χ is light, its quantum fluctuations, first generated on subhorizon scales, are gravitationally amplified and stretched to superhorizon scales because of the accelerated expansion of the Universe during inflation [165,167].

2.3.1. The power-spectrum

A useful quantity to characterize the properties of a perturbation field is the *power-spectrum*. For a given random field $f(t,\mathbf{x})$ which can be expanded in Fourier space (since we work in flat space) as[6]

$$f(t,\mathbf{x}) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,f_{\mathbf{k}}(t)\ , \tag{58}$$

the ("dimensionless") power-spectrum $\mathscr{P}_f(k)$ can be defined through

$$\langle f_{\mathbf{k}_1} f^*_{\mathbf{k}_2}\rangle \equiv \frac{2\pi^2}{k^3}\,\mathscr{P}_f(k)\delta^{(3)}(\mathbf{k}_1-\mathbf{k}_2)\,, \tag{59}$$

where the angled brackets denote ensemble averages. The power-spectrum measures the amplitude of the fluctuations at a given scale k; indeed from the definition (59) the mean square value of $f(t,\mathbf{x})$ in real space is

$$\langle f^2(t,\mathbf{x})\rangle = \int \frac{\mathrm{d}k}{k}\,\mathscr{P}_f(k)\ . \tag{60}$$

Thus, according to our definition the power-spectrum, $\mathscr{P}_f(k)$ is the contribution to the variance per unit logarithmic interval in the wave number k. This is standard notation in the literature for the inflationary power-spectrum. However, another definition of power-spectrum, given by the quantity $P_f(k)$ related to $\mathscr{P}_f(k)$ by the relation $P_f(k) = 2\pi^2\mathscr{P}_f(k)/k^3$, or $\langle f_{\mathbf{k}_1} f^*_{\mathbf{k}_2}\rangle = P_f(k)\delta^{(3)}(\mathbf{k}_1-\mathbf{k}_2)$, will be also used in this review.

[6] The alternative Fourier-transform definition $f(t,\mathbf{x}) = \int \mathrm{d}^3\mathbf{k}/(2\pi)^3\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,f_{\mathbf{k}}(t)$ is also used in this review.

To describe the slope of the power-spectrum it is standard practice to define a *spectral index* $n_f(k)$, through

$$n_f(k) - 1 \equiv \frac{\mathrm{d}\,\ln \mathscr{P}_f}{\mathrm{d}\,\ln k}\ . \tag{61}$$

In the case of a scalar field χ the power-spectrum $\mathscr{P}_{\delta\chi}(k)$ can be evaluated by combining Eqs. (35)–(37)

$$\langle \delta\chi_{\mathbf{k_1}} \delta\chi^*_{\mathbf{k_2}} \rangle = \frac{|u_k|^2}{a^2} \delta^{(3)}(\mathbf{k_1} - \mathbf{k_2})\ , \tag{62}$$

yielding

$$\mathscr{P}_{\delta\chi}(k) = \frac{k^3}{2\pi^2} |\delta\chi_k|^2\ , \tag{63}$$

where, as usual, $\delta\chi_k \equiv u_k/a$.

The expression in Eq. (63) is completely general. In the case of a de Sitter phase and a very light scalar field χ, with $m_\chi \ll 3/2H$ we find from Eq. (56) that the power-spectrum on superhorizon scales is given by

$$\mathscr{P}_{\delta\chi}(k) = \left(\frac{H}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\,\nu_\chi}\ , \tag{64}$$

where ν_χ is given by Eq. (57). Thus in this case the dependence on time is tiny, and the spectral index slightly deviates from unity

$$n_{\delta\chi} - 1 = 3 - 2\,\nu_\chi = 2\eta_\chi\ . \tag{65}$$

A useful expression to keep in mind is that of a massless free scalar field in de Sitter space. In this case from Eq. (53) with $\nu_\chi = 3/2$ we obtain

$$\delta\chi_k = (-H\tau)\left(1 - \frac{\mathrm{i}}{k\tau}\right)\frac{\mathrm{e}^{-\mathrm{i}k\tau}}{\sqrt{2k}}\ . \tag{66}$$

The corresponding two-point correlation function for the Fourier modes is

$$\langle \delta\chi(\mathbf{k_1}) \delta^*\chi(\mathbf{k_2}) \rangle = \delta^{(3)}(\mathbf{k_1} - \mathbf{k_2})\,\frac{H^2\tau^2}{2k_1}\left(1 + \frac{1}{k^2\tau^2}\right) \tag{67}$$

$$\approx \delta^{(3)}(\mathbf{k_1} - \mathbf{k_2})\,\frac{H^2}{2k_1^3} \quad (\text{for } k_1\tau \ll 1)\ , \tag{68}$$

with a power-spectrum which, on superhorizon scales, is given by

$$\mathscr{P}_{\delta\chi}(k) = \left(\frac{H}{2\pi}\right)^2\ , \tag{69}$$

which is exactly scale invariant.

We conclude this section with an important remark. Fluctuations of the scalar field can be generated on superhorizon scales as in Eq. (55) only if the scalar field is light. In fact it can be shown that for very

massive scalar fields $m_\chi \gg 3/2H$ (when ν_χ in Eq. (57) becomes imaginary) the fluctuations of the scalar field remain in the vacuum state and do not produce perturbations on cosmologically relevant scales. Indeed, the amplitude of the power-spectrum is damped exponentially as $\mathrm{e}^{-2m_\chi^2/H^2}$ and the spectral index is equal to 4 [229].

2.4. Quantum fluctuations of a generic scalar field in a quasi–de Sitter stage

So far, we have analyzed the time evolution and computed the spectrum of the quantum fluctuations of a generic scalar field χ assuming that the scale-factor evolves like in a pure de Sitter expansion, $a(\tau) = -1/(H\tau)$. However, during Inflation the Hubble rate is not exactly constant, but changes with time as $\dot{H} = -\epsilon H^2$ (quasi–de Sitter expansion). In this subsection, we will solve for the perturbations in a quasi–de Sitter expansion. According to the conclusions of the previous section we consider a scalar field χ with a very small effective mass, $\eta_\chi = (m_\chi^2/3H^2) \ll 1$, and we proceed by making an expansion to lowest order in η_χ and the inflationary parameter $|\epsilon| \ll 1$. Thus from the definition of the conformal time

$$a(\tau) \simeq -\frac{1}{H}\frac{1}{\tau(1-\epsilon)} \ . \tag{70}$$

and

$$\frac{a''}{a} = a^2H^2\left(2 + \frac{\dot{H}}{H^2}\right) \simeq \frac{2}{\tau^2}\left(1 + \frac{3}{2}\epsilon\right) \ . \tag{71}$$

In this way we obtain again the Bessel equation (50) where now ν_χ is given by

$$\nu_\chi \simeq \tfrac{3}{2} + \epsilon - \eta_\chi \ , \tag{72}$$

to lowest order in η_χ and ϵ. Notice that the time derivatives of the slow-roll parameters are next-order in the slow-roll parameters themselves, $\dot{\epsilon}, \dot{\eta} \sim \mathcal{O}(\epsilon^2, \eta^2)$, we can safely treat ν_χ as a constant to our order of approximation. Thus the solution is given by Eq. (53) with the new expression of ν_χ. On large scales and to lowest order in the slow-roll parameters we find

$$|\delta\chi_k| = \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu_\chi} \ . \tag{73}$$

Notice that the quasi–de Sitter expansion yields a correction of order ϵ in comparison with Eq. (56). Since on superhorizon scales from Eq. (73)

$$\delta\chi_{\mathbf{k}} \simeq \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{\eta_\chi-\epsilon} \simeq \frac{H}{\sqrt{2k^3}}\left[1 + (\eta_\chi - \epsilon)\ln\left(\frac{k}{aH}\right)\right], \tag{74}$$

we get

$$|\delta\dot{\chi}_{\mathbf{k}}| \simeq |H\eta_\chi\delta\chi_{\mathbf{k}}| \ll |H\delta\chi_{\mathbf{k}}|, \tag{75}$$

which shows that the fluctuations are (nearly) frozen on superhorizon scales. Therefore, a way to characterize the perturbations is to compute their power-spectrum on scales larger than the horizon, where

one finds

$$\mathscr{P}_{\delta\chi}(k) \simeq \left(\frac{H}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\,\nu_\chi} . \tag{76}$$

Let us conclude this subsection with a comment. Indeed the spectrum of the fluctuations of the scalar field χ in a quasi–de Sitter stage can be also obtained ignoring the variation of the Hubble rate and at the end replace H by its value, H_k, at the time when the fluctuation of wave number k leaves the horizon. The fact that the fluctuations get frozen on superhorizon scales guarantees that we get the exact result. From Eqs. (56) and (57) the power-spectrum obtained with this approach would read

$$\mathscr{P}_{\delta\chi}(k) = \left(\frac{H_k}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\,\nu_\chi} , \tag{77}$$

with $3 - 2\,\nu_\chi \simeq 2\eta_\chi$. In fact by using the relation

$$H \simeq H_k \left[1 + \left.\frac{\dot{H}}{H^2}\right|_{k=aH} \ln\left(\frac{aH}{k}\right)\right] \simeq H_k \left(\frac{k}{aH}\right)^{\epsilon} , \tag{78}$$

we reproduce our previous findings.

2.5. Correlation functions of a self interacting scalar field

The two-point correlation function or its Fourier transform, the power-spectrum, corresponds to the magnitude of a given cosmological perturbation $f(t, \mathbf{x})$. If such a quantity is *Gaussian distributed* then the power-spectrum is all that is needed in order to completely characterize it from a statistical point of view. In fact, in such a case if we consider higher-order correlation functions we find that all the odd correlation functions vanish, while the even correlation functions can be simply expressed in terms of the two-point function. Another way to say that is to introduce the *connected* part of the correlation functions, defined as the part of the expectation value $\langle f(t, \mathbf{x}_1) f(t, \mathbf{x}_2) \cdots f(t, \mathbf{x}_1) f(t, \mathbf{x}_N)\rangle$ which cannot be expressed in terms of expectation values of lower order. For a zero-mean random field, the second and third-order connected correlation functions coincide with the correlation functions themselves, while at fourth-order, for example, one can write

$$\begin{aligned}
&\langle f(t, \mathbf{x}_1) f(t, \mathbf{x}_2) f(t, \mathbf{x}_3) f(t, \mathbf{x}_4)\rangle \\
&\quad = \langle f(t, \mathbf{x}_1) f(t, \mathbf{x}_2)\rangle\langle f(t, \mathbf{x}_3) f(t, \mathbf{x}_4)\rangle + \langle f(t, \mathbf{x}_1) f(t, \mathbf{x}_3)\rangle\langle f(t, \mathbf{x}_2) f(t, \mathbf{x}_4)\rangle \\
&\qquad + \langle f(t, \mathbf{x}_1) f(t, \mathbf{x}_4)\rangle\langle f(t, \mathbf{x}_2) f(t, \mathbf{x}_3)\rangle + \langle f(t, \mathbf{x}_1) f(t, \mathbf{x}_2) f(t, \mathbf{x}_3) f(t, \mathbf{x}_4)\rangle_{\mathrm{c}} ,
\end{aligned} \tag{79}$$

where $\langle\cdot\rangle_{\mathrm{c}}$ denotes the connected part. Following the same hierarchical expansion one can express correlations functions of higher order in a similar manner. For the case of a Gaussian distributed perturbation $f(t, \mathbf{x})$ all the connected parts for $N > 2$ are zero. In particular it follows that the three-point function, or its Fourier transform, the *bispectrum* represents the lowest-order statistics able to distinguish non-Gaussian from Gaussian perturbations. Therefore a large fraction of this review will focus on the study of the bispectrum of the cosmological perturbations produced in different cosmological scenarios.

As an instructive example we start in this section by considering the bispectrum of a scalar field $\chi(t, \mathbf{x})$ during a de Sitter stage. Such a computation can be performed by using the techniques of quantum field theory in curved space–time. In the context of inflationary cosmologies these techniques have been used

in Refs. [11,82]. However, only recently some critical aspects of this approach have been clarified and a systematic formalism has been developed [182] (see also Ref. [46] for a critical investigation on the bases of such a calculation). Therefore we will now summarize how to calculate higher-order correlation functions for a scalar field in a de Sitter space–time following mainly Refs. [182,46].

Higher-order correlation functions are generated as soon as the scalar field has some interaction with itself (or other fields). This amounts to saying that the potential for χ contains some terms beyond the quadratic mass term, so that the interaction part of the potential can be written as

$$V_{\rm int}(\chi) = \frac{V^{(3)}}{3!}(\delta\chi)^3 + \frac{V^{(4)}}{4!}(\delta\chi)^4 + \cdots \ , \tag{80}$$

where $V^{(k)}$ is the kth derivative of the potential. The scalar field is quantized as in Eq. (36) in terms of the eigenmodes $u_k(\tau)$. We want to calculate the correlation function for N-points $\langle\delta\chi(1)\delta\chi(2)\cdots\delta\chi(N)\rangle$. The N-point correlation functions can be in fact expressed perturbatively in terms of those of the free scalar field which have been computed in Section 2.3. To completely take into account the effects of the interaction terms, the underlying idea is that it is necessary to calculate expectation values for the actual vacuum state, that is to say the interacting vacuum state, not just the free vacuum state $|0\rangle$ defined by the requirement that $a_{\mathbf{k}}|0\rangle = 0$ for all $\mathbf{k}$.[7] Such expectation values are defined in the following way by using the interaction picture [182,46]:

$$\langle\widetilde{\delta\chi}_{\mathbf{k}_1}\widetilde{\delta\chi}_{\mathbf{k}_2}\cdots\widetilde{\delta\chi}_{\mathbf{k}_N}\rangle \equiv \langle 0|U^{-1}(\tau_0,\tau)\widetilde{\delta\chi}_{\mathbf{k}_1}\widetilde{\delta\chi}_{\mathbf{k}_2}\cdots\widetilde{\delta\chi}_{\mathbf{k}_N}U(\tau_0,\tau)|0\rangle \ , \tag{81}$$

where $U(\tau_0,\tau)$ is the time evolution operator defined as

$$U(\tau_0,\tau) = \exp\left(-{\rm i}\int_{\tau_0}^{\tau}{\rm d}\tau' H_{\rm int}(\tau')\right) \ . \tag{82}$$

We have already moved to Fourier space where the calculations are easier. Here τ_0 is some early time at which the interactions of the field are supposed to switch on, while $H_{\rm int}=V^{(3)}/3!\,(\delta\chi)^3+V^{(4)}/4!\,(\delta\chi)^3+\cdots$ is the Hamiltonian in the interaction picture. Thus, as it has been pointed out in Ref. [182] the quantity in Eq. (81) doe not correspond to a scattering amplitude, where the initial (at $t\to-\infty$) and the final (at $t\to+\infty$) states are considered as free states. Moreover the expectation value defined in Eq. (81) is free of the critical divergences that occur when calculating the correlation functions on the free vacuum state in a de Sitter space–time, as explained in Ref. [46] (see also the references therein).

To first order in $H_{\rm int}$ the evolution operator can be expanded as

$$U(\tau_0,\tau) = I - {\rm i}\int_{\tau_0}^{\tau}{\rm d}\tau' H_{\rm int}(\tau') \ , \tag{83}$$

and it follows that the connected part of the N-point correlation function is given by

$$\langle\widetilde{\delta\chi}_{\mathbf{k}_1}\widetilde{\delta\chi}_{\mathbf{k}_2}\cdots\widetilde{\delta\chi}_{\mathbf{k}_N}\rangle = -{\rm i}\int_{\tau_0}^{\tau}{\rm d}\tau'\langle 0|[\langle\widetilde{\delta\chi}_{k_1}\widetilde{\delta\chi}_{k_2}\cdots\widetilde{\delta\chi}_{k_N}\rangle, H_{\rm int}(\tau')]0\rangle \ . \tag{84}$$

[7] In Refs. [155,187,89] the non-Gaussian signatures on the CMB arising from inflationary models with non-vacuum initial states for cosmological perturbations have been addressed.

The final result can be expressed in terms of the Green function [46]

$$G(k,\tau,\tau') = \frac{1}{2k}\left(1-\frac{\mathrm{i}}{k\tau}\right)\left(1+\frac{\mathrm{i}}{k\tau'}\right)\exp[\mathrm{i}k(\tau'-\tau)]\ , \tag{85}$$

which can be obtained from the definition $\langle 0|\widetilde{\delta\chi}(\tau,\mathbf{k})\widetilde{\delta\chi}(\tau',\mathbf{k}')|0\rangle = \delta^{(3)}(\mathbf{k}+\mathbf{k}')G(k,\tau,\tau')$ by employing perturbatively the solution of the free massless scalar field (66).

Let us now consider a specific example where the scalar field potential contains a cubic interaction term $(\lambda/3!)\chi^3$ [82,46] so that we can write

$$H_{\mathrm{int}} = \frac{\lambda}{3!}\delta\chi^3\ , \tag{86}$$

where λ is a coupling constant. From Eq. (84) it then follows

$$\begin{aligned}\langle\widetilde{\delta\chi}_{\mathbf{k}_1}\widetilde{\delta\chi}_{\mathbf{k}_2}\widetilde{\delta\chi}_{\mathbf{k}_3}\rangle = &-\mathrm{i}\lambda\delta^{(3)}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k}_3)\int_{-\infty}^{\tau}\frac{-\mathrm{d}\tau'}{H\tau'}[G(k_1,\tau,\tau')G(k_2,\tau,\tau')G(k_3,\tau,\tau')\\ &-G^*(k_1,\tau,\tau')G^*(k_2,\tau,\tau')G^*(k_3,\tau,\tau')]\ .\end{aligned} \tag{87}$$

Here τ corresponds to the conformal time at the end of inflation.

Such an integral depends on some combinations of the norms of the wavevectors like $\pi_1 = \sum_i k_i$, $\pi_2 = \sum_{i<j} k_i$, $\pi_3 = \sum_{i<j<k} k_i$. Actually it can be performed and expressed in the large scale limit, $k_i\tau \ll 1$ for all i, as [46]

$$\langle\delta\chi_{\mathbf{k}_1}\delta\chi_{\mathbf{k}_2}\delta\chi_{\mathbf{k}_3}\rangle = -\frac{\lambda H^2}{12}\frac{\delta^{(3)}(\sum_i\mathbf{k}_i)}{\prod k_i^3}\left[-\sum_i k_i^3(\gamma+\zeta_3(k_i)+\log[-k_t\tau])\right]\ , \tag{88}$$

where we have switched to the field $\delta\chi = \widetilde{\delta\chi}/a$. In formula (88) $k_T = k_1+k_2+k_3$, γ is the Euler constant $\gamma \approx 0.577$ and $\zeta_3(k_i)$ is a function which can be expressed in terms of the combinations π_i. As it has been shown in detail in Ref. [46] the function $\zeta(k_i)$ weakly depends on the wavevectors and it is always of the order of unity.[8]

Notice that it is standard use to express this result as the sum of products of the two-point correlation function on large scales given in Eq. (68)

$$\langle\delta\chi_{\mathbf{k}_1}\delta\chi_{\mathbf{k}_2}\delta\chi_{\mathbf{k}_3}\rangle = \nu_3(k_i)\sum_i\prod_{J\neq i}\frac{H^2}{2k_j^3}\ , \tag{90}$$

where

$$\nu_3(k_i) = \frac{\lambda}{3H^2}\left[\gamma+\zeta(k_i)+\log[-k_T\tau]\right]\ . \tag{91}$$

[8] The precise expression of $\zeta_3(k_i)$ is given by the appropriate limit $k_4 \to 0$ in the expression

$$\zeta_4(k_i) = \frac{-\pi_1^4+2\pi_1^2\pi_2+\pi_1\pi_3-3\pi_4}{\pi_1(\pi_1^3-3\pi_1\pi_2+3\pi_3)}\ , \tag{89}$$

where $\pi_4 = \sum_{i<j<k<l} k_i$. $\zeta_4(k_i)$ is the corresponding quantity that appears in the connected part of the correlation functions for a potential $\lambda/4!\,\chi^4$ (see Ref. [46] for more details).

The term $\log[-k_T\tau]$ corresponds to $N_{k_T} = \log(a_{\rm end}/a_{k_T})$ which is the number of e-foldings from the time the scale corresponding to k_T leaves the horizon during inflation and the end of inflation. Typically $N_{k_T} \approx 60$ for observable cosmological large scales and thus it dominates over the other terms which are typically of the order of unity so that one can approximate

$$v_3(k_i) \approx -\frac{\lambda N_{k_T}}{3H^2} \ . \tag{92}$$

This result, first found in Refs. [11,82], has actually a very transparent physical interpretation [46]. The first two terms in Eq. (91) can be interpreted as genuine quantum effects of the scalar field modes on scales smaller than the cosmological horizon which leave a characteristic (scale-dependent) imprint at the time of horizon crossing.[9] On the other hand after a few e-foldings after the modes leaves the horizon, the evolution of the field can be described in a classical way and it just corresponds to the term proportional to the number of e-foldings N_{k_T}.

In fact this can be shown by solving the Klein–Gordon equation for the scalar field χ in a de Sitter background in the large-scale limit *up to second order* in the perturbations. If we expand the scalar field as

$$\chi(\tau,\mathbf{x}) = \chi_0(\tau) + \delta\chi(\tau,\mathbf{x}) = \chi_0(\tau) + \delta^{(1)}\chi(\tau,\mathbf{x}) + \tfrac{1}{2}\,\delta^{(2)}\chi(\tau,\mathbf{x}) \ , \tag{93}$$

where we split the scalar field perturbation into a first and a second order part, the evolution equation for $\delta\chi^{(2)}$ on large scales reads in cosmic time (see Ref. [3])

$$\delta^{(\ddot{2})}\chi + 3H\delta^{(\dot{2})}\chi + \frac{\partial^2 V}{\partial\chi^2}\,\delta^{(2)}\chi = -\frac{\partial^3 V}{\partial\chi^3}(\delta^{(1)}\chi)^2 \ . \tag{94}$$

In a slow-roll approximation Eq. (94) becomes

$$3H\delta^{(\dot{2})}\chi \approx -\frac{\partial^3 V}{\partial\chi^3}(\delta^{(1)}\chi)^2 \ , \tag{95}$$

whose solution is

$$\delta^{(2)}\chi = -\frac{\lambda}{3H^2}\,N_k(\delta^{(1)}\chi)^2 + \delta^{(2)}\chi(t_k) \ , \tag{96}$$

where $N_k = \int_{t_k}^{t_{\rm end}} H\,{\rm d}t = H\Delta t$ is the number of e-folds between the end of inflation and the time t_k the scale of wave number k leaves the horizon during inflation and we have used the fact that $V = (\lambda/3!)\chi^3$. The integration constant $\delta^{(2)}\chi(t_k)$ is the value of the field at horizon-crossing and corresponds to the terms coming from quantum effects on subhorizon scales. However as it is evident from the result in Eq. (96)

[9] Following Ref. [182] we just recall here how to perform the integrals like (87). One can split them as integrals over the region outside the horizon, the region around horizon crossing and the region much smaller than the horizon. Moreover in order to take automatically into account that we are considering expectation values on the interacting vacuum one has to deform the τ integration contour so that it has some evolution in Euclidean time. This is achieved by the change $\tau \to \tau + {\rm i}\varepsilon|\tau|$, for large $|\tau|$. In this way since on subhorizon scales the fields oscillate rapidly, the integration over the region deep inside the horizon does not give any contribution.

these terms will be subdominant with respect to that corresponding to the scalar field dynamics occurring once the mode leaves the horizon. [10]

We conclude this section with some further comments. Such calculations are performed without taking into account any perturbation of the metric, such as the gravitational potential. Notice that if the scalar field χ is the inflaton field then according to these results the magnitude of the non-linearity parameter $f_{\rm NL}$ for the gravitational potential should be proportional to λ and then in terms of the slow-roll parameters $f_{\rm NL} \simeq \mathcal{O}(\xi^2)$, where $\xi^2 = M_{\rm P}^2 (V^{(1)} V^{(3)} / V^{(2)})$ which is second-order in the slow-roll parameters. However, it has been shown in Refs. [86,85] (see also Ref. [26]) that, when accounting also for the non-linearities in the metric perturbations the level of non-Gaussianity turns out to be $f_{\rm NL} \simeq \mathcal{O}(\epsilon, \eta)$. In fact the main contribution to the non-Gaussianity in single-field models of slow-roll inflation comes from the non-linear gravitational perturbations, rather than the inflaton self-interactions. The authors of Ref. [86] have used the so-called stochastic approach to inflation [275](for a more recent approach, see Ref. [189]). Such a result has been also recently obtained in a more rigorous way in Refs. [3,182] by studying the perturbations in the metric and in the inflaton field up to second order in deviations from the homogeneous background until the end of inflation. These results show a general principle holding for single-field models of slow-roll inflation. In order to have a period of inflation the inflaton potential must be very flat (i.e. $\epsilon, |\eta| \ll 1$), therefore the self-interaction terms in the inflaton potential and the gravitational coupling must very small and then non-linearities are suppressed too. On the other hand if the scalar field χ is different from the inflaton and it gives a negligible contribution to the total energy density, its self-interactions are not constrained by any inflationary slow-roll condition, and thus sizeable non-Gaussianities can be generated. This is the scenario which has been proposed for example in Ref. [11]. However, in this case the perturbations produced have a very little impact on the total energy density perturbation since the energy density of the scalar field is subdominant (the so-called isocurvature perturbations), and this kind of scenario is not in accordance with present observational data. There is one more interesting possibility for the self-interactions of a scalar field to play an important role in producing non-Gaussian signatures. If the subdominant scalar field χ is coupled to the inflaton field then it is possible that non-Gaussianities intrinsic in the scalar field χ are transferred to the inflaton sector, as it was first proposed in Ref. [30]. We will come back to this scenario in detail in Section 9.1.

---

[10] Notice that such a result, obtained computing second-order perturbations is completely equivalent to solving the equation of motion for the scalar field outside the horizon using a perturbative expansion in the coupling λ, as done in Ref. [46]. The equation of motion is

$$\ddot{\chi} + 3H\dot{\chi} = -\frac{\partial V}{\partial \chi} \ , \tag{97}$$

neglecting spatial gradients. This is equivalent to consider χ in this equation as a filtered field on scales that leave the horizon at a given time t_k [46]. At zeroth order in λ the solution $\chi_{(0)}$ is a Gaussian field which remains constant. The first-order correction then satisfies

$$\ddot{\chi}_{(1)} + 3H\dot{\chi}_{(1)} = -\frac{\partial V}{\partial \chi}(\chi_{(0)}) \ , \tag{98}$$

which for a potential $V(\chi) = \lambda\chi^3/3!$ gives

$$\chi_{(1)} = \chi_{(1)}(t_k) - \frac{\lambda}{2}(\chi_{(0)})^2 \frac{t - t_k}{3H} = \chi_{(1)}(t_k) - \frac{\lambda}{2}(\chi_{(0)})^2 \frac{N_k}{3H^2} \ . \tag{99}$$

The intrinsic non-linearities of the scalar fields present during inflation are only a possible source of non-Gaussianity. Indeed, even if the fluctuations of the scalar field χ are Gaussian distributed, it is possible that their energy density perturbations have some non-linearity. This is the case, for instance, of a scalar field χ different from the inflaton with a quadratic potential $V(\chi) \propto \chi^2$ leading to a vacuum expectation value $\langle\chi\rangle = 0$, as it has been proposed in Ref. [168] (see also [210]). In this case the energy density perturbations are not given by the usual linear contribution $\delta\rho_\chi \propto \chi\delta\chi$ but will be non-Gaussian with $\delta\rho_\chi \propto \delta\chi^2$. Actually such a quadratic contribution to the primordial energy density perturbations is the key feature of the non-Gaussianities in the curvaton scenario [176].

The gravitational dynamics itself introduces important non-linearities, which will contribute to the final non-Gaussianity in the large-scale CMB anisotropies. In fact as it has been shown in Ref. [31] it is just because of the non-linear gravitational dynamics that the tiny non-Gaussianity produced during inflation gets amplified in the post-inflationary evolution.

2.6. Metric perturbations and the energy–momentum tensor

In the previous sections we have shown how perturbations in a generic scalar field χ are generated on superhorizon scales during an inflationary period. This is the first step to understand the production and evolution of the cosmological perturbations in the different scenarios we are going to consider. In the standard single-field inflation, as well as in the curvaton and inhomogeneous reheating scenarios cosmological perturbations can be traced back initially to fluctuations of scalar fields; then they evolve in the radiation-dominated phase and, subsequently, in the matter and dark energy-dominated phases.

Let us first focus on the generation of the perturbations of scalar fields and make some remarks. As we have already emphasized in the discussion of the previous section, so far we have neglected the perturbations in the metric tensor around the homogeneous FRW background. In the case of the inflaton field taking into account the metric perturbations is of primary importance. The reason is very simple. The inflaton field drives the accelerated inflationary expansion, which means that it dominates the energy density of the Universe at that time. Thus any perturbation in the inflaton field $\delta\varphi$ implies a perturbation of the energy–momentum tensor $\delta T_{\mu\nu}$, and a perturbation in the energy–momentum tensor implies, through Einstein's equations of motion $G_{\mu\nu} = 8\pi G_{\rm N} T_{\mu\nu}$, a perturbation of the metric. On the other hand, perturbations in the metric affect the evolution of the inflaton fluctuations $\delta\varphi$ through the perturbed Klein–Gordon equation. We thus conclude that in the standard scenario of inflation perturbations of the inflaton field and perturbations of the metric are tightly coupled to each other and have to be studied together. Indeed this is the correct way to proceed. A very general scenario for the generation of the cosmological perturbations is one where other scalar fields χ_I are present besides the inflaton. This could be the case of inflation driven by several scalar fields whose contribution to the total energy density is comparable (multi-field inflation), or the case where the extra scalar fields are subdominant [164,274,198,231,166,129,132,200,168,228,152,93,28,29]. In such a general scenario a consistent way to study the production of cosmological fluctuations is to perturb both the scalar fields and the metric. The metric perturbations will then have a feedback also on the evolution of the subdominant scalar fields. Moreover, in such a general picture the different scalar fields can interact with one another through a generic potential $V(\varphi, \chi_I)$, while we have neglected such interactions for the scalar field χ so far.

During the radiation/matter-dominated eras, as we see again from the Einstein equations, a consistent study of the cosmological perturbations must take into account perturbations both in the energy–momentum-tensor and in the metric tensor. We shall see that the relation between the energy–momentum

perturbations and the metric is also justified in light of gauge issues. In fact we shall introduce some gauge-invariant quantities that mix both matter and metric perturbations. This will be essential in order to study the evolution of metric perturbations during the different stages, from the early period of inflation/reheating to the subsequent radiation and matter dominated epochs. As pointed out in the Introduction, in order to keep track of the non-Gaussianity of the cosmological perturbations throughout these different stages we perform our analysis *up to second order* in the perturbations. In particular we will focus on some quantities which are gauge-invariant up to second order, and which allow us to follow the evolution of the metric perturbations (the gravitational potentials) taking into account the different second-order contributions to the non-linearities of the perturbations.

3. Cosmological perturbations at first and second order

In order to study the perturbed Einstein's equations, we first write down the perturbations on a spatially flat FRW background following the formalism of Refs. [53,188,3]. We shall first consider the fluctuations of the metric, and then the fluctuations of the energy–momentum tensor. Hereafter greek indices will be taken to run from 0 to 3, while latin indices, labelling spatial coordinates, will run from 1 to 3. If not otherwise specified we will work with conformal time τ, and primes will denote differentiation with respect to τ.

3.1. The metric tensor

The components of a spatially flat FRW metric perturbed up to second order can be written as

$$\begin{aligned} g_{00} &= -a^2(\tau)(1+2\phi^{(1)}+\phi^{(2)}) \ , \\ g_{0i} &= a^2(\tau)(\hat{\omega}_i^{(1)}+\tfrac{1}{2}\,\hat{\omega}_i^{(2)}) \ , \\ g_{ij} &= a^2(\tau)[(1-2\psi^{(1)}-\psi^{(2)})\delta_{ij}+(\hat{\chi}_{ij}^{(1)}+\tfrac{1}{2}\,\hat{\chi}_{ij}^{(2)})] \ . \end{aligned} \tag{100}$$

The functions $\phi^{(r)}$, $\hat{\omega}_i^{(r)}$, $\psi^{(r)}$ and $\hat{\chi}_{ij}^{(r)}$, where $(r)=(1),(2)$, stand for the rth order perturbations of the metric. Notice that such an expansion could a priori include terms of arbitrary order [188], but for our purposes the first and second-order terms are sufficient. It is standard use to split the perturbations into the so-called scalar, vector and tensor parts according to their transformation properties with respect to the three-dimensional space with metric δ_{ij}, where scalar parts are related to a scalar potential, vector parts to transverse (divergence-free) vectors and tensor parts to transverse trace-free tensors. Thus in our case

$$\hat{\omega}_i^{(r)} = \partial_i\omega^{(r)}+\omega_i^{(r)} \ , \tag{101}$$

$$\hat{\chi}_{ij}^{(r)} = D_{ij}\chi^{(r)}+\partial_i\chi_j^{(r)}+\partial_j\chi_i^{(r)}+\chi_{ij}^{(r)} \ , \tag{102}$$

where ω_i and χ_i are transverse vectors, i.e. $\partial^i\omega_i^{(r)}=\partial^i\chi_i^{(r)}=0$, $\chi_{ij}^{(r)}$ is a symmetric transverse and trace-free tensor, i.e. $\partial^i\chi_{ij}^{(r)}=0$, $\chi_i^{i(r)}=0$ and $D_{ij}=\partial_i\partial_j-(1/3)\delta_{ij}\,\nabla^2$ is a trace-free operator.[11]

[11] Here and in the following latin indices are raised and lowered using δ^{ij} and δ_{ij}, respectively.

Let us recall that the reason why such a splitting has been introduced [20,128] is that, *at least* in linear theory, these different modes are decoupled from each other in the perturbed evolution equations, so that they can be studied separately. As we shall see throughout the following sections this property does not hold anymore beyond the linear regime where second-order perturbations *are* coupled—sourced—by first-order perturbations.

For our purposes the metric in Eq. (100) can be simplified. The fact that first-order vector perturbations have decreasing amplitudes and that are not generated in the presence of scalar fields, allows us to conclude that they can be safely disregarded. Moreover, the first-order tensor part gives a negligible contribution to second-order perturbations. Thus, in the following we will neglect $\omega_i^{(1)}$, $\chi_i^{(1)}$ and $\chi_{ij}^{(1)}$. However the same reasoning does not apply to second-order perturbations. Since in the non-linear case scalar, vector and tensor modes are dynamically coupled, the second-order vector and tensor contributions are generated by first-order scalar perturbations even if they were initially zero [188]. Thus we have to take them into account and we shall use the metric

$$\begin{aligned} g_{00} &= -a^2(\tau)(1+2\phi^{(1)}+\phi^{(2)}) \ , \\ g_{0i} &= a^2(\tau)(\partial_i\omega^{(1)}+\tfrac{1}{2}\,\partial_i\omega^{(2)}+\tfrac{1}{2}\,\omega_i^{(2)}) \ , \\ g_{ij} &= a^2(\tau)[(1-2\psi^{(1)}-\psi^{(2)})\delta_{ij}+D_{ij}(\chi^{(1)}+\tfrac{1}{2}\,\chi^{(2)})+\tfrac{1}{2}\,(\partial_i\chi_j^{(2)}+\partial_j\chi_i^{(2)}+\chi_{ij}^{(2)})] \ . \end{aligned} \tag{103}$$

The contravariant metric tensor is obtained by requiring (up to second order) that $g_{\mu\nu}g^{\nu\lambda}=\delta_\mu^\lambda$ and is given by

$$\begin{aligned} g^{00} &= -a^{-2}(\tau)(1-2\phi^{(1)}-\phi^{(2)}+4(\phi^{(1)})^2-\partial^i\omega^{(1)}\partial_i\omega^{(1)}) \ , \\ g^{0i} &= a^{-2}(\tau)[\partial^i\omega^{(1)}+\tfrac{1}{2}\,(\partial^i\omega^{(2)}+\omega^{i(2)})+2(\psi^{(1)}-\phi^{(1)})\partial^i\omega^{(1)}-\partial^i\omega^{(1)}D^i{}_k\chi^{(1)}] \ , \\ g^{ij} &= a^{-2}(\tau)[(1+2\psi^{(1)}+\psi^{(2)}+4(\psi^{(1)})^2)\delta^{ij}-D^{ij}(\chi^{(1)}+\tfrac{1}{2}\chi^{(2)}) \\ &\quad -\tfrac{1}{2}\,(\partial^i\chi^{j(2)}+\partial^j\chi^{i(2)}+\chi^{ij(2)})-\partial^i\omega^{(1)}\partial^j\omega^{(1)} \\ &\quad -4\psi^{(1)}D^{ij}\chi^{(1)}+D^{ik}\chi^{(1)}D_k^j\chi^{(1)}] \ . \end{aligned} \tag{104}$$

Using $g_{\mu\nu}$ and $g^{\mu\nu}$ one can calculate the connection coefficients $\Gamma^\alpha_{\beta\gamma}$ and the Einstein tensor components G^μ_ν up to second order in the metric fluctuations. We report their complete expressions in Appendix A; they can be also found in Ref. [3].

Let us conclude this section by noting that in the following we will often adopt the *Poisson gauge* [48,178] which is defined by the condition $\omega^{(r)}=\chi^{(r)}=\chi_i^{(r)}=0$. Then, one scalar degree of freedom is eliminated from g_{0i} and one scalar and two vector degrees of freedom from g_{ij}. This gauge generalizes the so-called longitudinal gauge to include vector and tensor modes.

3.2. The energy–momentum tensor

In this section we shall consider a fluid characterized by the energy–momentum tensor

$$T^\mu{}_\nu = (\rho+P)u^\mu u_\nu + P\,\delta^\mu{}_\nu \ , \tag{105}$$

where ρ is the energy density, P the pressure, and u^μ is the fluid four-velocity subject to the constraint $g^{\mu\nu}u_\mu u_\nu=-1$. Notice that we do not include any anisotropic stress term in our energy–momentum

tensor, i.e. we make the perfect fluid hypothesis, since in the different scenarios we are going to discuss anisotropic stresses are not present, as we only deal with scalar fields, matter and radiation components. Indeed, we devote a specific section to the energy–momentum tensor of a scalar field, given its importance for the standard scenario of inflation. Here, we also restrict ourselves to the case where the equation of state of the fluid $w = P/\rho$ is constant, with $w = \frac{1}{3}$ for a radiation fluid and $w = 0$ for collisionless matter (dust). We now expand the basic matter variables u^μ, ρ and P up to second order in the perturbations around the homogeneous background. For the velocity we write

$$u^\mu = \frac{1}{a}\left(\delta^\mu_0 + v^\mu_{(1)} + \frac{1}{2}\, v^\mu_{(2)}\right) \ . \tag{106}$$

From the normalization condition we obtain

$$\begin{aligned} v^0_{(1)} &= -\phi^{(1)} \ , \\ v^0_{(2)} &= -\phi^{(2)} + 3(\phi^{(1)})^2 + 2\partial_i \omega^{(1)} v^{(1)} + v^{(1)}_i v^i_{(1)} \ . \end{aligned} \tag{107}$$

Notice that the velocity perturbation $v^i_{(r)}$ also splits into a scalar (irrotational) and a vector (solenoidal) part, as

$$v^i_{(r)} = \partial^i v_{(r)} + v^i_{(r)\mathscr{S}} \ , \tag{108}$$

with $\partial_i v^i_{(r)\mathscr{S}} = 0$. According to what we said in the previous section we can neglect the linear vector velocity perturbation.[12]

Using the metric of Eq. (103) we find for $u_\mu = g_{\mu\nu} u^\nu$

$$\begin{aligned} u_0 &= a(-1 - \phi^{(1)} - \tfrac{1}{2}\,\phi^{(2)} + \tfrac{1}{2}\,(\phi^{(1)})^2 - \tfrac{1}{2}\, v^{(1)}_i v^i_{(1)}) \ , \\ u_i &= a(v^{(1)}_i + \partial_i \omega^{(1)} + \tfrac{1}{2}\, v^{(2)}_i + \tfrac{1}{2}\, \omega^{(2)}_i - \phi^{(1)} \partial_i \omega^{(1)} - 2\psi^{(1)} v^{(1)}_i + D_{ij} \chi^{(1)} v^j_{(1)}) \ . \end{aligned} \tag{109}$$

The energy density ρ can be split into a homogeneous background $\rho_0(\tau)$ and a perturbation $\delta\rho(\tau, x^i)$ as follows:

$$\rho(\tau, x^i) = \rho_0(\tau) + \delta\rho(\tau, x^i) = \rho_0(\tau) + \delta^{(1)}\rho(\tau, x^i) + \tfrac{1}{2}\,\delta^{(2)}\rho(\tau, x^i) \ , \tag{110}$$

where the perturbation has been expanded into a first and a second-order part. The same decomposition can be adopted for pressure P, where in our case $\delta P = w\delta\rho$.

Using expression (110) for the energy density and the expressions for the velocity into Eq. (105) we calculate $T^\mu{}_\nu$ up to second order and we find

$$T^\mu{}_\nu = T^{\mu(0)}{}_\nu + \delta^{(1)} T^\mu{}_\nu + \delta^{(2)} T^\mu{}_\nu \ , \tag{111}$$

where $T^{\mu(0)}{}_\nu$ corresponds to the background value, and

$$T^{0(0)}{}_0 + \delta^{(1)} T^0{}_0 = -\rho_0 - \delta^{(1)}\rho \ , \tag{112}$$

[12] Notice, however, that in the following expression for the perturbed energy–momentum tensor we are still completely general by including also linear vector and tensor perturbation modes.

$$\delta^{(2)}T^{0}{}_{0} = -\tfrac{1}{2}\,\delta^{(2)}\rho - (1+w)\rho_0 v_i^{(1)} v^i_{(1)} - (1+w)\rho_0 \partial_i \omega^{(1)} v^i_{(1)} \ , \tag{113}$$

$$T^{i(0)}{}_{0} + \delta^{(1)}T^{i}{}_{0} = -(1+w)\rho_0 v^i_{(1)} \ , \tag{114}$$

$$\delta^{(2)}T^{i}{}_{0} = -(1+w)\rho_0 \left[\frac{1}{2}\, v^i_{(2)} + \left(\psi^{(1)} + \frac{\delta^{(1)}\rho}{\rho_0}\right) v^i_{(1)}\right] \ , \tag{115}$$

$$T^{i(0)}{}_{j} + \delta^{(1)}T^{i}{}_{j} = w\rho_0 \left(1 + \frac{\delta^{(1)}\rho}{\rho_0}\right) \delta^{i)}{}_{j} \tag{116}$$

$$\delta^{(2)}T^{i}{}_{j} = \tfrac{1}{2}\, w\delta^{(2)}\rho\delta^{i)}{}_{j} + (1+w)\rho_0 v^i_{(1)}(v^{(1)}_j + \partial_j \omega^{(1)}) \ . \tag{117}$$

A comment is in order here. As it can be seen from Eqs. (103) and (113), (115) and (117), the second-order perturbations always contain two different contributions, quantities which are intrinsically of second order, and quantities which are given by the product of two first-order perturbations. As a consequence, when considering the Einstein equations to second order in Section 4.2, first-order perturbations behave as a source for the intrinsically second-order fluctuations. This is an important issue which was pointed out in different works on second-order perturbation theory [280,281,190,191,188] and it plays a central role in deriving our main results on non-Gaussianity of cosmological perturbations.

3.3. Gauge dependence at second order

In the previous sections we have defined the perturbation δT in a given quantity T considering the difference between the physical value of T (the perturbed one) and the background unperturbed value T_0, and in the specific we have then expanded such a perturbation in a first and a second-order part. However the theory of perturbations in general relativity intrinsically encodes a certain degree of arbitrariness in performing such a comparison between the physical and the reference background quantities. This is because we consider perturbations of the space–time itself on which a given quantity is defined. Thus we deal with two different space–times, the real physical space–time and the unperturbed background (which in our case is the FRW space–time), where T and T_0 are defined, respectively. In order to compare the value of T with the reference value T_0 it is necessary to establish a map, that is a one-to-one correspondence, between the physical and the background space–times. Such a map is a gauge choice, and a change of the map is a gauge transformation. From the point of view of a set of coordinates this means a change of the coordinates. However the gauge choice is not unique, since General Relativity *i*s a theory based on the freedom of changing locally the system of coordinates. Therefore the value of the perturbation in the generic quantity T depends on the gauge, or in other words, the perturbation δT will transform after a change of coordinates thus acquiring different values, which nonetheless are on the same footing. This is the issue of gauge-dependence, which holds at any order of perturbations. There are two options to avoid such an ambiguity: identify combinations of perturbations which are gauge-invariant quantities, that is to say quantities which are independent of the gauge transformation, or choosing a given gauge and perform the calculations in that gauge. The second option could introduce pure gauge-modes, which have no physical meaning, and must be eliminated from the solutions. The gauge-dependence has been widely studied within linear perturbation theory in different papers [246,276,20,78,277] and discussed specifically in the context of cosmological perturbations in some reviews [128,207]. Only recently the

gauge issue has been addressed in a systematic way beyond the linear regime [53,267,188] giving a full description of gauge transformations at second order.

It is not the purpose of this review to describe in detail gauge-transformations for second-order perturbations, and we refer the reader to detailed papers on the subject [53,267]. We want indeed to focus on some gauge-invariant quantities which play a major role when study the evolution of second-order perturbations and which allow us to determine accurately the resulting non-Gaussianity. Thus in the following we just give some of the transformation rules that cosmological perturbations around a flat FRW background obey to up to second order. They can be useful to check some of the gauge-invariant quantities we shall introduce.

3.3.1. Gauge transformations

Let us consider an infinitesimal coordinate transformation up to second order

$$\tilde{x}^\mu(\lambda) = x^\mu - \xi^\mu_{(1)} - \tfrac{1}{2}\,(\xi^\mu_{(1),\nu}\xi^\nu_{(1)} + \xi^\mu_{(2)}) \ , \tag{118}$$

where $\xi^{(\mu)}_{(r)}(\tau, x^i)$ are vector fields defining the gauge transformation, being regarded as quantities of the same order as the perturbation variables. Specifying their time and space components, one can write

$$\xi^0_{(r)} = \alpha_{(r)} \ , \tag{119}$$

and

$$\xi^i_{(r)} = \partial^i \beta_{(r)} + d^i_{(r)} \ , \tag{120}$$

where we have split the space component into a scalar and a vector part with $\partial_i d^{(r)i} = 0$. From a practical point of view fixing a gauge is equivalent to fixing a coordinate system. In particular the function $\xi^0_{(r)}$ selects constant-τ hypersurfaces, i.e. a time-slicing, while $\xi^i_{(r)}$ selects the spatial coordinates within those hypersurfaces.

If we now expand a generic tensor $T(\tau, x^i)$ defined in the real physical word up to second order as

$$T(\tau, x^i) = T_0 + \delta T(\tau, x^i) = T_0 + \delta^{(1)}T(\tau, x^i) + \tfrac{1}{2}\,\delta^{(2)}T(\tau, x^i) \ , \tag{121}$$

where T_0 is the background value, then its perturbations transform as [188]

$$\widetilde{\delta^{(1)}T} = \delta^{(1)}T + \pounds_{\xi_{(1)}} T_0 \ , \tag{122}$$

$$\widetilde{\delta^{(2)}T} = \delta^{(2)}T + 2\pounds_{\xi_{(1)}}\delta^{(1)}T + \pounds^2_{\xi_{(1)}} T_0 + \pounds_{\xi_{(2)}} T_0 \ , \tag{123}$$

where $\pounds_{\xi_{(r)}}$ is the Lie derivative along the vector $\xi^\mu_{(r)}$.

Thus, for example, the energy density perturbation transforms at first order as

$$\widetilde{\delta^{(1)}\rho} = \delta^{(1)}\rho + \rho_0'\alpha_{(1)} \ , \tag{124}$$

and at second order as [188]

$$\widetilde{\delta^{(2)}\rho} = \delta^{(2)}\rho + \rho_0'\alpha_{(2)} + \alpha_{(1)}(\rho_0''\alpha_{(1)} + \rho_0'\alpha_{(1)}' + 2\delta^{(1)}\rho') + \xi^i_{(1)}(\rho_0'\alpha^{(1)}_{,i} + 2\delta\rho_{,i}) \ . \tag{125}$$

By transforming the metric tensor perturbations $\delta^{(1)}g_{\mu\nu}$ and $\delta^{(2)}g_{\mu\nu}$ in the metric (103) according to Eqs. (122) and (123) one finds that the metric perturbation $\psi = \psi^{(1)} + \frac{1}{2}\psi^{(2)}$ transforms at first order as

$$\widetilde{\psi^{(1)}} = \psi^{(1)} - \frac{1}{3}\nabla^2\beta_{(1)} - \frac{a'}{a}\alpha_{(1)} \ , \tag{126}$$

and at second order as [188]

$$\begin{aligned}\widetilde{\psi^{(2)}} = {} & \psi^{(2)} + \alpha_{(1)}\left[2\left(\psi^{(1)\prime} + 2\frac{a'}{a}\psi^{(1)}\right) - \left(\frac{a''}{a} + \frac{a'^2}{a^2}\right)\alpha_{(1)} - \frac{a'}{a}\alpha'_{(1)}\right] \\ & + \xi^i_{(1)}\left(2\psi^{(1)}_{,i} - \frac{a'}{a}\alpha^{(1)}_{,i}\right) - \frac{1}{3}\left(-4\psi^{(1)} + \alpha_{(1)}\partial_0 + \xi^i_{(1)}\partial_i + 4\frac{a'}{a}\alpha_{(1)}\right)\nabla^2\beta_{(1)} \\ & - \frac{1}{3}(2\omega^{,i}_{(1)} - \alpha^{,i}_{(1)} + \xi^{i\prime}_{(1)})\alpha^{(1)}_{,i} - \frac{1}{3}(2\chi^{(1)}_{ij} + \xi^{(1)}_{i,j} + \xi^{(1)}_{j,i})\xi^{j,i}_{(1)} - \frac{a'}{a}\alpha_{(2)} \\ & - \frac{1}{3}\nabla^2\beta_{(2)} \ .\end{aligned} \tag{127}$$

3.4. Second-order gauge-invariant perturbations

In linear theory a gauge-invariant treatment of cosmological perturbations was introduced by Bardeen in his seminal work [20]. As far as non-linear perturbations are concerned the first results were found in Refs. [249,250,4], using a gradient-expansion technique (or long-wavelength approximation), and in Refs. [3,182], using a second-order perturbative approach. In these works a generalization of the so-called curvature perturbation to linear order was found in the context of single-field models of inflation, and its constancy in time was proved. Refs. [249,250,3,182] focused on the study of non-Gaussianity of cosmological perturbations. These were the first papers to fully account for—at least during the inflationary epoch—the different second-order effects both in the inflaton field and in the metric perturbations.

However no gauge-invariant theory for non-linear perturbations in the context of cosmological perturbations has been built up until some very recent papers on the subject [177,185,241,213,242]. Specifically the authors of Refs. [177,185] use a second-order perturbative approach, while in Ref. [241,242] a long-wavelength approximation is employed. Also, we refer the reader to Refs. [52,212], where the issue of gauge-invariance at second and higher order has been addressed from a broader point of view.

We now give the expressions of some quantities which are gauge-invariant up to second order in the perturbations and which we shall use to follow the non-Gaussianity of cosmological perturbations from an early period of inflation through reheating and deep into the radiation/matter dominated epochs. In particular we give particular relevance to the gauge-invariant definition of curvature perturbations.

3.5. The curvature perturbation on spatial slices of uniform density

At linear order the intrinsic spatial curvature on hypersurfaces on constant conformal time τ and for a flat Universe is given by [20,128]

$${}^{(3)}R = \frac{4}{a^2}\nabla^2\hat{\psi}^{(1)} \ , \tag{128}$$

where for simplicity of notation[13] we have indicated

$$\hat{\psi}^{(1)} = \psi^{(1)} + \tfrac{1}{6}\nabla^2\chi^{(1)} \ . \tag{129}$$

The quantity $\hat{\psi}^{(1)}$ is usually referred to as the *curvature perturbation*. However the curvature perturbation $\psi^{(1)}$ is *not* gauge invariant, but is defined only on a given slicing. In fact, under a transformation on constant time hypersurfaces $\tau \to \tau + \alpha_{(1)}$ (change of the slicing in Eq. (126))

$$\hat{\psi}^{(1)} \to \hat{\psi}^{(1)} - \mathscr{H}\alpha_{(1)} \ , \tag{130}$$

where we have used Eq. (126) and the transformation $\widetilde{\chi^{(1)}} = \chi^{(1)} + 2\beta_{(1)}$ [188]. If we consider the *slicing of uniform energy density* which is defined to be the slicing where there is no perturbation in the energy density, $\delta\rho = 0$, from Eq. (124) we have $\alpha_{(1)} = \delta^{(1)}\rho/\rho_0'$ and the curvature perturbation $\hat{\psi}^{(1)}$ on uniform density perturbation slices—usually indicated by $-\zeta^{(1)}$ is given by

$$-\zeta^{(1)} \equiv \widetilde{\hat{\psi}^{(1)}}|_\rho = \hat{\psi}^{(1)} + \mathscr{H}\frac{\delta^{(1)}\rho}{\rho_0'} \ . \tag{131}$$

This quantity is gauge-invariant and it is a clear example of how to find a gauge-invariant quantity by selecting in an unambiguous way a proper time slicing. It was first introduced in Refs. [22,21] as a conserved quantity on large scales for purely adiabatic perturbations.

Notice that such a combination can be regarded also as the density perturbation on uniform curvature slices, where $\psi^{(1)} = \chi^{(1)} = 0$, the so-called *spatially flat gauge* [128]. The energy density ρ here has to be regarded as the total energy density. If the matter content of a system is made of several fluids it is possible to define similarly the curvature perturbations associated with each individual energy density components ρ_i, which to linear order are given by [175,186]

$$\zeta_i^{(1)} = -\hat{\psi}^{(1)} - \mathscr{H}\left(\frac{\delta^{(1)}\rho_i}{\rho_i'}\right) \ . \tag{132}$$

Here and in the following, if not specified, we drop the subscript '0' referring to the background quantities for simplicity of notation. Notice that the total curvature perturbation in Eq. (131) is given in terms of the individual curvature perturbations as

$$\zeta^{(1)} = \sum_i \frac{\rho_i'}{\rho}\zeta_i^{(1)} \ . \tag{133}$$

We now come to the generalization at second order of the gauge-invariant curvature perturbation of Eq. (131). As in Eq. (103) we can define the gravitational potential ψ up to second order as $\psi = \psi^{(1)} + \frac{1}{2}\psi^{(2)}$

[13] Notice that our notation is different from that of Refs. [207,183,185] for the presence of D_{ij} in the metric (100), while it is closer to the one used in Refs. [20,128]. As far as the first-order perturbations are concerned the metric perturbations ψ and E of Refs. [207,183] are given in our notation as $\psi = \psi^{(1)} + (1/6)\nabla^2\chi^{(1)}$ and $E = \chi^{(1)}/2$, respectively. The same is true at second order for the perturbation variables ψ_2 and E_2 of Ref. [185], which in terms of our quantities are given by $\psi_2 = \psi^{(2)} + (1/6)\nabla^2\chi^{(2)}$ and $E_2 = \chi^{(2)}/2$. However, no difference appears in the calculations when using the Poisson gauge or the spatially flat gauge, or when considering the perturbation evolution on large scales.

and we expand the energy density as in Eq. (110). The authors of Refs. [177,185][14] have shown that the second-order curvature perturbation on uniform (total) density hypersurfaces is given by the quantity (up to a gradient term)

$$
\begin{aligned}
-\zeta^{(2)} &= \widetilde{\hat{\psi}^{(2)}}|_{\rho} \\
&= \hat{\psi}^{(2)} + \mathscr{H}\frac{\delta^{(2)}\rho}{\rho'} - 2\mathscr{H}\frac{\delta^{(1)}\rho'}{\rho'}\frac{\delta^{(1)}\rho}{\rho'} - 2\frac{\delta^{(1)}\rho}{\rho'}(\hat{\psi}^{(1)\prime} + 2\mathscr{H}\hat{\psi}^{(1)}) \\
&\quad + \left(\frac{\delta^{(1)}\rho}{\rho'}\right)^2\left(\mathscr{H}\frac{\rho''}{\rho'} - \mathscr{H}' - 2\mathscr{H}^2\right) ,
\end{aligned}
\tag{134}
$$

where as in Eq. (129) we have used the shorthand notation $\hat{\psi}^{(2)} = \psi^{(2)} + \frac{1}{6}\nabla^2\chi^{(2)}$. As explained in Refs. [177,185] the quantity $\zeta^{(2)}$ is gauge-invariant, being constructed on a well-defined time slicing corresponding to spatial hypersurfaces where $\delta^{(1)}\rho = \delta^{(2)}\rho = 0$. In a similar manner to linear order let us introduce the gauge-invariant curvature perturbations $\zeta_i^{(2)}$ at second order relative to a particular component. These quantities will be given by the same formula as Eq. (134) relatively to each energy density ρ_i

$$
\begin{aligned}
-\zeta_i^{(2)} &= \hat{\psi}^{(2)} + \mathscr{H}\frac{\delta^{(2)}\rho_i}{\rho_i'} - 2\mathscr{H}\frac{\delta^{(1)}\rho_i'}{\rho_i'}\frac{\delta^{(1)}\rho_i}{\rho_i'} - 2\frac{\delta^{(1)}\rho_i}{\rho_i'}(\hat{\psi}^{(1)\prime} + 2\mathscr{H}\hat{\psi}^{(1)}) \\
&\quad + \left(\frac{\delta^{(1)}\rho_i}{\rho_i'}\right)^2\left(\mathscr{H}\frac{\rho_i''}{\rho_i'} - \mathscr{H}' - 2\mathscr{H}^2\right) .
\end{aligned}
\tag{135}
$$

3.6. Adiabatic and entropy perturbations

The gauge-invariant curvature perturbations introduced in the previous sections are usually adopted to characterize the so-called adiabatic perturbations. In fact adiabatic perturbations are such that a net perturbation in the total energy density and—via the Einstein equations—in the intrinsic spatial curvature are produced. However, as we have seen, neither the energy density nor the curvature perturbations are gauge-invariant, hence the utility of using the variable $\zeta = \zeta^{(1)} + \frac{1}{2}\zeta^{(2)}$ to define such perturbations. Thus the notion of adiabaticity applies when the properties of a fluid, in the physical perturbed space–time, can be described *uniquely* in terms of its energy density ρ. For example, the pressure perturbation will be adiabatic if the pressure is a unique function of the energy density $P = P(\rho)$ (see Ref. [177] for an exhaustive discussion on this point).

On the other hand, by the same token, to define a non-adiabatic (or entropy) perturbation of a given quantity X it is necessary to "extract" that part of the perturbation which does not depend on the energy density. A very general prescription to do that is to consider the value of the perturbation $\delta X = \delta^{(1)}X + \frac{1}{2}\delta^{(2)}X$ on the hypersurfaces of uniform energy density

$$
\delta X_{\rm nad} \equiv \widetilde{\delta X}|_{\rho} ,
\tag{136}
$$

[14] The reader is also referred to Refs. [3,182,240] for related quantities and definitions.

since this quantity will vanish for adiabatic perturbations when $X = X(\rho)$. Specifically the non-adiabatic pressure perturbation will be given by the pressure perturbation on slices of uniform energy density $\delta\widetilde{P|_\rho}$. Being specified in a non-ambiguous slicing, the entropy perturbations defined in Eq. (136) turn out to be gauge-invariant. Notice that such a definition holds true both when considering quantities on the uniform total energy density hypersurfaces and when considering hypersurfaces of uniform energy density relative to each individual component when more than one fluid is present.

Before moving to the explicit expressions for the first- and second-order adiabatic and entropy perturbations an important remark is in order. In general the perturbations will not be exclusively of adiabatic or of entropy type, but both perturbation modes will be present. Indeed, as we will see in the next section the non-adiabatic pressure perturbation $\delta\widetilde{P|_\rho}$ sources the total curvature perturbation ζ on large scales. Such a coupling is the mechanism responsible for the generation of cosmological perturbations in the curvaton and in the inhomogeneous reheating scenarios, contrary to the standard single-field inflationary scenario where only adiabatic perturbations are involved.

3.6.1. Adiabatic and entropy perturbations at first order

At first order the non-adiabatic pressure perturbation is given by [20,128]

$$\delta^{(1)}P_{\rm nad} \equiv \widetilde{\delta^{(1)}P|_\rho} = \delta^{(1)}P - c_s^2\delta^{(1)}\rho \ , \tag{137}$$

where $c_s^2 = P_0'/\rho_0'$ is the adiabatic sound speed of the fluid. As a check of what we said above, notice that indeed this quantity is gauge-invariant.

It can be shown that in the presence of more than one fluid the total non-adiabatic pressure perturbation can be split into two parts

$$\delta^{(1)}P_{\rm nad} = \delta^{(1)}P_{\rm int} + \delta^{(1)}P_{\rm rel} \ . \tag{138}$$

The first part is given by the sum of the intrinsic entropy perturbation of each fluid

$$\delta^{(1)}P_{\rm int} = \sum_i \delta^{(1)}P_{{\rm intr},i} \ , \tag{139}$$

where

$$\delta^{(1)}P_{{\rm intr},i} = \delta^{(1)}P_i - c_i^2\delta^{(1)}\rho_i \tag{140}$$

is the intrinsic non-adiabatic pressure perturbation of that fluid (which is a gauge-invariant quantity) with $c_i^2 = p_i'/\rho_i'$ the adiabatic sound speed of the individual fluid. The second part is given by the relative entropy perturbation between different fluids [186]

$$\delta^{(1)}P_{\rm rel} = \frac{1}{6\mathscr{H}\rho'}\sum_{ij}\rho_i'\rho_j'(c_i^2 - c_i^2)\mathscr{S}_{ij}^{(1)} \ , \tag{141}$$

where $\mathscr{S}_{ij}$ is the relative energy density perturbation whose gauge-invariant definition is expressed in terms of the curvature perturbations $\zeta_i^{(1)}$ of Eq. (132) as [290,186]

$$\mathscr{S}_{ij}^{(1)} = 3(\zeta_i^{(1)} - \zeta_j^{(1)}) \ . \tag{142}$$

Notice that for fluids with no intrinsic entropy perturbations, the pressure perturbation will be adiabatic if the relative entropy perturbations vanish

$$\zeta_i^{(1)} = \zeta_j^{(1)} \ . \tag{143}$$

In such a case this is the condition to have *pure* adiabatic perturbations. As a consequence, from Eq. (133) we see that the total curvature is equally shared by the different components $\zeta^{(1)} = \zeta_i^{(1)}$.

On the other hand a *pure* isocurvature perturbation is such that the individual components compensate with each other in order to leave the curvature perturbation unperturbed. This is the reason why these are also referred to as *isocurvature* perturbations.

In Refs. [184,290] it has been shown how to derive the evolution equation for the curvature perturbation $\zeta^{(1)}$ simply from the continuity equation for the energy density, without making any use of Einstein's equations. The result is that even on large scales the curvature perturbation can evolve being sourced by the non-adiabatic pressure of the system according to [90,174,184,290]

$$\zeta^{(1)'} = -\frac{\mathscr{H}}{\rho + P}\,\delta^{(1)} P_{\rm nad} \ . \tag{144}$$

For purely adiabatic perturbations the curvature perturbation is conserved on large scales, thus making $\zeta^{(1)}$ the proper quantity to characterize the amplitude of adiabatic perturbations. Eq. (144) shows in particular that the notion of isocurvature perturbation is valid only at some initial epoch.[15] Indeed the fact that the non-adiabatic pressure perturbation sources the curvature perturbation on large scales was already known in the literature [20,207,195]. However, it was only recently that this issue has received renewed attention, being applied in the context of the curvaton scenario as an alternative way to produce adiabatic density perturbation starting from an initial entropy mode.

3.6.2. Adiabatic and entropy perturbations at second order

Up to second order in the perturbations it has been shown that the gauge-invariant non-adiabatic pressure perturbation is given by [185]

$$\delta^{(2)} P_{\rm nad} \equiv \widetilde{\delta^{(2)} P|_\rho} = \delta^{(2)} P - \frac{P'}{\rho'}\delta^{(2)}\rho + P'\left[2\left(\frac{\delta^{(1)'}\rho}{\rho'} - \frac{\delta^{(1)'}P}{P'}\right)\frac{\delta^{(1)}\rho}{\rho'} + \left(\frac{P''}{P} - \frac{\rho''}{\rho}\right)\left(\frac{\delta^{(1)}\rho}{\rho'}\right)^2\right] \ . \tag{145}$$

[15] On the other hand, as far as the evolution of the entropy perturbation itself is concerned, it has been shown that the non-adiabatic part of a perturbation is sourced on large scales only by other entropy perturbations, and that there is no source term coming from the overall curvature perturbation [175]. If we indicate generically an entropy perturbation as $\mathscr{S}$ then its equation of motion on large scales reads

$$\mathscr{S}' = \beta\mathscr{H}\mathscr{S} \ ,$$

where β is a time-dependent function which depends on the particular system under study. This result has also been obtained on very general grounds within the "separate Universe approach" of Ref. [290].

In a similar manner as in Eq. (143), for a set of fluids (with no intrinsic entropy perturbations) we can define pure adiabatic perturbations as those obeying the gauge-invariant condition

$$\zeta_i^{(2)} = \zeta_j^{(2)} \ . \tag{146}$$

Also at second order the curvature perturbation $\zeta^{(2)}$ on large scales evolves due to the non-adiabatic pressure perturbation, as shown in Ref. [185]

$$\zeta^{(2)'} = -\frac{\mathscr{H}}{\rho+P}\delta^{(2)}P_{\rm nad} - \frac{2}{\rho+P}\left[\delta^{(1)}P_{\rm nad} - 2(\rho+P)\zeta^{(1)}\right]\zeta^{(1)'} \ , \tag{147}$$

where the first-order curvature perturbation obeys Eq. (144).

The issue of the conservation of the curvature perturbation at second order (and beyond) for adiabatic perturbations has also been addressed in Ref. [177], while in Refs. [241,242] the evolution for the curvature perturbation has been obtained in the context of the long-wavelength approximation.

4. Evolution of cosmological perturbations up to second order

Let us now consider the evolution of cosmological perturbations on large scales up to second order. Our aim is to follow the non-linearity in the perturbations from an early period of inflation through the different post-inflationary stages till today. This will enable us to give a definite prediction for the level of non-Gaussianity in different scenarios for the generation of the cosmological perturbations, namely the standard single field inflation, the curvaton and the inhomogeneous reheating scenarios. In each of these the cosmological evolution can be divided into three main stages:

(1) A primordial epoch of accelerated expansion, when cosmological perturbations are produced on large scales from quantum fluctuations of a scalar field, which can be different from the inflaton, as in the curvaton and in the inhomogeneous reheating models.
(2) An epoch when the perturbations in the energy density of the scalar fields are transferred to a radiation fluid during the reheating stage. During this stage the inflaton field (and the curvaton field, if present) oscillates around the minimum of its potential behaving as non-relativistic matter and then it decays into light particles (the radiation fluid).
(3) After the reheating stage an overall adiabatic perturbation is generated and the Universe enters into a "post-inflationary" phase dominated by radiation and, subsequently, matter (and dark energy).

We shall follow the evolution of the different second-order effects throughout these phases using the gauge-invariant curvature perturbations introduced previously [31–33]. This is a highly efficient method for different reasons. As we will see from the continuity equation for the energy–momentum tensor, the overall curvature perturbation evolves on large scales due to a non-adiabatic pressure perturbation which can be expressed in terms of the individual curvature perturbations; thus we will be able to connect the different evolutionary stages. Moreover, by using gauge-invariant quantities we can pass from one gauge to another, in order to simplify some calculations, in a straightforward way. Finally, as stressed in Section 2.6, the curvature perturbations are a combination of the gravitational potential $\psi = \psi^{(1)} + \frac{1}{2}\psi^{(2)}$ and of the energy density perturbations $\delta\rho = \delta^{(1)}\rho + \frac{1}{2}\delta^{(2)}\rho$, which is very useful to obtain the final second-order contributions to the gravitational potentials from the different stages. When calculating

such contributions during the radiation/matter dominated phases we use the Einstein equations in order to relate the energy density fluctuations and the gravitational potential. As we shall see the prototype of this procedure is given by the standard scenario of single field inflation, where the curvature perturbation is in fact conserved on large scales, since in this case perturbations remain adiabatic. Our last step will be to define how our results must be compared to the observations, to search for possible non-Gaussian signatures in the CMB temperature anisotropy on large scales. To this aim, in Section 8.4 we will determine how the non-linearities in the gravitational potentials translate into non-linearities of the CMB temperature fluctuations on large angular scales [33].

In the following we derive the perturbed Einstein equations and the energy–momentum continuity equations up to second order, for a Universe filled by multiple interacting fluids, consisting of a (oscillating) scalar field and a radiation fluid, and for a Universe which is radiation/matter dominated. This is all what we need in order to study the three different cosmological scenarios, apart from a detailed analysis of the generation *during* inflation of second-order cosmological perturbations from the inflaton fluctuations. This analysis will be made separately in Section 5.3. We now strictly follow Refs. [31–33] which are the first works to systematically address the evolution of the second-order primordial cosmological perturbations in the different scenarios for perturbation generation.

4.1. First-order Einstein equations

Our starting point are the perturbed Einstein equations $\delta G^{\mu}{}_{\nu} = \kappa^2 \delta T^{\mu}{}_{\nu}$. Here $\kappa^2 \equiv 8\pi G_{\rm N}$. As the matter content we take the generic fluid defined by the energy–momentum tensor given in Section 3.2. The detailed expressions for the Einstein tensor components $\delta G^{\mu}{}_{\nu}$ from the metric in Eq. (103) are contained in Appendix A, and from there one can read the Einstein equations in different gauges. Here we only report those equations which we shall use to derive our main results. Specifically, in the Poisson gauge defined in Section 3.1 the first-order the $(0-0)$- and the $(i-0)$-components of Einstein equations are

$$\frac{1}{a^2}[6\mathscr{H}^2\phi^{(1)} + 6\,\mathscr{H}\psi^{(1)\prime} - 2\nabla^2\psi^{(1)}] = -\kappa^2\delta^{(1)}\rho \ , \tag{148}$$

$$\frac{2}{a^2}(\mathscr{H}\partial^i\phi^{(1)} + \partial^i\psi^{(1)\prime}) = -\kappa^2\,(1+w)\,\rho_0 v^i_{(1)} \ , \tag{149}$$

where $w \equiv P/\rho$ is the equation of state of the fluid.

In the Poisson gauge at first-order the non-diagonal part of the $(i-j)$-component of Einstein equations, gives

$$\psi^{(1)} = \phi^{(1)} \ , \tag{150}$$

and, on superhorizon scales, Eq. (148) gives

$$\psi^{(1)} = -\frac{1}{2}\,\frac{\delta^{(1)}\rho}{\rho_0} = \frac{3(1+w)}{2}\,\mathscr{H}\,\frac{\delta^{(1)}\rho}{\rho'} \tag{151}$$

where in the last step we have used the background continuity equation $\rho' = -3\mathscr{H}\rho(1+w)$.

Using the spatially flat gauge $\psi^{(1)} = \chi^{(1)} = 0$, from the $(0-0)$-component of Einstein equation we get a similar result for the gravitational potential $\phi^{(1)}$[16]

$$\phi^{(1)} = -\frac{1}{2}\frac{\delta^{(1)}\rho}{\rho_0} \ . \tag{152}$$

Notice that Eqs. (148), (149), (151) and (152) indeed hold also when referring to the total energy density ρ and the total velocity perturbation and equation of state w in the case of a multiple component system.

4.2. Second-order Einstein equations

Let us consider the Einstein equations perturbed at second-order $\delta^{(2)}G^{\mu}{}_{\nu} = \kappa^2\delta^{(2)}T^{\mu}{}_{\nu}$. The second-order expression for the Einstein tensor $\delta^{(2)}G^{\mu}{}_{\nu}$ can be found in any gauge in Appendix A.

We first consider the Einstein equations in the Poisson gauge which will be used in particular to express the non-linearities in the gravitational potential $\phi^{(2)}$ during the radiation/matter dominated phases.

- The $(0-0)$-component of Einstein equations leads to

$$\begin{aligned} &3\mathscr{H}^2\phi^{(2)} + 3\mathscr{H}\psi^{(2)\prime} - \nabla^2\psi^{(2)} - 12\mathscr{H}^2(\psi^{(1)})^2 - 3(\nabla\psi^{(1)})^2 \\ &\quad - 8\psi^{(1)}\nabla^2\psi^{(1)} - 3(\psi^{(1)\prime})^2 = \kappa^2 a^2\delta^{(2)}T^0{}_0 \ . \end{aligned} \tag{153}$$

- At second order the gravitational potentials $\phi^{(2)}$ and $\psi^{(2)}$ differ even in the Poisson gauge for the presence of source terms which are quadratic in the first-order perturbations. In fact it is possible to find the second-order equivalent of the linear constraint $\psi^{(1)} = \phi^{(1)}$ using the traceless part of the $(i-j)$-components of Einstein equations. One finds the following constraint relating the gravitational potentials $\psi^{(2)}$ and $\phi^{(2)}$[17] [31]:

$$\begin{aligned} \psi^{(2)} - \phi^{(2)} = &-4(\psi^{(1)})^2 - \nabla^{-2}(2\partial^i\psi^{(1)}\partial_i\psi^{(1)} + 3(1+w)\mathscr{H}^2 v^i_{(1)}v_{(1)i}) \\ &+ 3\nabla^{-4}\partial_i\partial^j(2\partial^i\psi^{(1)}\partial_j\psi^{(1)} + 3(1+w)\mathscr{H}^2 v^i_{(1)}v_{(1)j}) \ . \end{aligned} \tag{154}$$

  In particular in the case of a matter-dominated phase when $w = 0$ the linear gravitational potential $\psi^{(1)} = \phi^{(1)}$ is constant in time and, using Eq. (149), constraint (154) reads

$$\begin{aligned} \psi^{(2)} - \phi^{(2)} = &-\tfrac{2}{3}(\psi^{(1)})^2 + \tfrac{10}{3}\nabla^{-2}(\psi^{(1)}\nabla^2\psi^{(1)}) \\ &- 10\nabla^{-4}(\partial^i\partial_j(\psi^{(1)}\partial_i\partial^j\psi^{(1)})) \ . \end{aligned} \tag{155}$$

If we use the spatially flat gauge $\psi^{(1)} = \chi^{(1)} = 0$ and $\psi^{(2)} = \chi^{(2)} = 0$ we obtain for the $(0-0)$-component of Einstein equation on large scales

$$\phi^{(2)} = -\frac{1}{2}\frac{\delta^{(2)}\rho}{\rho_0} + 4(\phi^{(1)})^2 \ . \tag{156}$$

---

[16] From this section onward we do not use different symbols for the perturbations evaluated in different gauges, rather we will specify every time the gauge we are using.

[17] Such a constraint has first been derived for a Universe filled by a scalar field in Ref. [3].

4.3. Energy–momentum tensor conservation

We now derive the time-evolution on large scales of the gauge-invariant curvature perturbations ζ_i introduced in Section 3.5. Indeed the equations of motion for these quantities are a direct consequence of the energy continuity equation. In particular, we will focus on a system composed by a scalar field oscillating around the minimum of its potential and a radiation fluid having in mind the physical case of reheating. We will then describe the evolution of the curvature perturbations in the subsequent radiation/matter dominated phases.

Let us consider the system composed by the oscillating scalar field φ and the radiation fluid. Averaged over several oscillations the effective equation of state of the scalar field φ is $w_\varphi = \langle P_\varphi/\rho_\varphi\rangle = 0$, where P_φ and ρ_φ are the scalar field pressure and energy density, respectively. The scalar field is thus equivalent to a fluid of non-relativistic particles [283]. Moreover it is supposed to decay into radiation (light particles) with a decay rate Γ. We can thus describe this system as a pressureless and a radiation fluid which interact via energy transfer triggered by the decay rate Γ. We follow the gauge-invariant approach developed in Ref. [186] to study cosmological perturbations at first-order for the general case of an arbitrary number of interacting fluids and we shall extend the analysis to second order in the perturbations. Indeed the system under study encompasses the dynamics of the three main mechanisms for the generation of the primordial cosmological density perturbations on large scales, namely the standard scenario of single field inflation [101,174], the curvaton scenario [79,176,201,175], and the recently introduced scenario of "inhomogeneous reheating" [76,130,77,192,10]. Each component has energy–momentum tensor $T^{\mu\nu}_{(\varphi)}$ and $T^{\mu\nu}_{(\gamma)}$. The total energy momentum $T^{\mu\nu} = T^{\mu\nu}_{(\varphi)} + T^{\mu\nu}_{(\gamma)}$ is covariantly conserved

$$T^{\mu\nu}_{\ \ ;\mu} = 0 \ , \tag{157}$$

but allowing for an interaction between the two fluids [128]

$$\begin{aligned} T^{\mu\nu}_{(\varphi);\mu} &= Q^{\nu}_{(\varphi)} \ , \\ T^{\mu\nu}_{(\gamma);\mu} &= Q^{\nu}_{(\gamma)} \ , \end{aligned} \tag{158}$$

where $Q^{\nu}_{(\varphi)}$ and $Q^{\nu}_{(\gamma)}$ are the generic energy–momentum transfer coefficients for the scalar field and radiation sector, respectively, and are subject to the constraint

$$Q^{\nu}_{(\varphi)} + Q^{\nu}_{(\gamma)} = 0 \ , \tag{159}$$

derived from Eq. (157). The energy–momentum transfer $Q^{\nu}_{(\varphi)}$ and $Q^{\nu}_{(\gamma)}$ can be decomposed for convenience as [128]

$$\begin{aligned} Q^{\nu}_{(\varphi)} &= \hat{Q}_\varphi u^\nu + f^{\nu}_{(\varphi)} \ , \\ Q^{\nu}_{(\gamma)} &= \hat{Q}_\gamma u^\nu + f^{\nu}_{(\gamma)} \ , \end{aligned} \tag{160}$$

where the f^ν's are required to be orthogonal to the total velocity of the fluid u^ν. The energy continuity equations for the scalar field and radiation can be obtained from $u_\nu T^{\mu\nu}_{(\varphi);\mu} = u_\nu Q^{\nu}_{(\varphi)}$ and $u_\nu T^{\mu\nu}_{(\gamma);\mu} = u_\nu Q^{\nu}_{(\gamma)}$

and hence from Eq. (160)

$$u_\nu T^{\mu\nu}_{(\varphi);\mu} = \hat{Q}_\varphi \ ,$$
$$u_\nu T^{\mu\nu}_{(\gamma);\mu} = \hat{Q}_\gamma \ . \tag{161}$$

In the case of an oscillating scalar field decaying into radiation the energy transfer coefficient $\hat{Q}_\varphi$ is given by [103]

$$\hat{Q}_\varphi = -\Gamma\rho_\varphi \ ,$$
$$\hat{Q}_\gamma = \Gamma\rho_\varphi \ , \tag{162}$$

where Γ is the decay rate of the scalar field into radiation.

4.3.1. Background equations

The evolution of our spatially flat FRW background Universe is governed by the Friedmann constraint equation

$$\mathscr{H}^2 = \frac{8\pi G_{\rm N}}{3}\rho a^2 \ , \tag{163}$$

and by the energy continuity equation derived from Eq. (157)

$$\rho' = -3\mathscr{H}(\rho + P) \ , \tag{164}$$

where ρ and P are the total energy density and pressure of the system. The total energy density and the total pressure are related to the energy density and pressure of the scalar field and radiation by

$$\rho = \rho_\varphi + \rho_\gamma \ ,$$
$$P = P_\varphi + P_\gamma \ , \tag{165}$$

where P_γ is the radiation pressure. The energy continuity equations for the energy density of the scalar field ρ_φ and radiation ρ_γ in the background are

$$\rho'_\varphi = -3\mathscr{H}(\rho_\phi + P_\phi) + aQ_\varphi \ , \tag{166}$$
$$\rho'_\gamma = -4\mathscr{H}(\rho_\gamma + P_\gamma) + aQ_\gamma \ , \tag{167}$$

where Q_φ and Q_γ indicate the background values of the transfer coefficients $\hat{Q}_\varphi$ and $\hat{Q}_\gamma$, respectively.

4.4. Evolution of first-order curvature perturbations on large scales

The curvature perturbations $\zeta_i^{(1)}$ associated with the energy density of the scalar field and the radiation fluid are

$$\zeta^{(1)}_\varphi = -\hat{\psi}^{(1)} - \mathscr{H}\left(\frac{\delta^{(1)}\rho_\varphi}{\rho'_\varphi}\right) \ , \tag{168}$$

$$\zeta_\gamma^{(1)} = -\hat{\psi}^{(1)} - \mathscr{H}\left(\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma'}\right) . \tag{169}$$

Notice that the total curvature perturbation $\zeta^{(1)}$ can be expressed as a weighted sum of the single curvature perturbations of the scalar field and radiation fluid as [290,186]

$$\zeta^{(1)} = f\zeta_\varphi^{(1)} + (1-f)\zeta_\gamma^{(1)} \tag{170}$$

where

$$f = \frac{\rho_\varphi'}{\rho'}, \quad 1-f = \frac{\rho_\gamma'}{\rho'} \tag{171}$$

define the contribution of the scalar field and radiation to the total curvature perturbation $\zeta^{(1)}$, respectively. We now perturb at first order the continuity equations (161) for the scalar field and radiation energy densities, including the energy transfer. To this aim we first expand the transfer coefficients $\hat{Q}_\varphi$ and $\hat{Q}_\gamma$ up to first order in the perturbations around the homogeneous background as

$$\hat{Q}_\varphi = Q_\varphi + \delta^{(1)}Q_\varphi , \tag{172}$$

$$\hat{Q}_\gamma = Q_\gamma + \delta^{(1)}Q_\gamma . \tag{173}$$

Eqs. (161) give—on wavelengths larger than the horizon scale—

$$\begin{aligned}&\delta^{(1)}\rho_\varphi' + 3\mathscr{H}(\delta^{(1)}\rho_\varphi + \delta^{(1)}P_\varphi) - 3(\rho_\varphi + P_\varphi)\psi^{(1)'} \\ &\quad = aQ_\varphi\phi^{(1)} + a\delta^{(1)}Q_\varphi ,\end{aligned} \tag{174}$$

$$\begin{aligned}&\delta^{(1)}\rho_\gamma' + 3\mathscr{H}(\delta^{(1)}\rho_\gamma + \delta^{(1)}P_\gamma) - 3(\rho_\gamma + P_\gamma)\psi^{(1)'} \\ &\quad = aQ_\gamma\phi^{(1)} + a\delta^{(1)}Q_\gamma .\end{aligned} \tag{175}$$

Notice that the oscillating scalar field and radiation have fixed equations of state with $\delta^{(1)}P_\varphi = 0$ and $\delta^{(1)}P_\gamma = \delta^{(1)}\rho_\gamma/3$. This corresponds to vanishing intrinsic non-adiabatic pressure perturbations, as defined in Eq. (140).

Before proceeding further let us make a cautionary remark. In Eqs. (174)–(175) and in the following, as in Ref. [32], in the long-wavelength limit we are neglecting gradient terms which, upon integration over time, may give rise to non-local operators when evaluating second-order perturbations. However, these gradient terms will not affect statistical quantities in momentum-space, such as the gravitational potential bispectrum on large scales, as discussed in details in Section 8.

Following the procedure of Ref. [186] we can rewrite Eqs. (174) and (175) in terms of the gauge-invariant curvature perturbations $\zeta_\varphi^{(1)}$ and $\zeta_\gamma^{(1)}$

$$\zeta_\varphi^{(1)'} = \frac{a\mathscr{H}}{\rho_\varphi'}\left[\delta^{(1)}Q_\varphi - \frac{Q_\varphi'}{\rho_\varphi'}\delta^{(1)}\rho_\varphi + Q_\varphi\frac{\rho'}{2\rho}\left(\frac{\delta^{(1)}\rho_\varphi}{\rho_\varphi'} - \frac{\delta^{(1)}\rho}{\rho'}\right)\right] , \tag{176}$$

$$\zeta_\gamma^{(1)'} = \frac{a\mathscr{H}}{\rho_\gamma'}\left[\delta^{(1)}Q_\gamma - \frac{Q_\gamma'}{\rho_\gamma'}\delta^{(1)}\rho_\gamma + Q_\gamma\frac{\rho'}{2\rho}\left(\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma'} - \frac{\delta^{(1)}\rho}{\rho'}\right)\right] , \tag{177}$$

where we have used the perturbed $(0-0)$-component of Einstein equations for superhorizon wavelengths $\psi^{(1)'}+\mathscr{H}\phi^{(1)}=-\mathscr{H}/2\,\delta^{(1)}\rho/\rho$ (see Appendix A). Notice that from the constraint in Eq. (159) the perturbed energy transfer coefficients obey

$$\delta^{(1)}Q_\gamma=-\delta^{(1)}Q_\varphi\ . \tag{178}$$

4.4.1. Perturbations in the decay rate

If the energy transfer coefficients $\hat{Q}_\varphi$ and $\hat{Q}_\gamma$ are given in terms of the decay rate Γ as in Eq. (162), the first-order perturbation are, respectively,

$$\delta^{(1)}Q_\varphi=-\Gamma\delta^{(1)}\rho_\varphi-\delta^{(1)}\Gamma\rho_\varphi\ , \tag{179}$$

$$\delta^{(1)}Q_\gamma=\Gamma\delta^{(1)}\rho_\varphi+\delta^{(1)}\Gamma\rho_\varphi\ , \tag{180}$$

where notice in particular that we have allowed for a perturbation in the decay rate Γ,

$$\Gamma(\tau,\mathbf{x})=\Gamma(\tau)+\delta^{(1)}\Gamma(\tau,\mathbf{x})\ . \tag{181}$$

Perturbations in the inflaton decay rate are indeed the key feature of the "inhomogeneous reheating" scenario [76,130,77,10]. In fact from now on we shall consider the background value Γ of the decay rate as constant in time, $\Gamma\approx\Gamma_*$ as this is the case for the standard case of inflation and the inhomogeneous reheating mechanism. In such a case $\delta^{(1)}\Gamma$ is automatically gauge-invariant.[18] Plugging expressions (179)–(180) into Eqs. (176)–(177), and using Eq. (170), we find that the first order curvature perturbations for the scalar field and radiation obey on large-scales [33]

$$\zeta_\varphi^{(1)\prime}=\frac{a\Gamma}{2}\,\frac{\rho_\varphi}{\rho_\varphi'}\,\frac{\rho'}{\rho}(\zeta^{(1)}-\zeta_\varphi^{(1)})+a\mathscr{H}\,\frac{\rho_\varphi}{\rho_\varphi'}\,\delta^{(1)}\Gamma\ , \tag{182}$$

$$\zeta_\gamma^{(1)\prime}=-\frac{a}{\rho_\gamma'}\left[\Gamma\rho'\,\frac{\rho_\varphi'}{\rho_\gamma'}\left(1-\frac{\rho_\varphi}{2\rho}\right)(\zeta^{(1)}-\zeta_\varphi^{(1)})+\mathscr{H}\rho_\varphi\delta^{(1)}\Gamma\right]\ . \tag{183}$$

From Eq. (170) it is thus possible to find the equation of motion for the total curvature perturbation $\zeta^{(1)}$ using the evolution of the individual curvature perturbations in Eqs. (182) and (183)

$$\begin{aligned}\zeta^{(1)\prime}&=f'(\zeta_\varphi^{(1)}-\zeta_\gamma^{(1)})+f\zeta_\varphi^{(1)\prime}+(1-f)\zeta_\gamma^{(1)\prime}\\&=-\mathscr{H}f(\zeta^{(1)}-\zeta_\varphi^{(1)})\ .\end{aligned} \tag{184}$$

Notice that Eq. (184) can be rewritten as

$$\zeta^{(1)\prime}=\mathscr{H}f(1-f)(\zeta_\varphi^{(1)}-\zeta_\gamma^{(1)})=\frac{\mathscr{H}}{3}\,f(1-f)\mathscr{S}_{\varphi\gamma}\ , \tag{185}$$

[18] The authors of Ref. [192] have introduced a gauge-invariant generalization at first order in the case of $\Gamma'\neq 0$ which reads $\delta\Gamma^{(1)}_{\rm GI}=\delta^{(1)}\Gamma-\Gamma'\delta\rho_\varphi/\rho_\varphi'$. Indeed such a time-variation can have interesting effects on the overall curvature perturbation evolution $\zeta^{(1)\prime}$. See Ref. [192] for more details.

which explicitly shows that, in general, the total curvature perturbation can evolve on large scales due to a non-adiabatic pressure given by the relative entropy perturbation $\mathcal{S}_{\varphi\gamma}$ defined in Eq. (142). In fact by comparison with Eq. (144) the expression for the non-adiabatic pressure perturbation at first order reads

$$\delta^{(1)}P_{\rm nad} = -\frac{(3\rho_\varphi + 4\rho_\gamma)}{3} f(1-f)(\zeta_\varphi^{(1)} - \zeta_\gamma^{(1)}) \ . \tag{186}$$

From Eqs. (182) and (183) one can also obtain an equation of motion for the relative entropy perturbation, being $\mathcal{S}_{\varphi\gamma} = 3(\zeta_\varphi^{(1)} - \zeta_\gamma^{(1)})$, and use Eq. (185) to close the system of equations. However notice that during the decay of the scalar field into the radiation fluid, ρ_γ' may vanish and Eq. (183) for $\zeta_\gamma^{(1)}$, and hence, the evolution equation for $\mathcal{S}_{\varphi\gamma}$ become singular. Therefore it is convenient to close the system of equations by using the two first-order Eqs. (182) and (184) for the evolution of $\zeta_\varphi^{(1)}$ and $\zeta^{(1)}$.

In the scenarios for the generation of cosmological perturbations that we are going to study in detail, an adiabatic perturbation is produced after the "reheating phase" and thus the total curvature perturbation ζ is conserved on large scales during the radiation and matter-dominated phases, as it is evident from Eqs. (144) and (147) for a vanishing non-adiabatic pressure perturbation. In particular from the definition of the curvature perturbation at linear order $\zeta^{(1)} = -\hat{\psi}^{(1)} - \delta^{(1)}\rho/\rho'$ and using Eq. (151) in the Poisson gauge we determine

$$\psi^{(1)} = -\frac{3(1+w)}{5+3w}\,\zeta^{(1)} \ . \tag{187}$$

Such a relation is very useful to relate the gravitational potential $\psi^{(1)}$ during either the radiation or the matter-dominated epoch to the gauge-invariant curvature perturbation $\zeta^{(1)}$ at the end of the "reheating" phase. In fact, as we will see, in the case of standard single field inflation the perturbations are always adiabatic through the different phases and thus the curvature perturbation $\zeta^{(1)}$ remains always constant on superhorizon scales, so that we can write $\zeta^{(1)} = \zeta_I^{(1)}$, where the subscript "*I*" means that $\zeta^{(1)}$ is evaluated during the inflationary stage. On the other hand in the curvaton and in the inhomogeneous reheating scenarios the curvature perturbation $\zeta^{(1)}$ initially evolve on large scales due to a non-vanishing entropy perturbation, and thus the value of $\zeta^{(1)}$ during the radiation and matter-dominated phase will be determined by the curvature perturbation produced at the end of the "reheating" phase.

4.5. Evolution of second-order curvature perturbations on large scales

We now generalize to second order in the density perturbations the results of the previous section. In particular we obtain an equation of motion on large scales for the individual second-order curvature perturbations which include also the energy transfer between the scalar field and the radiation component [33].

Since the curvature perturbations $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$ are gauge-invariant, we choose to work in the spatially flat gauge $\psi^{(1)} = \chi^{(1)} = 0$ and $\psi^{(2)} = \chi^{(2)} = 0$ if not otherwise specified. Note that from Eqs. (168) and (169) $\zeta_\varphi^{(1)}$ and $\zeta_\gamma^{(1)}$ are thus given by

$$\zeta_\varphi^{(1)} = -\mathscr{H}\left(\frac{\delta^{(1)}\rho_\varphi}{\rho_\varphi'}\right) \ , \tag{188}$$

$$\zeta_\gamma^{(1)} = -\mathscr{H}\left(\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma'}\right) . \tag{189}$$

Eqs. (188)–(189) and the energy continuity equations at first order, Eqs. (174)–(175), in the spatially flat gauge $\psi^{(1)} = \chi^{(1)} = 0$ yield

$$\frac{\delta^{(1)}\rho'}{\rho'} = 3f\zeta_\varphi^{(1)} + 4(1-f)\zeta_\gamma^{(1)} , \tag{190}$$

$$\mathscr{H}\frac{\delta^{(1)}\rho}{\rho'} = -f\zeta_\varphi^{(1)} - (1-f)\zeta_\gamma^{(1)} . \tag{191}$$

We can thus rewrite the total second-order curvature perturbation $\zeta^{(2)}$ in Eq. (134) as

$$\begin{aligned}\zeta^{(2)} = & -\mathscr{H}\frac{\delta^{(2)}\rho}{\rho'} \\ & -[f\zeta_\varphi^{(1)} + (1-f)\zeta_\gamma^{(1)}][f^2\zeta_\varphi^{(1)} + (1-f)(2+f)\zeta_\gamma^{(1)}] ,\end{aligned} \tag{192}$$

where we have used the background continuity Eqs. (166)–(167) to find $\mathscr{H}\rho''/\rho' - \mathscr{H}' - 2\mathscr{H}^2 = -\mathscr{H}^2(6-f)$.

Following the same procedure, the individual curvature perturbations for the scalar field and the radiation fluid as defined in Eq. (135) are given by [33]

$$\begin{aligned}\zeta_\varphi^{(2)} = & -\mathscr{H}\frac{\delta^{(2)}\rho_\varphi}{\rho_\varphi'} + [2-3(1+w_\varphi)](\zeta_\varphi^{(1)})^2 - 2\left(a\frac{Q_\varphi\phi^{(1)}}{\rho_\varphi'} + a\frac{\delta^{(1)}Q_\varphi}{\rho_\varphi'}\right)\zeta_\varphi^{(1)} \\ & -\left[a\frac{Q_\varphi'}{\mathscr{H}\rho_\varphi'} - \frac{a}{2}\frac{Q_\varphi}{\mathscr{H}\rho_\varphi'}\frac{\rho'}{\rho}\right](\zeta_\varphi^{(1)})^2 ,\end{aligned} \tag{193}$$

$$\begin{aligned}\zeta_\gamma^{(2)} = & -\mathscr{H}\frac{\delta^{(2)}\rho_\gamma}{\rho_\gamma'} + [2-3(1+w_\gamma)](\zeta_\gamma^{(1)})^2 - 2\left(a\frac{Q_\gamma\phi^{(1)}}{\rho_\gamma'} + a\frac{\delta^{(1)}Q_\gamma}{\rho_\gamma'}\right)\zeta_\gamma^{(1)} \\ & -\left[a\frac{Q_\gamma'}{\mathscr{H}\rho_\gamma'} - \frac{a}{2}\frac{Q_\gamma}{\mathscr{H}\rho_\gamma'}\frac{\rho'}{\rho}\right](\zeta_\gamma^{(1)})^2 ,\end{aligned} \tag{194}$$

where $w_\gamma = \frac{1}{3}$ is the radiation equation of state. Using Eqs. (193) and (194) to express the perturbation of the total energy density $\delta^{(2)}\rho$ one obtains the following expression for the total curvature perturbation $\zeta^{(2)}$ [33]:

$$\begin{aligned}\zeta^{(2)} = & f\zeta_\varphi^{(2)} + (1-f)\zeta_\gamma^{(2)} + f(1-f)(1+f)(\zeta_\varphi^{(1)} - \zeta_\gamma^{(1)})^2 \\ & + 2\left(a\frac{Q_\varphi\phi^{(1)}}{\rho'} + a\frac{\delta^{(1)}Q_\varphi}{\rho'}\right)[\zeta_\varphi^{(1)} - \zeta_\gamma^{(1)}] \\ & + \left(a\frac{Q_\varphi'}{\mathscr{H}\rho'} - \frac{a}{2}\frac{Q_\varphi}{\mathscr{H}\rho}\right)[(\zeta_\varphi^{(1)})^2 - (\zeta_\gamma^{(1)})^2] .\end{aligned} \tag{195}$$

Expressing the $(0-0)$-component of Einstein equations (152) in the spatially flat gauge at first-order $\psi^{(1)}=\chi^{(1)}=0$ in terms of the total curvature $\zeta^{(1)}$

$$\phi^{(1)}=-\frac{1}{2}\frac{\delta^{(1)}\rho}{\rho}=\frac{1}{2}\frac{\rho'}{\mathscr{H}\rho}\zeta^{(1)}\ , \tag{196}$$

and using the explicit expressions for the first-order perturbed coefficients in terms of the decay rate Γ, Eqs. (179)–(180), we finally obtain [33]

$$\begin{aligned}\zeta^{(2)}&=f\zeta_\varphi^{(2)}+(1-f)\zeta_\gamma^{(2)}+f(1-f)(1+f)(\zeta_\varphi^{(1)}-\zeta_\gamma^{(1)})^2\\&\quad+\frac{a\Gamma}{\mathscr{H}}f(\zeta_\varphi^{(1)}-\zeta_\gamma^{(1)})^2-2a\delta^{(1)}\Gamma\frac{\rho_\varphi}{\rho'}(\zeta_\varphi^{(1)}-\zeta_\gamma^{(1)})\\&\quad+\frac{a\Gamma}{\mathscr{H}}(1-2f)\frac{\rho_\varphi}{2\rho}(\zeta_\varphi^{(1)}-\zeta_\gamma^{(1)})^2\ .\end{aligned} \tag{197}$$

Eq. (197) is an important result. It generalizes to second order in the perturbations the weighted sum in Eq. (170) and extends the expression found in Ref. [32] in the particular case of the curvaton scenario, under the sudden-decay approximation, where the energy transfer was neglected. Similarly to linear order such an expression will be useful to describe the large-scale evolution of $\zeta^{(2)}$ sourced by a non-adiabatic pressure perturbation through the evolution of the density perturbations in the scalar field and radiation.

As already mentioned in the previous section, the expressions for the second-order perturbations have been found here by using the long-wavelength limit for the first-order perturbations. Indeed second-order quantities expressed in terms of first-order perturbations will depend also on the short-wavelength behavior of the first-order perturbations, as it is evident going to momentum space. Thus, for example, even if $\zeta_i^{(1)}$ are constant on large scales at linear order, it would not be strictly correct to consider the second-order part of ζ depending on $\zeta_i^{(1)}$ as constant. However our procedure is fully justified when applied to the evaluation of the bispectrum on superhorizon scales as we shall discuss in Section 8.

Let us now give the equations of motion on large scales for the individual second-order curvature perturbations $\zeta_\varphi^{(2)}$ and $\zeta_\gamma^{(2)}$. The energy transfer coefficients $\hat{Q}_\varphi$ and $\hat{Q}_\gamma$ in Eqs. (160) perturbed at second order around the homogeneous backgrounds are given by

$$\hat{Q}_\varphi=Q_\varphi+\delta^{(1)}Q_\varphi+\tfrac{1}{2}\delta^{(2)}Q_\varphi\ , \tag{198}$$

$$\hat{Q}_\gamma=Q_\gamma+\delta^{(1)}Q_\gamma+\tfrac{1}{2}\delta^{(2)}Q_\gamma\ . \tag{199}$$

Note that from Eq. (159) it follows that $\delta^{(2)}Q_\gamma=-\delta^{(2)}Q_\gamma$. Thus the energy continuity equations (161) perturbed at second order give on large scales [33]

$$\begin{aligned}&\delta^{(2)}\rho_\varphi'+3\mathscr{H}(\delta^{(2)}\rho_\varphi+\delta^{(2)}P_\varphi)-3(\rho_\varphi+P_\varphi)\psi^{(2)'}\\&\quad-6\psi^{(1)'}[\delta^{(1)}\rho_\varphi+\delta^{(1)}P_\varphi+2(\rho_\varphi+P_\varphi)\psi^{(1)}]\\&=a\delta^{(2)}Q_\varphi+aQ_\varphi\phi^{(2)}-aQ_\varphi(\phi^{(1)})^2+2a\phi^{(1)}\delta^{(1)}Q_\varphi\ ,\end{aligned} \tag{200}$$

$$\begin{aligned}&\delta^{(2)}\rho_\gamma'+3\mathscr{H}(\delta^{(2)}\rho_\gamma+\delta^{(2)}P_\gamma)-3(\rho_\gamma+P_\gamma)\psi^{(2)'}\\&\quad-6\psi^{(1)'}[\delta^{(1)}\rho_\gamma+\delta^{(1)}P_\gamma+2(\rho_\gamma+P_\gamma)\psi^{(1)}]\\&=a\delta^{(2)}Q_\gamma+aQ_\gamma\phi^{(2)}-aQ_\gamma(\phi^{(1)})^2+2a\phi^{(1)}\delta^{(1)}Q_\gamma\ ,\end{aligned} \tag{201}$$

where $\phi^{(2)}$ is the second-order perturbation in the gravitational potential $\phi = \phi^{(1)} + \frac{1}{2}\phi^{(2)}$. Note that Eqs. (200) and (201) hold true in a generic gauge. We can now recast these equations in terms of the gauge-invariant curvature perturbations $\zeta_\varphi^{(2)}$ and $\zeta_\gamma^{(2)}$ in a straightforward way by choosing the spatially flat gauge $\psi^{(r)} = \chi^{(r)} = 0$.

The $(0-0)$-component of Einstein equations in the spatially flat gauge at first order is given by Eq. (196), and at second order on large-scales it reads

$$\phi^{(2)} = -\frac{1}{2}\frac{\delta^{(2)}\rho}{\rho} + 4(\phi^{(1)})^2 \ . \tag{202}$$

Using Eqs. (196) and (202) with expressions (193)–(194) we find from the energy continuity equations (200)–(201) that the individual second-order curvature perturbations obey on large-scales [33]

$$\begin{aligned}\zeta_\varphi^{(2)\prime} = & -\frac{a\mathscr{H}}{\rho'}\left[\left(\delta^{(2)}Q_\varphi - \frac{Q_\varphi'}{\rho_\varphi'}\delta^{(2)}\rho_\varphi\right) + Q_\varphi\frac{\rho'}{2\rho}\left(\frac{\delta^{(2)}\rho_\varphi}{\rho_\varphi'} - \frac{\delta^{(2)}\rho}{\rho'}\right)\right] \\ & - 3aQ_\varphi\frac{\mathscr{H}}{\rho_\varphi'}(\phi^{(1)})^2 - 2a\frac{\mathscr{H}}{\rho_\varphi'}\delta^{(1)}Q_\varphi\phi^{(1)} - 2\zeta_\varphi^{(1)}\zeta_\varphi^{(1)\prime} \\ & - 2\left[\zeta_\varphi^{(1)}\left(a\frac{Q_\varphi\phi^{(1)}}{\rho_\varphi'} + a\frac{\delta^{(1)}Q_\varphi}{\rho_\varphi'}\right)\right]' - \left[(\zeta_\varphi^{(1)})^2\left(a\frac{Q_\varphi'}{\mathscr{H}\rho_\varphi'} - \frac{a}{2}\frac{Q_\varphi}{\mathscr{H}\rho_\varphi'}\frac{\rho'}{\rho}\right)\right]'\end{aligned} \tag{203}$$

and

$$\begin{aligned}\zeta_\gamma^{(2)\prime} = & -\frac{a\mathscr{H}}{\rho'}\left[\left(\delta^{(2)}Q_\gamma - \frac{Q_\gamma'}{\rho_\gamma'}\delta^{(2)}\rho_\gamma\right) + Q_\gamma\frac{\rho'}{2\rho}\left(\frac{\delta^{(2)}\rho_\gamma}{\rho_\gamma'} - \frac{\delta^{(2)}\rho}{\rho'}\right)\right] \\ & - 3aQ_\gamma\frac{\mathscr{H}}{\rho_\gamma'}(\phi^{(1)})^2 - 2a\frac{\mathscr{H}}{\rho_\gamma'}\delta^{(1)}Q_\gamma\phi^{(1)} - 4\zeta_\gamma^{(1)}\zeta_\gamma^{(1)\prime} \\ & - 2\left[\zeta_\gamma^{(1)}\left(a\frac{Q_\gamma\phi^{(1)}}{\rho_\gamma'} + a\frac{\delta^{(1)}Q_\gamma}{\rho_\gamma'}\right)\right]' - \left[\left(\zeta_\gamma^{(1)}\right)^2\left(a\frac{Q_\gamma'}{\mathscr{H}\rho_\gamma'} - \frac{a}{2}\frac{Q_\gamma}{\mathscr{H}\rho_\gamma'}\frac{\rho'}{\rho}\right)\right]' ,\end{aligned} \tag{204}$$

where we have used the fact that $w_\varphi = 0$ and $w_\gamma = 1/3$.

Eqs. (203) and (204) allow to follow the time-evolution of the gauge-invariant curvature perturbations at second order.

The results contained in the previous section can now be used to study the evolution of the second-order curvature perturbations during the reheating phase after a period of standard single field inflation, and in the alternative scenarios for the generation of the primordial adiabatic perturbations which have been recently proposed, namely the curvaton scenario [79,176,201] and the inhomogeneous (or "modulated") reheating [76,130,77,192,10]. In fact, in each of these scenarios a scalar field oscillates around the minimum of its potential and eventually decays into radiation. The evolution at second order of the curvature perturbations is necessary in order to follow the non-linearity of the cosmological perturbations and thus to accurately compute the level of non-Gaussianity, including all the relevant second-order effects. We shall now consider in detail these contributions in the three mentioned scenarios.

5. The standard scenario

The *standard scenario* is associated to one-single-field models of inflation, and the observed density perturbations are due to fluctuations of the inflaton field itself. When inflation ends, the inflaton oscillates about the minimum of its potential and decays, thereby reheating the Universe. The initial inflaton fluctuations are adiabatic on large-scales and are transferred to the radiation fluid during reheating. In such a standard scenario the inflaton decay rate has no spatial fluctuations.

5.1. The first-order curvature perturbation

During inflation the inflaton field dominates the energy density of the Universe and therefore the energy density perturbations produced by the inflaton quantum fluctuations generate an adiabatic curvature perturbation. Let us consider the inflaton field $\varphi(\tau, \mathbf{x})$ with a potential $V(\varphi)$ and minimally coupled to gravity. The evolution equation for the inflaton field is the Klein–Gordon equation

$$\Box\varphi = \frac{\partial V}{\partial\varphi} \ . \tag{205}$$

Perturbing Eq. (205) at linear order we obtain that the inflaton fluctuations obey

$$\delta^{(1)}\varphi'' + 2\mathscr{H}\delta^{(1)}\varphi' - \nabla^2\delta^{(1)}\varphi + a^2\delta^{(1)}\varphi\,\frac{\partial^2 V}{\partial\varphi^2}\,a^2 + 2\phi^{(1)}\,\frac{\partial V}{\partial\varphi} - \varphi_0'[\phi^{(1)\prime} + 3\psi^{(1)\prime} + \nabla^2\omega^{(1)}] = 0 \ . \tag{206}$$

A straightforward way to calculate the curvature perturbation generated on large scales is to solve the Klein–Gordon equation in the spatially flat gauge defined by the requirement $\psi^{(1)} = 0$ and $\chi^{(1)} = 0$. In fact in this gauge the perturbations of the scalar field correspond to the Sasaki–Mukhanov gauge-invariant variables [253,205]

$$Q_\varphi = \delta^{(1)}\varphi + \frac{\varphi'}{\mathscr{H}}\,\hat{\psi}^{(1)} \ . \tag{207}$$

As usual we introduce the field $\widetilde{Q}_\varphi = a\,Q_\varphi$. The Klein–Gordon equation in the spatially flat gauge now reads (in Fourier space) [278]

$$\widetilde{Q}_\varphi'' + \left(k^2 - \frac{a''}{a} + \mathscr{M}_\varphi^2 a^2\right)\tilde{Q}_\varphi = 0 \ , \tag{208}$$

where

$$\mathscr{M}_\varphi^2 = V_{\varphi\varphi} - \frac{8\pi G_{\mathrm{N}}}{a^3}\left(\frac{a^3}{H}\dot{\varphi}^2\right) \tag{209}$$

is an effective mass of the inflaton field in this gauge, which to lowest order in the slow-roll parameters is given by

$$\frac{\mathscr{M}_\varphi^2}{H^2} = 3\eta - 6\epsilon \ , \tag{210}$$

where $\epsilon = (1/16\pi G_{\rm N})(V_\varphi/V)^2$ and $\eta = (1/8\pi G_{\rm N})(V_{\varphi\varphi}/V)$ are the inflaton slow-roll parameters. This equation has the same form as Eq. (41) and thus we can just follow the same procedure described in detail in Section 2.3, simply replacing m_χ^2 with $\mathscr{M}_\varphi^2$. The equation of motion for $\widetilde{Q}_\varphi$ or for the corresponding eigenvalues $u_k(\tau)$ thus becomes

$$u_k'' + \left(k^2 - \frac{\nu_\varphi^2 - \frac{1}{4}}{\tau^2}\right) u_k = 0 \ , \tag{211}$$

with

$$\nu_\varphi \simeq \tfrac{3}{2} + 3\epsilon - \eta \ . \tag{212}$$

From Section 2.3 we conclude that on superhorizon scales and to lowest order in the slow-roll parameters the inflaton fluctuations are

$$|Q_\varphi(k)| = \frac{H}{\sqrt{2k^3}}\left(\frac{k}{aH}\right)^{3/2-\nu_\varphi} \ . \tag{213}$$

In order to calculate the curvature perturbation on large scales we can consider the curvature perturbation on comoving hypersurfaces, which in the case of a single scalar field reads [170,173,159,174]

$$\mathscr{R}^{(1)} = \hat{\psi}^{(1)} + \frac{\mathscr{H}}{\varphi'}\,\delta^{(1)}\varphi \ . \tag{214}$$

Notice that the comoving curvature perturbation $\mathscr{R}^{(1)}$ and the uniform energy density curvature perturbation $\zeta^{(1)}$ are simply related by (see for example Ref. [93])

$$-\zeta^{(1)} = \mathscr{R}^{(1)} + \frac{2\rho}{9(\rho+p)}\left(\frac{k}{aH}\right)^2 \phi^{(1)} \ , \tag{215}$$

where here $\phi^{(1)}$ is the gravitational potential in the longitudinal gauge. Therefore on large scales $\mathscr{R}^{(1)} \simeq -\zeta^{(1)}$. From Eq. (207) it is evident that

$$\mathscr{R}^{(1)} = \frac{\mathscr{H}}{\varphi'} Q_\varphi \ . \tag{216}$$

Thus we obtain the power-spectrum of the curvature perturbation on large scales

$$\mathscr{P}_\mathscr{R} = \left(\frac{H^2}{2\pi\dot{\varphi}}\right)^2 \left(\frac{k}{aH}\right)^{3-2\nu_\varphi} \simeq \left(\frac{H^2}{2\pi\dot{\varphi}}\right)_*^2 , \tag{217}$$

where the asterisk stands for the epoch a given perturbation mode leaves the horizon during inflation. From Eq. (217) one immediately reads the spectral index of the curvature perturbation to lowest order in the slow-roll parameters

$$n_\mathscr{R} - 1 \equiv \frac{{\rm d}\,\ln \mathscr{P}_\mathscr{R}}{{\rm d}\,\ln k} = 3 - 2\nu_\varphi = -6\epsilon + 2\eta \ . \tag{218}$$

Notice that from our results one can check that during inflation the curvature perturbation mode is constant on superhorizon scales $\mathscr{R}^{(1)\prime} \simeq -\zeta^{(1)\prime} \simeq 0$ (from which the last equality in Eq. (217) follows).

This is a well-known result: the curvature mode is the quantity which allows to connect observable perturbations to primordial perturbations produced during inflation [22,156,174]. This result comes from the fact that in single-field slow-roll models of inflation the intrinsic entropy perturbation of the inflaton field is negligible on large scales [174,35,93,27]. We will now show that the above result also holds during the reheating phase on large scales.

5.2. Reheating after inflation

When inflation ends, the inflaton oscillates about the minimum of its potential and decays into radiation, thereby reheating the Universe. In such a standard scenario the inflaton decay rate has no spatial fluctuations. Eqs. (184) and (182) with $\delta^{(1)}\Gamma = 0$ now read

$$\zeta^{(1)'} = -\mathscr{H} f(\zeta^{(1)} - \zeta_{\varphi}^{(1)}) \,, \tag{219}$$

$$\zeta_{\varphi}^{(1)'} = \frac{a\Gamma}{2}\frac{\rho_{\varphi}}{\rho_{\varphi}'}\frac{\rho'}{\rho}(\zeta^{(1)} - \zeta_{\varphi}^{(1)}) \,. \tag{220}$$

At the beginning of the reheating phase, after the end of inflation, the total curvature perturbation is initially given by the curvature perturbation of the inflaton fluctuations $\zeta_{\rm in}^{(1)} = \zeta_{\varphi,{\rm in}}^{(1)}$. Therefore Eqs. (219) and (220) show that during the reheating phase $\zeta^{(1)} = \zeta_{\varphi}^{(1)} = \zeta_{\varphi,{\rm in}}^{(1)}$ is a fixed-point of the time-evolution. Such a result has been obtained in this way at first order in Ref. [192] (see also Refs. [127,103]) and extended to second order in the perturbations in Ref. [31], under the sudden-decay approximation.

5.3. The second-order curvature perturbation and non-Gaussianity during inflation

A complete analysis of the perturbations produced during single-field slow-roll inflation up to second order has been performed in Ref. [3]. Such an analysis fully accounts for the inflaton self-interactions as well as for the second-order fluctuations of the background metric. Moreover it also provides a gauge-invariant expression for the second-order comoving curvature perturbation thus allowing to calculate the bispectrum of such a quantity during inflation. The results of Ref. [3] agree with those of Ref. [182], where the three-point function for the curvature perturbation is calculated using a different procedure. In Ref. [182] the starting point is the Lagrangian and one evaluates the cubic contributions to the curvature perturbations. In fact Refs. [3,182] represent a step forward in the computation of the non-linearities produced during inflation. Before then, the problem of calculating the bispectrum of perturbations produced during inflation had been addressed by either looking at the effect of inflaton self-interactions (which necessarily generate non-linearities in its quantum fluctuations) in a fixed de Sitter background [82], or using the so-called stochastic approach to inflation [86],[19] where back-reaction effects of field fluctuations on the background metric are partially taken into account. An intriguing result of the stochastic approach—which is indeed confirmed by the second-order analyses of Refs. [3,182]—is that the dominant source of non-Gaussianity actually comes from non-linear gravitational perturbations, rather than by inflaton self-interactions.

Before going into details, let us give here an estimate of the size of the non-Gaussianity that we expect to be produced during inflation.

[19] See also Refs. [291,87,189], for a more recent analysis.

The inflaton field can be split into a homogeneous background $\varphi_0(\tau)$ and a perturbation $\delta\varphi(\tau, x^i)$ as

$$\varphi(\tau, x^i) = \varphi_0(\tau) + \delta\varphi(\tau, x^i) = \varphi_0(\tau) + \delta^{(1)}\varphi(\tau, x^i) + \tfrac{1}{2}\,\delta^{(2)}\varphi(\tau, x^i)\ , \tag{221}$$

where the perturbation has been expanded into a first and a second-order part, respectively. First of all notice that, at first order in the perturbations, using Eq. (149) for a scalar field in the longitudinal gauge ($\dot{\psi}^{(1)} + H\psi^{(1)} = \kappa^2/2\,\dot{\varphi}_0\delta^{(1)}\varphi$) and the perturbed Klein–Gordon equation (B.4) one obtains $\psi^{(1)} = \epsilon H\delta^{(1)}\varphi/\dot{\varphi}_0$ to lowest order in the slow-roll parameters and on large scales. On the other hand, from the definition of the comoving curvature perturbation at first order, Eq. (214), it follows that $\mathscr{R}^{(1)} = H\delta^{(1)}\varphi/\dot{\varphi}_0$ to lowest order in the slow-roll parameters and on large scales, and hence under these approximations $\psi^{(1)} = \epsilon\mathscr{R}^{(1)}$. Let us consider the perturbed Klein–Gordon equation at second order (in the Poisson gauge) on large scales (see Eq. (B.5) in Appendix B)

$$\begin{aligned}&\delta^{(2)\ddot{}}\varphi + 3H\delta^{(2)\dot{}}\varphi + 2\,\frac{\partial V}{\partial\varphi}\,\phi^{(2)} - \dot{\varphi}_0\dot{\phi}^{(2)} - 3\dot{\varphi}_0\dot{\psi}^{(2)} - 8\dot{\varphi}_0\psi^{(1)}\dot{\psi}^{(1)} - 8\dot{\psi}^{(1)}\delta^{(1)\dot{}}\varphi \\ &\quad - \frac{8}{a^2}\,\psi^{(1)}\partial_i\partial^i\delta^{(1)}\varphi = -4\,\frac{\partial^{(2)}V}{\partial\varphi^2}\,\psi^{(1)}\delta^{(1)}\varphi - \frac{\partial^2 V}{\partial\varphi^2}\,\delta^{(2)}\varphi - \frac{\partial^3 V}{\partial\varphi^3}\,(\delta^{(1)}\varphi)^2\ .\end{aligned} \tag{222}$$

Now in order to give our estimate we consider a second-order curvature perturbation $\mathscr{R}^{(2)} \sim H\delta^{(2)}\varphi/\dot{\varphi}_0$. For simplicity, let us just focus on the first source term on the R.H.S. of Eq. (222). Under a slow-roll approximation from Eq. (222) we see that (up to numerical coefficients of order unity)

$$\dot{\mathscr{R}}^{(2)} \sim \left(\frac{\partial^2 V}{\partial\varphi^2}\right)\left(\frac{\psi^{(1)}}{H}\right)\left(H\,\frac{\delta^{(1)}\varphi}{\dot{\varphi}_0}\right) \sim H\eta\epsilon(\mathscr{R}^{(1)})^2\ , \tag{223}$$

where in the last step we have used $\psi^{(1)} = \epsilon\mathscr{R}^{(1)}$ with $\mathscr{R}^{(1)} \sim H\delta^{(1)}\varphi/\dot{\varphi}_0$ and the definition of the slow-roll parameters. Recalling that the time derivatives of the slow-roll parameters are next order in the parameters, $\dot{\epsilon}, \dot{\eta} = \mathcal{O}(\epsilon^2, \eta^2)$, from Eq. (223) we obtain

$$\mathscr{R}^{(2)} \sim \mathcal{O}(\epsilon, \eta)(\mathscr{R}^{(1)})^2\ . \tag{224}$$

From this simple calculation we therefore see that the non-Gaussianity level in the standard scenario is of the order of the slow-roll parameters ϵ and η

$$\mathrm{NG} \sim \mathcal{O}(\epsilon, \eta), \tag{225}$$

in qualitative agreement with the predictions of Refs. [86,85], within the stochastic inflation approach.

Let us now turn to the exact results on the level of non-Gaussianity by summarizing some of the findings of Ref. [3]. It is possible to extend at second order the gauge-invariant large-scale comoving curvature perturbation $\mathscr{R}^{(1)}$ defined in Eq. (214) by introducing a quantity $\mathscr{R} = \mathscr{R}^{(1)} + \frac{1}{2}\mathscr{R}^{(2)}$ defined as [3]

$$\mathscr{R} = \mathscr{R}^{(1)} + \frac{1}{2}\left[\mathscr{H}\,\frac{\delta^{(2)}\varphi}{\varphi_0'} + \psi^{(2)}\right] + \frac{1}{2}\,\frac{(\psi^{(1)\prime} + 2\mathscr{H}\psi^{(1)} + \mathscr{H}\delta^{(1)}\varphi'/\varphi_0')^2}{\mathscr{H}' + 2\mathscr{H}^2 - \mathscr{H}\varphi_0''/\varphi_0'}\ . \tag{226}$$

Such a quantity is gauge-invariant with respect to an infinitesimal second-order shift of the time coordinate, $\tau \to \tau - \xi^0_{(1)} + \frac{1}{2}(\xi^{0\prime}_{(1)}\xi^0_{(1)} - \xi^0_{(2)})$.

By solving the Einstein equations during inflation in the longitudinal (Poisson) gauge on large scales, and by performing an expansion to lowest order in the slow-roll parameters the second-order curvature perturbation is determined in terms of its first-order counterpart. The result is [3]

$$\mathscr{R}^{(2)} = (\eta - 3\epsilon)(\mathscr{R}^{(1)})^2 + \mathscr{I} \ , \tag{227}$$

where

$$\begin{aligned}\mathscr{I} = & -\frac{2}{\epsilon}\int \frac{1}{a^2}\,\psi^{(1)}\nabla^2\psi^{(1)}\,\mathrm{d}t - \frac{4}{\epsilon}\int \frac{1}{a^2}\,(\partial_i\psi^{(1)}\partial^i\psi^{(1)})\,\mathrm{d}t \\ & -\frac{4}{\epsilon}\int (\ddot{\psi}^{(1)})^2\,\mathrm{d}t + (\epsilon-\eta)\triangle^{-1}\partial_i R^{(1)}\partial^i R^{(1)}\end{aligned} \tag{228}$$

contains also terms which are $\mathscr{O}(\epsilon,\eta)$. Notice that the integrals in Eq. (228) give rise to non-local operators which are not necessarily suppressed on large scales being of the form $\nabla^{-2}[\nabla(\cdot)\nabla(\cdot)]$ or $\nabla^{-2}[(\cdot)\nabla^2(\cdot)]$. This is due to the fact that a given perturbation mode during inflation first is subhorizon, where it oscillates, and then at a given epoch it leaves the horizon.

The total comoving curvature perturbation thus receives a contribution which is quadratic in $\mathscr{R}^{(1)}$ and it will then have a non-Gaussian (χ^2) component. We conclude that during inflation a tiny intrinsic non-linearity is produced, being the slow-roll parameters $\epsilon, |\eta| \ll 1$. This does not come as a surprise, indeed, and it has a very transparent interpretation. Since the inflaton field is driving inflation, its potential must be very flat, with very small slow-roll parameters. This amounts to saying that the interaction terms in the inflaton potential must be suppressed, hence also the non-linearities eventually producing non-Gaussian features. Alternative mechanisms to generate a higher level of non-Gaussian adiabatic perturbations in the inflationary framework could be the presence of some features in the inflaton potential [249,131,248,291] in the part corresponding to the last ~ 60 e-foldings, or the presence of more than a single scalar field during inflation [30]. In both cases the restrictions coming from the slow-roll conditions can be avoided. We shall come back later to these alternative scenarios. In the case of the standard single-field models of inflation in order to characterize the primordial non-Gaussianity we can expand $\mathscr{R}$ in Fourier space as

$$\begin{aligned}\mathscr{R}(\mathbf{k}) = & \mathscr{R}^{(1)}(\mathbf{k}) + \frac{1}{(2\pi)^3}\int \mathrm{d}^3k_1\,\mathrm{d}^3k_2\,\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}) \\ & \times f_{\mathrm{NL}}^{\mathscr{R}}(\mathbf{k}_1,\mathbf{k}_2)\mathscr{R}^{(1)}(\mathbf{k}_1)\mathscr{R}^{(1)}(\mathbf{k}_2) \ ,\end{aligned} \tag{229}$$

where we have introduced a momentum-dependent non-linearity parameter $f_{\mathrm{NL}}^{\mathscr{R}}(\mathbf{k}_1,\mathbf{k}_2)$, which from Eq. (227) reads

$$f_{\mathrm{NL}}^{\mathscr{R}}(\mathbf{k}_1,\mathbf{k}_2) = \tfrac{1}{2}(\eta - 3\epsilon) + I(\mathbf{k}_1,\mathbf{k}_2) \ , \tag{230}$$

where $I(\mathbf{k}_1,\mathbf{k}_2)$ is directly related to the function $\mathscr{I}$ and is of first order in the slow-roll parameters. Thus the level of non-Gaussianity generated during inflation is typically $f_{\mathrm{NL}}^{\mathscr{R}} \sim \mathscr{O}(10^{-1}-10^{-2})$. Eq. (230) can also be recast in the form

$$f_{\mathrm{NL}}^{\mathscr{R}}(\mathbf{k}_1,\mathbf{k}_2) = \tfrac{1}{4}(n_{\mathscr{R}} - 1) + I(\mathbf{k}_1,\mathbf{k}_2) \ , \tag{231}$$

where we have made use of the expression of the spectral index $n_{\mathscr{R}} - 1 = -6\epsilon + 2\eta$ in terms of the slow-roll parameters. Notice however that the result in Eq. (230) refers only to the non-Gaussianity generated *during inflation*.[20]

In order to determine the level of the non-Gaussianity which can be actually compared with observations it is necessary to consider the subsequent evolution of the (gravitational potential) perturbations after inflation ends, through reheating and the radiation/matter dominated epochs. Usually in the literature the matching has been performed by simply extending the linear relation on large-scales between the gravitational potential and the curvature perturbation in the matter dominated epoch $\phi^{(1)} = -(3/5)\mathscr{R}^{(1)}$ to second order [82,86,3,182]. We warn the reader that such a procedure is indeed not correct. In fact one has to take into account also after inflation a fully second-order relativistic analysis of the cosmological perturbations. As we shall consider in the next sections, in the case of the standard scenario of inflation the matching between the inflationary epoch and the radiation/matter dominate phases, where observable quantities are defined, is achieved by exploiting the conservation on large scales of the curvature perturbation ζ up to second order. As a result, as it has been shown in Ref. [31], the small initial non-Gaussianity generated during single-field inflation is actually largely *enhanced* by the second-order gravitational dynamics in the post-inflationary phases. Such an enhancement produces a non-linearity parameter in the CMB temperature anisotropy on large scales which is $f_{\rm NL} \sim \mathcal{O}(1)$. Nonetheless, the results contained in Refs. [3,182] are useful in that they allow to determine the initial conditions on the non-Gaussianity produced during slow-roll inflation.

5.3.1. *Inflaton effective Lagrangian*

In this paragraph we want to mention the possibility discussed in Ref. [72] that the non-Gaussianity produced during inflation might receive additional contributions from some high-energy corrections which can modify the minimal inflaton Lagrangian given in Eq. (18). Such corrections can arise if inflation takes place at relatively high energies, and they can be parametrized by an effective inflaton Lagrangian in which one integrates out degrees of freedom with momenta larger than some scale M corresponding to the scale of the new physics. This is realized by including some higher order operators suppressed by the appropriate power of M. Such operators must not spoil the flatness of the potential in order to have an inflationary phase, and, as argued in Ref. [72], the best candidates are operators that just modify the kinetic part of the action such as those of the form $(\nabla\varphi)^4$ with a scale M which can be taken as low as $\dot{\varphi}^2$. Therefore a possible effective action reads

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[-\frac{1}{2}(\nabla\varphi)^2 - V(\varphi) + \frac{1}{8M^4}(\nabla\varphi)^2(\nabla\varphi)^2 + \cdots\right]. \tag{232}$$

Notice that when M^2 tends to $\dot{\varphi}$ then such an effective description ceases to make sense, because one should keep track of all the higher terms in the action. The higher dimension operators represent additional self-interaction terms which will produce some non-Gaussianities during inflation. In fact starting from action (232), following Ref. [182], it is possible to calculate the contributions to the three-point function

[20] The generalization of the calculation of the non-linearity parameter during inflation to two-field models of inflation has been recently presented in Ref. [80].

for the curvature perturbation $\mathscr{R}$ coming from the higher dimension operators to lowest order in $\dot{\varphi}^2/M^4$. In Ref. [72] it has been found that the typical magnitude of such contributions is[21]

$$f_{\rm NL}^{\mathscr{R}} \sim \frac{\dot{\varphi}^2}{M^4} \ . \tag{233}$$

We see that the net effect of the introduction of a new scale M is that the slow-roll parameters in the standard result (230) are now replaced by a new order parameter which does not have to be restricted by slow-roll conditions. However the lower limit allowed for M^4 is $\dot{\varphi}^2$ and therefore also in this case

$$f_{\rm NL}^{\mathscr{R}} \lesssim 1 \ . \tag{234}$$

We must stress here again the important point we discussed at the end of the previous section. As for Eq. (230), also estimate (233) refers to the level of non-Gaussianity only during the inflationary epoch. In order to make a full and sensible comparison with the observations it is necessary to perform a fully second-order analysis of the evolution of the perturbations after inflation ends, through the radiation and matter dominated epochs. As we shall see, the non-linearities arising from the post-inflationary dynamics will anyway enhance the observable non-Gaussianity level in the CMB temperature anisotropy in such a way as to hide the initial imprint such as that in (233).

5.4. Reheating after inflation

The first step to follow the evolution of non-linearities on large scales after inflation is to analyze how the curvature perturbation ζ evolves on large scales during reheating.

In Ref. [31] it was shown that also at second order the curvature perturbation $\zeta^{(2)}$ remains constant during inflation, under the inflaton sudden-decay approximation. Under such an approximation the individual energy density perturbations (and hence the corresponding curvature perturbations) are separately conserved until the decay of the scalar field, which amounts to saying that in the equations for the curvature perturbations Eqs. (203) and (204) one can drop the energy transfer triggered by the decay rate $a\Gamma/\mathscr{H} \ll 1$. Going beyond the sudden-decay approximation, the first-order results $\zeta_\varphi^{(1)} = \zeta^{(1)}$ in Eq. (197) yield

$$\zeta^{(2)} = f\zeta_\varphi^{(2)} + (1-f)\zeta_\gamma^{(2)} \ . \tag{235}$$

The equation of motion for $\zeta^{(2)}$ on large scales is obtained by differentiating this expression and by using Eqs. (203) and (204), with $\delta^{(2)}\Gamma = 0$ and $\zeta_\varphi^{(1)} = \zeta^{(1)}$; it reads

$$\zeta^{(2)'} = -\mathscr{H} f(\zeta^{(2)} - \zeta_\varphi^{(2)}) \ . \tag{236}$$

In the same way as at first order from Eqs. (235) and (236) it follows that the second-order curvature perturbation $\zeta^{(2)}$ remains constant on large-scales during the reheating phase, being given at the end of inflation by the curvature perturbation in the inflaton field $\zeta_{\rm in}^{(2)} = \zeta_{\varphi,{\rm in}}^{(2)}$.

[21] The exact expression also contains a scale-dependent part of the same magnitude, whose precise form however differs from the one obtained in the standard case, see Ref. [72] for more details.

5.5. Post-inflationary evolution of the second-order curvature perturbation

The superhorizon-scale evolution of the primordial non-linearity generated during inflation during the radiation and matter dominated phases has been studied in Ref. [31]. Following their approach, we consider the energy–momentum tensor for a perfect fluid with constant but otherwise generic equation of state, as defined in Section 3.2.

We will explicitly show that during the radiation and matter dominated epochs the second-order curvature perturbation $\zeta^{(2)}$ is conserved on large scales. From now on we shall adopt the Poisson gauge. Our starting point is the energy continuity equation at second order

$$\begin{aligned}
&\delta^{(2)}\rho' + 3\mathscr{H}(1+w)\delta^{(2)}\rho - 3(1+w)\rho_0\psi^{(2)'} - 6(1+w)\psi^{(1)'}[\delta^{(1)}\rho + 2\rho_0\psi^{(1)}] \\
&\quad = -2(1+w)\rho_0(v_i^{(1)}v_{(1)}^i)' - 2(1+w)(1-3w)\mathscr{H}\rho_0 v_i^{(1)}v_{(1)}^i \\
&\qquad + 4(1+w)\rho_0\partial_i\psi^{(1)}v_{(1)}^i + 2\,\frac{\rho_0}{\mathscr{H}^2}(\psi^{(1)}\nabla^2\psi^{(1)\prime} - \psi^{(1)\prime}\nabla^2\psi^{(1)})\ ,
\end{aligned} \tag{237}$$

where we have also used the $(0-i)$-component of Einstein equation (see Appendix A). This equation can be rewritten in a more suitable form as

$$\begin{aligned}
&\left[\psi^{(2)} + \mathscr{H}\,\frac{\delta^{(2)}\rho}{\rho_0'} + (1+3w)\mathscr{H}^2\left(\frac{\delta\rho^{(1)}}{\rho_0'}\right)^2 - 4\mathscr{H}\left(\frac{\delta\rho^{(1)}}{\rho_0'}\right)\psi^{(1)}\right]' \\
&\quad = \frac{2}{3}\,(v_i^{(1)}v_{(1)}^i)' + \frac{2}{3}\,(1-3w)\mathscr{H}v_i^{(1)}v_{(1)}^i - \frac{4}{3}\,\partial_i\psi^{(1)}v_{(1)}^i + \frac{16}{27(1+w)^2\mathscr{H}}\,\psi^{(1)}\nabla^2\psi^{(1)} \\
&\qquad - \frac{2}{3(1+w)\mathscr{H}^2}\left[\left(1 - \frac{8}{9\,(1+w)}\right)\psi^{(1)}\nabla^2\psi^{(1)\prime} - \left(1 - \frac{4\,(1+3w)}{9(1+w)}\right)\psi^{(1)\prime}\nabla^2\psi^{(1)}\right] \\
&\qquad + \frac{8(1+3w)}{27(1+w)^2\mathscr{H}^3}\left[\frac{\left(\nabla^2\psi^{(1)}\right)^2}{3} - \psi^{(1)\prime}\nabla^2\psi^{(1)\prime} + \frac{\nabla^2\psi^{(1)\prime}\nabla^2\psi^{(1)}}{3\mathscr{H}}\right]\ ,
\end{aligned} \tag{238}$$

where the argument on the L.H.S. can be further simplified to

$$\psi^{(2)} + \mathscr{H}\,\frac{\delta^{(2)}\rho}{\rho_0'} - (5+3w)\mathscr{H}^2\left(\frac{\delta^{(1)}\rho}{\rho_0'}\right)^2 = \psi^{(2)} + \mathscr{H}\,\frac{\delta^{(2)}\rho}{\rho_0'} - \frac{4}{5+3w}\,(\zeta_I^{(1)})^2\ , \tag{239}$$

and the final form has been obtained employing Eqs. (151) and (187).

Notice that the quantity in Eq. (239) is in fact the curvature perturbation defined in Eq. (134) in the case of the generic fluid with constant equation of state that we are considering here. From Eqs. (238) and (239) we find

$$\psi^{(2)} + \mathscr{H}\,\frac{\delta^{(2)}\rho}{\rho_0'} - (5+3w)\mathscr{H}^2\left(\frac{\delta^{(1)}\rho}{\rho_0'}\right)^2 = \mathscr{C} + \frac{2}{3}\,(v_i^{(1)}v_{(1)}^i) + \int^\tau \mathrm{d}\tau'\mathscr{S}(\tau')\ , \tag{240}$$

where $\mathscr{C}$ is a constant in time, $\mathscr{C}' = 0$, on large scales and

$$\mathscr{S} = \frac{2}{3}(1-3w)\mathscr{H} v_i^{(1)} v_{(1)}^i - \frac{4}{3}\partial_i\psi^{(1)} v_{(1)}^i + \frac{16}{27(1+w)^2\mathscr{H}}\psi^{(1)}\nabla^2\psi^{(1)} - \frac{2}{3(1+w)\mathscr{H}^2}\left[\left(1-\frac{8}{9(1+w)}\right)\psi^{(1)}\nabla^2\psi^{(1)\prime} - \left(1-\frac{4(1+3w)}{9(1+w)}\right)\psi^{(1)\prime}\nabla^2\psi^{(1)}\right] + \frac{8(1+3w)}{27(1+w)^2\mathscr{H}^3}\left[\frac{\left(\nabla^2\psi^{(1)}\right)^2}{3} - \psi^{(1)\prime}\nabla^2\psi^{(1)\prime} + \frac{\nabla^2\psi^{(1)\prime}\nabla^2\psi^{(1)}}{3\mathscr{H}}\right] . \tag{241}$$

We are interested in the determination of the non-linearities after the inflationary stage. We have seen in the previous section that also during the reheating phase the curvature perturbation $\zeta^{(2)}$ is conserved. Therefore we are allowed to fix the constant $\mathscr{C}$ by matching the conserved quantity on large scales at the end of inflation ($\tau = \tau_I$)

$$\mathscr{C} = \psi_I^{(2)} + \mathscr{H}_I\frac{\delta^{(2)}\rho_I}{\rho_{0I}'} - 2(\zeta_I^{(1)})^2 , \tag{242}$$

where we have used the fact that during inflation $w_I \simeq -1$ and we have disregarded gradient terms which turn out to be negligible for the computation of the large-scale bispectrum.

In fact the inflationary quantity $(\psi_I^{(2)} + \mathscr{H}_I\delta^{(2)}\rho_I/\rho_{0I}')$ has been computed in Refs. [3,182]

$$\psi_I^{(2)} + \mathscr{H}_I\frac{\delta^{(2)}\rho_I}{\rho_{0I}'} \simeq (\eta - 3\epsilon)(\zeta_I^{(1)})^2 + \mathscr{O}(\epsilon,\eta)\ \text{(non-local terms)} , \tag{243}$$

in terms of the slow-roll parameters $\epsilon = 1 - \mathscr{H}_I'/\mathscr{H}_I^2$ and $\eta = 1 + \epsilon - \left(\varphi''/\mathscr{H}_I\varphi'\right)$ where $\mathscr{H}_I$ is the Hubble parameter during inflation and φ is the inflaton field [174]. Since during inflation the slow-roll parameters are tiny, we can safely disregard the intrinsically second-order terms originated from the inflationary epoch. Thus from Eqs. (240) and (242) we obtain a relation between the gravitational potential $\psi^{(2)}$ and the energy density perturbation $\delta^{(2)}\rho$ during the radiation/matter dominated epochs

$$\psi^{(2)} - \frac{1}{3(1+w)}\frac{\delta^{(2)}\rho}{\rho_0} = -\frac{2}{3}\frac{5+3w}{1+w}(\psi^{(1)})^2 + \frac{2}{3}(v_i^{(1)}v_{(1)}^i) + \int_{\tau_I}^{\tau}\mathscr{S}(\tau')\,\mathrm{d}\tau' , \tag{244}$$

where we have made use of Eqs. (151) and (187), with $\zeta^{(1)} = \zeta_I^{(1)}$ since the curvature perturbation is conserved on large-scales.

6. The curvaton scenario

Let us now consider the so-called curvaton mechanism [79,176,201] to generate an initially adiabatic perturbation deep in the radiation era, as an alternative to the standard inflationary picture. In fact in the curvaton scenario the cosmological perturbations are produced from fluctuations of a scalar field σ (different from the inflaton) during a period of inflation, in the case where the perturbations from the inflaton field are considered to be negligible. The scalar field is subdominant during inflation and thus its fluctuations are initially of isocurvature type. Therefore a curvature perturbation is sourced on large

scales according to Eqs. (144) and (147). The curvature perturbation will become relevant when the energy density of the curvaton field is a significant fraction of the total energy. This happens after the end of inflation when the curvaton field begins to oscillate around the minimum of its potential once its mass has dropped below the Hubble rate, behaving like non-relativistic matter. Finally, well before primordial nucleosynthesis, the curvaton field is supposed to completely decay into thermalised radiation thus generating a final adiabatic perturbation.[22] From this epoch onwards the "standard" radiation dominated phase takes place.

6.1. Generating the curvature perturbation at linear order

During inflation the curvaton field σ is supposed to give a negligible contribution to the energy density and to be an almost free scalar field, with a small effective mass $m_\sigma^2 = |\partial^2 V/\partial\sigma^2| \ll H_I^2$ [176,175], where $H_I = \dot{a}/a$ is the Hubble rate during inflation.

The unperturbed curvaton field satisfies the equation of motion

$$\sigma'' + 2\mathscr{H}\sigma' + a^2 \frac{\partial V}{\partial \sigma} = 0 \ . \tag{245}$$

It is also usually assumed that the curvaton field is very weakly coupled to the scalar fields driving inflation and that the curvature perturbation from the inflaton fluctuations is negligible [176,175]. Notice that these are just the conditions under which we worked in Section 2.4 when calculating the spectrum of perturbations generated by the quantum fluctuations of a generic light scalar field during inflation. Thus, if we expand the curvaton field up to first order in the perturbations around the homogeneous background as

$$\sigma(\tau, \mathbf{x}) = \sigma(\tau) + \delta^{(1)}\sigma \ , \tag{246}$$

the linear perturbations satisfy on large-scales the equation

$$\delta^{(1)}\sigma'' + 2\mathscr{H}\delta^{(1)}\sigma' + a^2 \frac{\partial^2 V}{\partial \sigma^2}\, \delta^{(1)}\sigma = 0 \ . \tag{247}$$

The fluctuations $\delta\sigma$ on superhorizon scales will be Gaussian distributed and, from the results of Section 2.4, they will have a nearly scale-invariant spectrum—see Eq. (77)—

$$\mathscr{P}_{\delta\sigma}(k) \approx \frac{H_*^2}{4\pi^2} \ , \tag{248}$$

where the subscript $*$ denotes the epoch of horizon exit $k = aH$. Once inflation is over the inflaton energy density will be converted into radiation (γ) and the curvaton field will remain approximately constant

[22] In the curvaton scenario it is indeed possible that some residual isocurvature perturbations survive after the curvaton decay. This could be the case for example if the curvaton field decays when subdominant into a component of cold dark matter (CDM) which does not thermalize with the existing radiation. This is due to the fact that an isocurvature perturbation is present initially, while in the standard scenario of inflation it is not possible since the perturbations initially are adiabatic. If this is the case, non-Gaussianity in the isocurvature perturbations are expected as well. We refer the reader to Refs. [176,175,100] for more details. Here we will just consider the simplest setting of the curvaton scenario where only adiabatic perturbations are left after the curvaton decay.

until $H^2 \sim m_\sigma^2$. At this epoch the curvaton field begins to oscillate around the minimum of its potential which can be safely approximated by the quadratic term $V \approx m_\sigma^2\sigma^2/2$. During this stage the energy density of the curvaton field just scales as non-relativistic matter $\rho_\sigma \propto a^{-3}$ [283]. The energy density in the oscillating field is

$$\rho_\sigma(\tau, \mathbf{x}) \approx m_\sigma^2\sigma^2(\tau, \mathbf{x}) \ , \tag{249}$$

and it can be expanded into a homogeneous background $\rho_\sigma(\tau)$ and a first-order perturbation $\delta^{(1)}\rho_\sigma$ as

$$\rho_\sigma(\tau, \mathbf{x}) = \rho_\sigma(\tau) + \delta^{(1)}\rho_\sigma(\tau, \mathbf{x}) = m_\sigma^2\sigma + 2m_\sigma^2\sigma\delta^{(1)}\sigma \ . \tag{250}$$

As it follows from Eqs. (245) and (247) for a quadratic potential the ratio $\delta^{(1)}\sigma/\sigma$ remains constant and the resulting relative energy density perturbation is

$$\frac{\delta^{(1)}\rho_\sigma}{\rho_\sigma} = 2\left(\frac{\delta^{(1)}\sigma}{\sigma}\right)_* \ , \tag{251}$$

where the $*$ stands for the value at horizon crossing.

Perturbations in the energy density of the curvaton field produce in fact a primordial density perturbation well after the end of inflation. The primordial adiabatic density perturbation is associated with a perturbation in the spatial curvature ψ and it is characterized in a gauge-invariant manner by the curvature perturbation ζ on hypersurfaces of uniform total density ρ, introduced in Section 3.5. At linear order ζ is defined by Eq. (131) and on large scales its evolution is sourced by the non-adiabatic pressure perturbation $\delta^{(1)}P_{\rm nad} = \delta^{(1)}P - c_s^2\delta^{(1)}\rho$, obeying the equation of motion (144). In the curvaton scenario the curvature perturbation is generated well after the end of inflation during the oscillations of the curvaton field because the pressure of the mixture of matter (curvaton) and radiation produced by the inflaton decay is not adiabatic. A convenient way to study this mechanism is to consider the curvature perturbations ζ_i associated with each individual energy density components defined in Eq. (132). In fact the weighted sum in Eq. (170) during the oscillations of the curvaton field can be written as [177,175]

$$\zeta^{(1)} = (1-f)\zeta_\gamma^{(1)} + f\zeta_\sigma^{(1)} \ , \tag{252}$$

with the quantity f defining the relative contribution of the curvaton field to the total curvature perturbation is now given by

$$f = \frac{3\rho_\sigma}{4\rho_\gamma + 3\rho_\sigma} \tag{253}$$

according to Eq. (171).

From now on we shall work under the approximation of sudden decay of the curvaton field. Under this approximation the curvaton and the radiation components ρ_σ and ρ_γ satisfy separately the energy conservation equations

$$\begin{aligned} \rho_\gamma' &= -4\mathscr{H}\rho_\gamma \ , \\ \rho_\sigma' &= -3\mathscr{H}\rho_\sigma \ , \end{aligned} \tag{254}$$

and the curvature perturbations $\zeta_i^{(1)}$ remains constant on superhorizon scales until the decay of the curvaton, as it follows from Eqs. (182)–(183) in the limit $a\Gamma/\mathscr{H} \ll 1$.

Therefore from Eq. (252) it follows that the first-oder curvature perturbation evolves on large-scales as

$$\zeta^{(1)'} = f'(\zeta_\sigma^{(1)} - \zeta_\gamma^{(1)}) = \mathscr{H} f(1-f)(\zeta_\sigma^{(1)} - \zeta_\gamma^{(1)}) \ , \tag{255}$$

where we have used the conservation of the curvature perturbations. By comparison with Eq. (144) one obtains the expression for the non-adiabatic pressure perturbation at first order [176,175]

$$\delta^{(1)} P_{\rm nad} = \rho_\sigma (1-f)(\zeta_\gamma^{(1)} - \zeta_\sigma^{(1)}) \ . \tag{256}$$

Since in the curvaton scenario it is supposed that the curvature perturbation in the radiation produced at the end of inflation is negligible

$$\zeta_\gamma^{(1)} = -\hat{\psi}^{(1)} + \frac{1}{4}\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma} = 0 \ . \tag{257}$$

Similarly the value of $\zeta_\sigma^{(1)}$ is fixed by the fluctuations of the curvaton during inflation

$$\zeta_\sigma^{(1)} = -\hat{\psi}^{(1)} + \frac{1}{3}\frac{\delta^{(1)}\rho_\sigma}{\rho_\sigma} = \zeta_{\sigma I}^{(1)} \ , \tag{258}$$

where I stands for the value of the fluctuations during inflation. From Eq. (252) the total curvature perturbation during the curvaton oscillations is given by

$$\zeta^{(1)} = f\zeta_\sigma^{(1)} \ . \tag{259}$$

As it is clear from Eq. (259) initially, when the curvaton energy density is subdominant, the density perturbation in the curvaton field $\zeta_\sigma^{(1)}$ gives a negligible contribution to the total curvature perturbation, thus corresponding to an isocurvature (or entropy) perturbation. On the other hand during the oscillations, $\rho_\sigma \propto a^{-3}$ increases with respect to the energy density of radiation $\rho_\gamma \propto a^{-4}$, and the perturbations in the curvaton field are then converted into the curvature perturbation. Well after the decay of the curvaton, during the conventional radiation and matter-dominated eras, the total curvature perturbation will remain constant on superhorizon scales at a value which, in the sudden-decay approximation, is fixed by Eq. (259) at the epoch of curvaton decay

$$\zeta^{(1)} = f_{\rm D}\,\zeta_\sigma^{(1)} \ , \tag{260}$$

where D stands for the epoch of the curvaton decay.

Going beyond the sudden-decay approximation it is possible to introduce a transfer parameter r defined as [175,186]

$$\zeta^{(1)} = r\zeta_\sigma^{(1)} \ , \tag{261}$$

where $\zeta^{(1)}$ is evaluated well after the epoch of the curvaton decay and $\zeta_\sigma^{(1)}$ is evaluated well before this epoch. The numerical study of the coupled perturbation equations has been performed in Ref. [186] showing that the sudden-decay approximation is exact when the curvaton dominates the energy density before it decays ($r = 1$), while in the opposite case

$$r \approx \left(\frac{\rho_\sigma}{\rho}\right)_{\rm D} . \tag{262}$$

6.2. Second-order curvature perturbation from the curvaton fluctuations

As we have shown in Section 5 in the standard scenario where the generation of cosmological perturbations is induced by fluctuations of a single inflaton field (and there is no curvaton) the evolution of the perturbations is purely adiabatic, and the total curvature perturbation $\zeta^{(2)}$ is indeed conserved. Thus, following Ref. [31], we have used the conserved quantity $\zeta^{(2)}$ to follow the evolution on large scales of the primordial non-linearity in the cosmological perturbations from a period of inflation to the matter dominated era. On the contrary in the curvaton and inhomogeneous reheating scenarios the total curvature perturbation $\zeta^{(2)}$ evolves on large scales due to a non-adiabatic pressure. In the present scenario the conversion of the curvaton isocurvature perturbations into a final curvature perturbation at the epoch of the curvaton decay can be followed through the sum (252) of the individual curvature perturbations weighted by the ratio f of Eq. (253).

Let us now extend such a result at second order in the perturbations. Since the quantities $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$ are gauge-invariant, we choose to work in the spatially flat gauge $\psi^{(r)} = \chi^{(r)} = 0$ if not otherwise specified. Note that from Eqs. (251) and (258) the value of $\zeta_\sigma^{(1)}$ is thus given by

$$\zeta_\sigma^{(1)} = \frac{1}{3}\frac{\delta^{(1)}\rho_\sigma}{\rho_\sigma} = \frac{2}{3}\frac{\delta^{(1)}\sigma}{\sigma} = \frac{2}{3}\left(\frac{\delta^{(1)}\sigma}{\sigma}\right)_* , \tag{263}$$

where we have used the fact that $\zeta_\sigma^{(1)}$ (or equivalently $\delta^{(1)}\sigma/\sigma$) remains constant, while from Eq. (257) in the spatially flat gauge

$$\zeta_\gamma^{(1)} = \frac{1}{4}\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma} . \tag{264}$$

During the oscillations of the scalar field Eq. (197) with $\delta^{(1)}\Gamma = 0$ reduces to

$$\zeta^{(2)} = f\zeta_\sigma^{(2)} + (1-f)\zeta_\gamma^{(2)} + f(1-f)(1+f)(\zeta_\sigma^{(1)} - \zeta_\gamma^{(1)})^2 , \tag{265}$$

where we have used the sudden-decay limit $a\Gamma/\mathscr{H} \ll 1$ and within such an approximation f is given by Eq. (253). Similarly from Eqs. (193)–(194) the expression of the individual curvature perturbations in the spatially flat-gauge now read

$$\zeta_\sigma^{(2)} = \frac{1}{3}\frac{\delta^{(2)}\rho_\sigma}{\rho_\sigma} - (\zeta_\sigma^{(1)})^2 , \tag{266}$$

$$\zeta_\gamma^{(2)} = \frac{1}{4}\frac{\delta^{(2)}\rho_\gamma}{\rho_\gamma} - 2(\zeta_\gamma^{(1)})^2 . \tag{267}$$

Such quantities are gauge-invariant and, in the sudden-decay approximation they are separately conserved until the curvaton decay.

Therefore from Eq. (265) it follows that $\zeta^{(2)}$ evolves according to Eq. [32]

$$\zeta^{(2)'} = f'(\zeta_\sigma^{(2)} - \zeta_\gamma^{(2)}) + f'(1-3f^2)(\zeta_\sigma^{(1)} - \zeta_\gamma^{(1)})^2 . \tag{268}$$

Note that Eq. (268) can be rewritten as Eq. (147) derived in Ref. [185] with $\delta^{(1)}P_{\rm nad}$ given by Eq. (256) and

$$\delta^{(2)}P_{\rm nad} = \rho_\sigma(1-f)[(\zeta_\gamma^{(2)}-\zeta_\sigma^{(2)}) + (f^2+6f-1) \times (\zeta_\sigma^{(1)}-\zeta_\gamma^{(1)})^2 + 4\zeta_\gamma^{(1)}(\zeta_\sigma^{(1)}-\zeta_\gamma^{(1)})] \,, \tag{269}$$

is the gauge-invariant non-adiabatic pressure perturbation on uniform density hypersurfaces on large scales which, as one can easily check, coincides with the generic expression in Eq. (145) which has been provided in Ref. [185].

The second-order curvature perturbation in the standard radiation or matter eras remain constant on superhorizon scales and, in the sudden-decay approximation, it is thus given by the quantity in Eq. (265) evaluated at the epoch of the curvaton decay

$$\zeta^{(2)} = f_{\rm D}\zeta_\sigma^{(2)} + f_{\rm D}(1-f_{\rm D}^2)(\zeta_\sigma^{(1)})^2 \,, \tag{270}$$

where we have used the curvaton hypothesis that the curvature perturbation in the radiation produced at the end of inflation is negligible so that $\zeta_\gamma^{(1)} \approx 0$ and $\zeta_\gamma^{(2)} \approx 0$. The curvature perturbation $\zeta_\sigma^{(1)}$ is given by Eq. (263), while $\zeta_\sigma^{(2)}$ in Eq. (266) is obtained by expanding the energy density of the curvaton field, Eq. (249), up to second order in the curvaton fluctuations

$$\begin{aligned}\rho_\sigma(\mathbf{x},t) &= \rho_\sigma(\tau) + \delta^{(1)}\rho_\sigma(\tau,x^i) + \tfrac{1}{2}\,\delta^{(2)}\rho_\sigma(\tau,x^i) \\ &= m_\sigma^2\sigma + 2m_\sigma^2\sigma\delta^{(1)}\sigma + m_\sigma^2(\delta^{(1)}\sigma)^2 \,.\end{aligned} \tag{271}$$

It follows that

$$\frac{\delta^{(2)}\rho_\sigma}{\rho_\sigma} = \frac{1}{2}\left(\frac{\delta^{(1)}\rho_\sigma}{\rho_\sigma}\right)^2 = \frac{9}{2}\,(\zeta_\sigma^{(1)})^2 \,, \tag{272}$$

where we have used Eq. (263), and hence from Eq. (266) we obtain

$$\zeta_\sigma^{(2)} = \frac{1}{2}\,(\zeta_\sigma^{(1)})^2 = \frac{1}{2}\,(\zeta_\sigma^{(1)})_I^2 \,, \tag{273}$$

where we have emphasized that also $\zeta_\sigma^{(2)}$ is a conserved quantity whose value is determined by the curvaton fluctuations during inflation. Plugging Eq. (273) into Eq. (270) the curvature perturbation during the standard radiation or matter-dominated eras turns out to be [32]

$$\zeta^{(2)} = f_{\rm D}\left(\frac{3}{2} - f_{\rm D}^2\right)(\zeta_\sigma^{(1)})^2 \,. \tag{274}$$

From now on we switch from the spatially flat gauge $\psi = \chi = 0$ to the Poisson gauge defined in Section 3.1. Such a procedure is possible since the curvature perturbations $\zeta_i^{(2)}$ are gauge-invariant quantities. In particular this is evident from the expression found in Eq. (274). In fact we are interested in the non-linearities produced in the gravitational potentials in the Poisson gauge. By doing so we are in the position to obtain a relation between the gravitational potential $\psi^{(2)}$ and the energy density $\delta^{(2)}\rho$ in the radiation/matter dominated epochs.

From Eq. (134) we find that during the matter-dominated era

$$\zeta^{(2)} = -\psi^{(2)} + \frac{1}{3}\frac{\delta^{(2)}\rho}{\rho} + \frac{5}{9}\left(\frac{\delta^{(1)}\rho}{\rho}\right)^2 = -\psi^{(2)} + \frac{1}{3}\frac{\delta^{(2)}\rho}{\rho} + \frac{20}{9}(\psi^{(1)})^2 \,, \tag{275}$$

where in the last step we have used the first-order solution (151) on large scales in the Poisson gauge. On the other hand, the curvature perturbation in the radiation/matter-dominated eras remains constant at a value which is fixed by Eq. (274). Thus Eq. (275) combined with Eq. (274), yields [32]

$$\psi^{(2)} - \frac{1}{3}\frac{\delta^{(2)}\rho}{\rho} = \frac{1}{9}\left[20 - \frac{75}{2f_{\rm D}} + 25 f_{\rm D}\right](\psi^{(1)})^2 \,, \tag{276}$$

where we have used

$$f_{\rm D}\zeta^{(1)}_{\sigma} = -\tfrac{5}{3}\psi^{(1)} \tag{277}$$

from Eq. (260) and the usual linear relation between the curvature perturbation and the gravitational potential $\zeta^{(1)} = -\frac{5}{3}\psi^{(1)}$ during the matter-dominated era, see Eq. (187).

7. The inhomogeneous reheating scenario: $\delta\Gamma \neq 0$

Recently, another mechanism for the generation of cosmological perturbations has been proposed [76,130,77,192,10]. It acts during the reheating stage after inflation and it was dubbed the "inhomogeneous reheating" mechanism in Ref. [77] and "modulated reheating" in Ref. [130]. This mechanism works as follows. As in the curvaton scenario it is supposed that the perturbations coming from the inflaton fluctuations are negligible. To reheat the Universe the inflaton has to couple to ordinary particles and has to decay into radiation with a decay rate Γ which depends on the couplings of the inflaton field. In the standard scenario of inflation such a coupling is constant. In fact it may be determined by the vacuum expectation value of fields χ's of the underlying theory. It could be the case of supersymmetric theories or theories inspired by superstrings, as discussed in some details in Refs. [76,130], respectively, with the scalar fields $\chi's$ being some scalar super-partner or the so-called moduli fields. If those fields are light during inflation fluctuations $\delta\chi \sim H/2\pi$, where H is the Hubble rate during inflation, are left imprinted on superhorizon scales, as we have recalled in Section 2.4. These perturbations lead to spatial fluctuations in the decay rate Γ of the inflaton field to ordinary matter

$$\frac{\delta\Gamma}{\Gamma} \sim \frac{\delta\chi}{\chi} \,, \tag{278}$$

thus producing fluctuations in the radiation and in the reheating temperature in different regions of the Universe. These fluctuations are of isocurvature type and will be converted into curvature fluctuations after reheating, once the thermalized radiation starts do dominate the energy density.[23]

[23] Indeed the idea that the total curvature perturbation may be affected on large scales by entropy perturbations when there exists a scalar field affecting particle masses or couplings constants controlling the reheating process has first been suggested in Ref. [103].

7.1. Generating the curvature perturbation at linear order from decay rate fluctuations

Using the cosmic time as time variable, the first order Eq. (182) for $\zeta_\varphi^{(1)}$ on large scales reads

$$\dot{\zeta}_\varphi^{(1)} = \frac{\Gamma}{2}\frac{\rho_\varphi}{\dot{\rho}_\varphi}\frac{\dot{\rho}}{\rho}(\zeta^{(1)} - \zeta_\varphi^{(1)}) + H\frac{\rho_\varphi}{\dot{\rho}_\varphi}\delta^{(1)}\Gamma \ . \tag{279}$$

We shall now adopt a "mixed sudden-decay approximation". We shall treat the pressureless scalar field and radiation fluids as if they were not interacting until the decay of the inflaton, when $\Gamma \approx H$. Since at the beginning of the reheating phase the energy density in radiation is negligible this means that $f = \dot{\rho_\varphi}/\dot{\rho} \approx 1$ and there is indeed only a single fluid with, from Eq. (170), $\zeta^{(1)} \approx \zeta_\varphi^{(1)}$ and $\zeta_\gamma^{(1)} \approx 0$. In fact under such an approximation we can neglect all the terms proportional to the decay rate Γ, *but* we allow for the spatial fluctuations of the decay rate. Thus the first-order Eq. (279) reads

$$\dot{\zeta}_\varphi^{(1)} \simeq -\tfrac{1}{3}\,\delta^{(1)}\Gamma \ , \tag{280}$$

where we have used $\dot{\rho_\varphi} = -3H\rho_\varphi$ in the sudden-decay approximation. Integration over time yields

$$\zeta_\varphi^{(1)} = -\frac{t}{3}\,\delta^{(1)}\Gamma = -\frac{2}{9}\frac{\delta^{(1)}\Gamma}{H} \simeq \zeta^{(1)} \ , \tag{281}$$

where we have used the fact that during the oscillations of the scalar field which dominates the energy density $H = \frac{2}{3}\,t$. The inhomogeneous reheating mechanism produces at linear level a gravitational potential which after the reheating phase, in the radiation-dominated epoch, is given by (in the longitudinal gauge) [76]

$$\psi^{(1)} = \frac{1}{9}\frac{\delta^{(1)}\Gamma}{\Gamma_*} \ , \tag{282}$$

where Γ_* stands for the value of the background decay rate, which in this scenario is approximately constant, being determined by the very light scalar field(s) χ. During the radiation-dominated epoch the usual relation between the gravitational potential and the curvature perturbation in Eq. (187) yields

$$\psi^{(1)} = -\tfrac{2}{3}\,\zeta^{(1)} \ , \tag{283}$$

and thus from Eq. (281) we can set the ratio $\Gamma_*/H_{\rm D} = \frac{3}{4}$ at the time of the inflaton decay in order to reproduce the result (282) of Ref. [76]. Therefore from Eq. (281) it follows that the value of $\zeta^{(1)}$ is [33]

$$\zeta^{(1)} \simeq -\frac{1}{6}\frac{\delta^{(1)}\Gamma}{\Gamma_*} \ . \tag{284}$$

7.2. Second-order curvature perturbation from inhomogeneous reheating

We now expand the decay rate as

$$\Gamma = \Gamma_* + \delta\Gamma = \Gamma_* + \delta^{(1)}\Gamma + \tfrac{1}{2}\,\delta^{(2)}\Gamma \ , \tag{285}$$

and perturbing the energy transfer coefficient $\hat{Q}_\varphi = -\Gamma\rho_\varphi$ up to second order it follows from Eqs. (198) and (285)

$$\delta^{(2)}Q_\varphi = -\rho_\varphi\delta^{(2)}\Gamma - \Gamma_*\delta^{(2)}\rho_\varphi - 2\delta^{(1)}\Gamma\delta^{(1)}\rho_\varphi \ . \tag{286}$$

Plugging Eq. (286) into Eq. (203), the equation of motion on large scales for the curvature perturbation $\zeta_\varphi^{(2)}$ allowing for possible fluctuations of the decay rate $\delta^{(1)}\Gamma$ and $\delta^{(2)}\Gamma$ turns out to be [33]

$$\begin{aligned}\dot{\zeta}_\varphi^{(2)} = {} & \frac{H}{\dot{\rho}_\varphi}\left(\delta^{(2)}\Gamma\rho_\varphi + 2\delta^{(1)}\Gamma\delta^{(1)}\rho_\varphi\right) - \frac{\Gamma_*\rho_\varphi}{2}\frac{\dot{\rho}}{\rho}\left(\frac{\delta^{(2)}\rho_\varphi}{\dot{\rho}_\varphi} - \frac{\delta^{(2)}\rho}{\dot{\rho}}\right) \\ & + 3\Gamma_*\rho_\varphi\frac{H}{\dot{\rho}_\varphi}\phi^{(1)2} + \frac{2H}{\dot{\rho}_\varphi}\left(\delta^{(1)}\Gamma\rho_\varphi + \delta^{(1)}\rho_\varphi\Gamma_*\right)\phi^{(1)} - 2\zeta_\varphi^{(1)}\dot{\zeta}_\varphi^{(1)} \\ & + 2\left[\zeta_\varphi^{(1)}\left(\Gamma_*\frac{\rho_\varphi}{\dot{\rho}_\varphi}\phi^{(1)} + \delta^{(1)}\Gamma\frac{\rho_\varphi}{\dot{\rho}_\varphi} + \Gamma_*\frac{\delta^{(1)}\rho_\varphi}{\dot{\rho}_\varphi}\right)\right]^{\cdot} + \left[\zeta_\varphi^{(1)2}\frac{\Gamma}{H}\left(1 - \frac{\rho_\varphi}{\dot{\rho}_\varphi}\frac{\dot{\rho}}{\rho}\right)\right]^{\cdot} ,\end{aligned} \tag{287}$$

where we have used the fact that the decay rate Γ in the scenario under consideration remains constant. We shall use the result previously found in Eq. (284) to solve this equation. In fact under the sudden-decay approximation and using Eq. (196) the second-order Eq. (287) simplifies to

$$\dot{\zeta}_\varphi^{(2)} \simeq -\frac{1}{3}\delta^{(2)}\Gamma - \zeta_\varphi^{(1)}\delta^{(1)}\Gamma - 2(\zeta_\varphi^{(1)}\dot{\zeta}_\varphi^{(1)}) - \frac{2}{3}\left(\frac{\delta^{(1)}\Gamma}{H}\zeta_\varphi^{(1)}\right)^{\cdot} \ . \tag{288}$$

Notice that the fluctuations $\delta\Gamma = \delta^{(1)}\Gamma + \frac{1}{2}\delta^{(2)}\Gamma$ indeed depend on the underlying physics for the coupling of the inflaton field to the other scalar field(s) χ. Let us take for example $\Gamma(t,\mathbf{x}) \propto \chi^2(t,\mathbf{x})$. If the scalar field χ is very light, its homogeneous value can be treated as constant $\chi(t) \approx \chi_*$ and during inflation quantum fluctuations $\delta^{(1)}\chi$ around its homogeneous value χ_* are left imprinted on superhorizon scales. Therefore non-linear fluctuations $(\delta^{(1)}\chi)^2$ of the decay rate Γ are produced as well

$$\Gamma(t,\mathbf{x}) \propto \chi^2(t,\mathbf{x}) = \chi_*^2 + 2\chi_*\delta^{(1)}\chi + (\delta^{(1)}\chi)^2 \ . \tag{289}$$

From Eqs. (285) and (289) it follows

$$\begin{aligned}\frac{\delta^{(1)}\Gamma}{\Gamma_*} &= 2\frac{\delta^{(1)}\chi}{\chi_*} \ , \\ \frac{\delta^{(2)}\Gamma}{\Gamma_*} &= 2\left(\frac{\delta^{(1)}\chi}{\chi_*}\right)^2 = \frac{1}{2}\left(\frac{\delta^{(1)}\Gamma}{\Gamma_*}\right)^2 \ .\end{aligned} \tag{290}$$

Using the first-order solution $\zeta_\varphi^{(1)} = -(t/3)\delta^{(1)}\Gamma$ and Eq. (290) in Eq. (288), the evolution of $\zeta_\varphi^{(2)}$ on large scales is

$$\dot{\zeta}_\varphi^{(2)} \simeq -\frac{1}{6\Gamma_*}(\delta^{(1)}\Gamma)^2 + \frac{1}{3}(\delta^{(1)}\Gamma)^2 t - 2(\zeta_\varphi^{(1)}\dot{\zeta}_\varphi^{(1)}) - \frac{2}{3}\left(\frac{\delta^{(1)}\Gamma}{H}\zeta_\varphi^{(1)}\right)^{\cdot} \ . \tag{291}$$

Integration over time is straightforward and yields

$$\zeta_\varphi^{(2)} = -\frac{t}{6}\frac{(\delta^{(1)}\Gamma)^2}{\Gamma_*} + \frac{1}{6}(\delta^{(1)}\Gamma)^2 t^2 - (\zeta_\varphi^{(1)})^2 - \frac{2}{3}\zeta_\varphi^{(1)}\frac{\delta^{(1)}\Gamma}{H} \ . \tag{292}$$

Now recall that at the time of inflaton decay $\Gamma_*/H_{\rm D} = 3/4$, and since $H = 2/3\,t$, it follows $t_{\rm D} = 1/2\Gamma_*$. Thus $\zeta_\varphi^{(2)}$ in Eq. (292) evaluated at the time $t_{\rm D}$ of inflaton decay is

$$\zeta_\varphi^{(2)} \simeq -\frac{1}{24}\left(\frac{\delta^{(1)}\Gamma}{\Gamma_*}\right)^2 - \zeta_\varphi^{(1)2} - \frac{1}{2}\zeta_\varphi^{(1)}\frac{\delta^{(1)}\Gamma}{\Gamma_*} \ . \tag{293}$$

Finally, using Eq. (284), we find that the total curvature perturbation $\zeta^{(2)}$ in the sudden-decay approximation is given by [33]

$$\zeta^{(2)} \simeq \zeta_\varphi^{(2)} \simeq \tfrac{1}{2}(\zeta^{(1)})^2 \ . \tag{294}$$

Eq. (294) gives the value at which the curvature perturbation remains constant during the radiation and dominated phases. Notice that our results are gauge-invariant, involving the curvature perturbations, as it is clear for example from Eq. (294). Thus, as in the previous section, we can now switch from the spatially flat gauge to the Poisson gauge, to obtain a relation analogous to the ones in Eqs. (244) and (276) between the energy density perturbation $\delta^{(2)}\rho$ and the gravitational potential $\psi^{(2)}$ during the matter-dominated epoch. By combining Eq. (294) with expression (275) for the curvature perturbation $\zeta^{(2)}$ in the Poisson gauge during the matter-dominated phase we find

$$\psi^{(2)} - \frac{1}{3}\frac{\delta^{(2)}\rho}{\rho} = \frac{5}{6}(\psi^{(1)})^2 \ . \tag{295}$$

8. Non-linearities in the gravitational potential

Let us now focus on the calculation of the non-linearity in the gravitational potential $\phi = \phi^{(1)} + \frac{1}{2}\phi^{(2)}$ (or ψ) in the Poisson gauge. In fact with our results we can express the gravitational potential ϕ in momentum space as

$$\phi(\mathbf{k}) = \phi^{(1)}(\mathbf{k}) + \frac{1}{(2\pi)^3}\int {\rm d}^3k_1\,{\rm d}^3k_2\,\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}) \times f_{\rm NL}^{\phi}(\mathbf{k}_1,\mathbf{k}_2)\phi^{(1)}(\mathbf{k}_1)\phi^{(1)}(\mathbf{k}_2) \ , \tag{296}$$

where we have defined an effective "momentum-dependent" non-linearity parameter $f_{\rm NL}^{\phi}$. Here the linear lapse function $\phi^{(1)} = \psi^{(1)}$ is a Gaussian random field. Notice that indeed a momentum-dependent function must be added to the R.H.S. of Eq. (296) in order to satisfy the requirement that $\langle\phi\rangle = 0$. From Eq. (296) it follows that the gravitational potential bispectrum reads

$$\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\rangle = (2\pi)^3\,\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3) \times [2f_{\rm NL}^{\phi}(\mathbf{k}_1,\mathbf{k}_2)P_\phi(k_1)P_\phi(k_2) + \text{cyclic}] \ , \tag{297}$$

where $P_\phi(k)$ is related to the dimensionless power-spectrum of the gravitational potential as defined in Eq. (59) by $P_\phi(k) = \mathscr{P}_\phi(k)\,2\pi^2/k^3$.

We want to make here an important remark. The non-linearity parameter $f_{\rm NL}^{\phi}$ defines the non-Gaussianity in the gravitational potential, but it does not define the non-Gaussianity level of the CMB temperature fluctuations. In order to predict such an observable it is necessary to make a further step, and determine how the perturbations in the gravitational potentials translate into second-order fluctuations of the CMB temperature. We will carry out this calculation in the next section. We now give the expression for the non-linearities in the gravitational potential ϕ. In each of the scenarios considered, solving the evolution for the curvature perturbation $\zeta^{(2)}$ for each of the scenarios considered, we have obtained the relations (244), (276), and (295) between the gravitational potential $\psi^{(2)}$ and the energy density $\delta^{(2)}\rho$ in terms of the linear gravitational potential squared $(\phi^{(1)})^2$. We can now close our system and fully determine the variables $\psi^{(2)}$, $\phi^{(2)}$ and $\delta^{(2)}\rho$ by using the (0-0)-component of Einstein equation (153) and the constraint (154) relating the gravitational potentials $\phi^{(2)}$ and $\psi^{(2)}$.

8.1. The standard scenario

Combining Eq. (244) obtained from the conservation of $\zeta^{(2)}$ with Eqs. (153) and (154) we single out an equation for the gravitational potential $\phi^{(2)}$ on large-scales

$$\begin{aligned}\phi^{(2)\prime}+\frac{5+3w}{2}\,\mathscr{H}\,\phi^{(2)} &= (5+3w)\mathscr{H}(\psi^{(1)})^2+\frac{3}{2}\,\mathscr{H}(1+w)[\nabla^{-2}(2\partial^i\psi^{(1)}\partial_i\psi^{(1)}\\ &\quad+3(1+w)\mathscr{H}^2v^i_{(1)}v_{(1)i})-3\nabla^{-4}\partial_i\partial^j(2\partial^i\psi^{(1)}\partial_j\psi^{(1)}\\ &\quad+3(1+w)\mathscr{H}^2v^i_{(1)}v_{(1)j})]+\frac{3}{2}\,\mathscr{H}(1+w)\int_{\tau_I}^{\tau}\mathscr{S}(\tau')\,\mathrm{d}\tau'-\mathscr{S}_1'\\ &\quad+\frac{1}{\mathscr{H}}\,(\nabla\psi^{(1)})^2+\frac{8}{3\mathscr{H}}\,\psi^{(1)}(\nabla^2\psi^{(1)})+\frac{\nabla^2\mathscr{S}_1}{3\mathscr{H}}+\frac{1}{\mathscr{H}}\,(\psi^{(1)\prime})^2\ , \end{aligned}\tag{298}$$

where $\mathscr{S}_1$ denotes the R.H.S. of Eq (154).

We want to integrate this equation from τ_I to a time τ in the matter-dominated epoch. The general solution is given by the solution of the homogeneous equation plus a particular solution

$$\begin{aligned}\phi^{(2)} &= \phi^{(2)}(\tau_I)\,\exp\left[-\int_{\tau_I}^{\tau}\frac{5+3w}{2}\,\mathscr{H}\,\mathrm{d}\tau'\right]\\ &\quad+\exp\left[-\int_{\tau_I}^{\tau}\frac{5+3w}{2}\,\mathscr{H}\,\mathrm{d}\tau'\right]\times\int_{\tau_I}^{\tau}\exp\left[\int_{\tau_I}^{\tau'}\frac{5+3w}{2}\,\mathscr{H}\,\mathrm{d}s\right]b(\tau')\,\mathrm{d}\tau'\ , \end{aligned}\tag{299}$$

where $b(\tau)$ stands for the source term in the R.H.S of Eq. (298).

Notice that the homogeneous solution during both the radiation and the matter-dominated epoch decreases in time. Therefore we can neglect the homogeneous solution and focus on the contributions from the source term $b(\tau)$. At a time τ in the matter-dominated epoch $\exp[-\int_{\tau_I}^{\tau}\mathrm{d}\tau'\mathscr{H}(5+3w)/2]\propto\tau^{-5}$. Thus if we are interested in the gravitational potential $\phi^{(2)}$ during the matter-dominated epoch the contributions in the particular solution coming from the radiation-dominated epoch can be considered negligible. Recalling that during the matter-dominated epoch the linear gravitational potential $\psi^{(1)}$ is constant in time,

it turns out that

$$\phi^{(2)} \simeq 2(\psi^{(1)})^2 + \frac{3}{5}\left[\nabla^{-2}\left(\frac{10}{3}\,\partial^i\psi^{(1)}\partial_i\psi^{(1)}\right) - 3\nabla^{-4}\partial_i\partial^j\left(\frac{10}{3}\,\partial^i\psi^{(1)}\partial_j\psi^{(1)}\right)\right] + \exp\left[-\int_{\tau_I}^{\tau}\frac{5+3w}{2}\,\mathscr{H}\,\mathrm{d}\tau'\right]\times\int_{\tau_I}^{\tau}\exp\left[\int_{\tau_I}^{\tau'}\frac{5+3w}{2}\,\mathscr{H}\,\mathrm{d}s\right]\left\{\frac{3}{2}\,\mathscr{H}(1+w) \times\int_{\tau_I}^{\tau'}\mathscr{S}(s)\,\mathrm{d}s + \frac{1}{\mathscr{H}}\,(\nabla\psi^{(1)})^2 + \frac{8}{3\mathscr{H}}\,\psi^{(1)}(\nabla^2\psi^{(1)}) + \frac{\nabla^2\mathscr{S}_1}{3\mathscr{H}}\right\}\mathrm{d}\tau'\ , \tag{300}$$

where we have used Eq. (149) to express the first-order velocities in terms of the gravitational potential, and we have taken into account that during the matter-dominated epoch $\mathscr{S}_1' = 0$.

As the gravitational potential $\psi^{(1)}$ on superhorizon scales is generated during inflation, it is clear that the origin of the non-linearity traces back to the inflationary quantum fluctuations.

The gravitational potential will then have a non-Gaussian (χ^2)-component. Going to momentum space, from Eq. (300) we directly read the non-linearity parameter of the gravitational potential $\phi = \phi^{(1)} + \frac{1}{2}\,\phi^{(2)}$ for scales entering the horizon during the matter-dominated stage [31,33]

$$f_{\mathrm{NL}}^{\phi}(\mathbf{k}_1,\mathbf{k}_2) \simeq -\tfrac{1}{2} + g(\mathbf{k}_1,\mathbf{k}_2)\ , \tag{301}$$

where

$$g(\mathbf{k}_1,\mathbf{k}_2) = 4\,\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k^2} - 3\,\frac{(\mathbf{k}_1\cdot\mathbf{k}_2)^2}{k^4} + \frac{3}{2}\,\frac{k_1^4+k_2^4}{k^4}\ , \tag{302}$$

with $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$. Notice that in deriving Eq. (301) we have neglected the contribution from the last term in Eq. (300), since as we explain in Section 8.3.1 this term is fully negligible when evaluating the bispectrum of the gravitational potential on large scales. Moreover in the final bispectrum expression, the diverging terms arising from the infrared behavior of $f_{\mathrm{NL}}^{\phi}(\mathbf{k}_1,\mathbf{k}_2)$ are automatically regularized once the monopole term is subtracted from the definition of ϕ, by requiring that $\langle\phi\rangle = 0$. The non-Gaussianity provided by expression (305) will add to the known Newtonian and relativistic second-order contributions which are relevant on subhorizon scales (a simple example being the Rees–Sciama effect [236]), whose complete and detailed analysis has been given in Refs. [233,234,199].

From Eq. (301) we conclude that the tiny non-Gaussianity generated during the inflationary epoch driven by a single scalar field, discussed in Section 5.3, gets enhanced in the post-inflationary evolution giving rise to a non-negligible signature of large-scale non-linearity in the gravitational potentials. Once again, inflation provides the key generating mechanism to produce superhorizon seeds, which are later amplified by gravity.

Finally it is interesting to note that as long as we are interested in the gravitational potential bispectrum on large-scales, it is possible to obtain the same result as in Eq. (301) using the following approximate solution to the (0-0)-component of Einstein equations (153) on large-scales:

$$\phi^{(2)} = -\frac{1}{2}\,\frac{\delta^{(2)}\rho}{\rho_0} + 4(\psi^{(1)})^2\ , \tag{303}$$

combining it with the equation obtained from the conservation of $\zeta^{(2)}$ Eq. (244) and with the constraint Eq. (155), both evaluated at the matter-dominated phase, in order to close the system of equations for

the variables $\psi^{(2)}$, $\phi^{(2)}$ and $\delta^{(2)}\rho$. The same holds true also for the other scenarios. Therefore, in the following we shall use relations (276) and (295), between the gravitational potential $\psi^{(2)}$ and the energy density $\delta^{(2)}\rho$ in the matter-dominated epoch, together with Eq. (303) and the constraint Eq. (155), which relates the gravitational potentials $\phi^{(2)}$ and $\psi^{(2)}$ in the matter-dominated epoch. As shown for the standard scenario, the terms which are neglected with such an approximation give a negligible contribution to the large-scale bispectrum of the gravitational potential.

8.2. The curvaton scenario

We use Eq. (276) obtained from the evolution of $\zeta^{(2)}$ in the sudden-decay approximation with Eqs. (303) and (155) and we conclude that in the curvaton scenario during the matter-dominated epoch [32]

$$\phi^{(2)} = \left[\frac{10}{3} + \frac{5}{3}\,f_{\rm D} - \frac{5}{2\,f_{\rm D}}\right](\psi^{(1)})^2 - 2\nabla^{-2}(\psi^{(1)}\nabla^2\psi^{(1)}) + 6\nabla^{-2}(\partial^i\partial_j(\psi^{(1)}\partial_i\partial^j\psi^{(1)}))\ , \tag{304}$$

where $f_{\rm D}$ is given by Eq. (253) at the time of the curvaton decay and defines the fractional energy density of the curvaton field. From Eq. (304) we obtain the non-linearity parameter for the gravitational potential $\phi = \phi^{(1)} + \frac{1}{2}\,\phi^{(2)}$ [32]

$$f_{\rm NL}^{\phi} = \left[\frac{7}{6} + \frac{5}{6}\,r - \frac{5}{4r}\right] + g(\mathbf{k}_1, \mathbf{k}_2)\ , \tag{305}$$

where we have replaced $f_{\rm D}$ with $r \approx (\rho_\sigma/\rho)_{\rm D}$ to go beyond the sudden-decay approximation, and the function $g(\mathbf{k}_1, \mathbf{k}_2)$ is the same as in Eq. (302). As far as the momentum-independent part of the non-linearity parameter $f_{\rm NL}^{\phi}$ is concerned, we note that in the limit $r \ll 1$ we obtain $f_{\rm NL}^{\phi} = -5/4r$ which reproduces the estimate provided in [176,175], while, in the limit $r \simeq 1$, we obtain $f_{\rm NL}^{\phi} = \frac{3}{4}$ for $r \simeq 1$.[24] Such a difference is due to the fact that we have taken into account all the relevant second-order gravitational effects.

An important comment is in order here. As it is evident from Eq. (305) the level of non-Gaussianity increases for decreasing values of the parameter r, that is to say with a lower efficiency for generating the density perturbations. This relation between the inefficiency and the non-Gaussianity is in fact quite a general feature, that has been pointed out in Refs. [139,300]. It is due to the fact that, in order to keep the density fluctuations at the observed level, as we decrease the efficiency to generate perturbations the second-order terms become more and more relevant in comparison with the linear contributions, thus increasing the level of non-Gaussianity.

8.3. The inhomogeneous reheating scenario

In the inhomogeneous reheating scenario where $\Gamma \propto \chi^2$ by combining Eq. (295) with Eqs. (303) and (155) we find during the matter dominated epoch

$$\phi^{(2)} = \tfrac{5}{2}\,(\psi^{(1)})^2 - 2\nabla^{-2}(\psi^{(1)}\nabla^2\psi^{(1)}) + 6\nabla^{-2}(\partial^i\partial_j(\psi^{(1)}\partial_i\partial^j\psi^{(1)}))\ . \tag{306}$$

[24] Notice that formula (36) in Ref. [175] for the estimate of the non-linearity parameter contains a sign misprint and should read $f_{\rm NL}^{\phi} \simeq -5/4r$, giving $f_{\rm NL}^{\phi} \simeq -5/4$ for $r \simeq 1$.

We then read the non-linearity parameter for the gravitational potential $\phi = \phi^{(1)} + \frac{1}{2}\phi^{(2)}$ [33]

$$f_{\rm NL}^{\phi} = \tfrac{3}{4} + g(\mathbf{k}_1, \mathbf{k}_2) \ , \tag{307}$$

with $g(\mathbf{k}_1, \mathbf{k}_2)$ defined in Eq. (302).

We would like to remind that the result in Eq. (307) has been obtained under certain minimal conditions for the inhomogeneous reheating to take place. This includes the assumption that during inflation $\Gamma_* \ll H$, and that the decay rate is completely determined by a scalar field χ as $\Gamma \propto \chi^2$. However, the curvature perturbation produced in the inhomogeneous reheating scenario does have a dependence on the ratio Γ_*/H, which one can parametrize as [300]

$$\zeta^{(1)} = -\alpha \frac{\delta^{(1)}\Gamma}{\Gamma_*} \ , \tag{308}$$

where α is positive and decreases as the ratio Γ_*/H at the end of inflation increases, with $\alpha = \frac{1}{6}$ in the limit $\Gamma_*/H \to 0$, thus recovering Eq. (284). Moreover the scalar field χ might set actually only a decay channel into which the inflaton decays; in addition the decay rate could have another channel which does not fluctuate, so that

$$\Gamma = \Gamma_0 + \Gamma_1 \left(\frac{\chi}{\chi_*}\right)^2 \ , \tag{309}$$

as considered in Refs. [76,300]. The resulting linear curvature perturbation reads

$$\zeta^{(1)} = -2\alpha \frac{\Gamma_1}{\Gamma_*} \frac{\delta^{(1)}\chi}{\chi_*} \ . \tag{310}$$

As argued in Ref. [300], Eq. (310) shows that the efficiency to generate the density perturbations can actually be very small either when the decay rate is not much smaller than the Hubble rate during inflation, or because the scalar field χ controls only one of the channels in which the inflaton field decays. Therefore, according to the previous considerations one expects that due to these effects the level of non-Gaussianity can be in fact higher than in Eq. (307). Moreover, as discussed in Ref. [300] the decay rate could depend by several scalar fields and one should also account for the possible presence of intrinsic non-Gaussianities in the scalar field(s) χ, produced by self-interactions of the type described in Section 2.5. In our formalisms this means that in expanding $\Gamma(t, \mathbf{x})$ as in Eq. (289) there might be an additional non-linear term given by $\delta^{(2)}\chi$ sourced by the self-couplings of the scalar field χ. As shown in Ref. [300] all these "variations on the theme" should increase the non-Gaussianity at a level very close to the limits set by WMAP.[25]

In fact following the same steps which lead to Eq. (307) we are able to extend the result to the more general case which includes the dependence on the α parameter and a decay rate as in Eq. (309), while keeping track of the different second-order effects arising in the determination of the non-linearity parameter of the gravitational potential. Using Eqs. (308) and (309) we find

$$f_{\rm NL}^{\phi} = \tfrac{3}{4} + I + g(\mathbf{k}_1, \mathbf{k}_2) \ , \tag{311}$$

[25] Notice however that the analysis in Ref. [300] focuses on the curvature perturbation ζ and does not take into account all the second-order effects which contribute to the level of the non-Gaussianity in the gravitational potential and in the CMB anisotropies which is actually the observable quantity.

where

$$I = -\frac{5}{2} + \frac{5}{12}\frac{\bar{\Gamma}}{\alpha\Gamma_1} \ , \tag{312}$$

$\bar{\Gamma}$ being the mean value of the decay rate. Thus the "minimal case" (307) is recovered for $\alpha = \frac{1}{6}$ and $\Gamma_1 = \bar{\Gamma}$ ($I = 0$).

8.3.1. Some remarks on the large-scale limit

Let us now clarify here again our procedure in deriving the expression of second-order quantities or equations of motion. Indeed, when dealing with second-order perturbations which are expressed in terms of first-order quantities, also the short-wavelength behavior of the first-order perturbations must be taken into account, as it becomes evident going to momentum space. The crucial point here is which is the final quantity one is interested in. We are interested in calculating the bispectrum of the gravitational potential and of the temperature anisotropies on large scales as a measure of the non-Gaussianity of the cosmological perturbations on those scales. The bispectrum of these quantities is twice the kernel which appears when expressing these second-order quantities in Fourier space, that is to say, e.g., $f_{\rm NL}^{\phi}(\mathbf{k}_1, \mathbf{k}_2)$ in Eq. (296). This means that, when calculating the bispectrum, we can evaluate the kernel in the long-wavelength limit, irrespective of the integration over the whole range of momenta. This is the reason why we have used the long-wavelength approximation in the equations of motion when deriving the expressions of second-order quantities in terms of first-order perturbations. Thus, the final result for the bispectrum is not affected by our procedure.

8.4. Second-order temperature fluctuations on large scales and the correct definition of the measured $f_{\rm NL}$

In this subsection we provide the expression for the second-order temperature fluctuations on large-scales which will allow the exact definition of the non-linearity parameter $f_{\rm NL}$. From now on, we will adopt the *Poisson gauge* defined in Section 3.1.

The second-order expression for the temperature fluctuation field in the Poisson gauge has been obtained in Ref. [199], by implementing the general formalism introduced in Ref. [234]. We are interested here in the large-scale limit of that expression, which allows to unambiguously define the primordial non-Gaussian contribution. Keeping only the large-scale limit of the linear and second-order terms in Eqs. (2.27) and (2.28) of Ref. [199], (see also Eqs. (4.11) and (4.12) of Ref. [234]), we obtain

$$\frac{\Delta T}{T} = \phi_{\mathscr{E}}^{(1)} + \tau_{\mathscr{E}}^{(1)} + \frac{1}{2}\left(\phi_{\mathscr{E}}^{(2)} + \tau_{\mathscr{E}}^{(2)}\right) - \frac{1}{2}\left(\phi_{\mathscr{E}}^{(1)}\right)^2 + \phi_{\mathscr{E}}^{(1)}\tau_{\mathscr{E}}^{(1)} \ , \tag{313}$$

where $\phi_{\mathscr{E}} = \phi_{\mathscr{E}}^{(1)} + \frac{1}{2}\phi_{\mathscr{E}}^{(2)}$ is the lapse perturbations at emission on the last scattering surface and $\tau_{\mathscr{E}} = \tau_{\mathscr{E}}^{(1)} + \frac{1}{2}\tau_{\mathscr{E}}^{(2)}$ is the intrinsic fractional temperature fluctuation at emission

$$\tau_{\mathscr{E}} \equiv \left.\frac{\Delta T}{T}\right|_{\mathscr{E}} \ . \tag{314}$$

Let us recall that, at linear order $\phi^{(1)} = \psi^{(1)}$. In Eq. (313) we dropped all those terms which represent *integrated* contributions such as Integrated Sachs–Wolfe, Rees–Sciama and more complicated

second-order integrated effects [172,211,196,268,69]. A full account of these effects is indeed provided by the general expressions for non-linear temperature anisotropies given in Refs. [233,234,199] and will not be reported here. The form of the CMB temperature bispectrum arising from some non-linear effects has been inferred recently in Ref. [73] (see also Ref. [18]), for a particular triangle configuration. Notice that for a ΛCDM cosmology the integrated Sachs–Wolfe effect would also give a contribution on large-scales. In order to compute this effect one should study the complete evolution of the gravitational potentials from last scattering till Λ (dark energy) domination.

It is important here to stress that the non-linearity parameter $f_{\rm NL}$ as introduced e.g. in Refs. [143,137] singles out the large-scale part of the second-order CMB anisotropies. One should be able to distinguish secondary integrated terms from the large-scale effects thanks to their specific angular-scale dependence. For the very same reason, we disregarded gravitational-lensing and Shapiro time-delay effects, Doppler terms and all those second-order effects which are characterized by a high-ℓ harmonic content. We finally dropped contributions at the observer position, which only modify the *monopole* term.

To obtain the intrinsic anisotropy in the photon temperature, we can expand the photon energy density $\rho_\gamma \propto T^4$ up to second order and write

$$\tau_{\mathscr{E}}^{(1)} = \frac{1}{4}\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma}\bigg|_{\mathscr{E}} , \tag{315}$$

where ρ_γ is the mean photon energy density, and

$$\tau_{\mathscr{E}}^{(2)} = \frac{1}{4}\frac{\delta^{(2)}\rho_\gamma}{\rho_\gamma}\bigg|_{\mathscr{E}} - 3(\tau_{\mathscr{E}}^{(1)})^2 = \frac{1}{4}\frac{\delta^{(2)}\rho_\gamma}{\rho_\gamma}\bigg|_{\mathscr{E}} - \frac{3}{16}\left(\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma}\bigg|_{\mathscr{E}}\right)^2 . \tag{316}$$

Next, we need to relate the photon energy density fluctuation to the lapse perturbation, which we can easily do by implementing the adiabaticity condition up to second order. At first order the adiabaticity condition reads $\zeta_m^{(1)} = \zeta_\gamma^{(1)}$ and we obtain

$$\frac{\delta^{(1)}\rho_\gamma}{\rho_\gamma} = \frac{4}{3}\frac{\delta^{(1)}\rho_m}{\rho_m} , \tag{317}$$

where ρ_m is the average energy density of the matter component. At second order the adiabaticity condition imposes $\zeta_m^{(2)} = \zeta_\gamma^{(2)}$, as explained in Section 3.6.2. From Eq. (135) applied to matter and radiation we find

$$\frac{\delta^{(2)}\rho_\gamma}{\rho_\gamma} = \frac{4}{3}\frac{\delta^{(2)}\rho_m}{\rho_m} + \frac{4}{9}\left(\frac{\delta^{(1)}\rho_m}{\rho_m}\right)^2 . \tag{318}$$

In the large-scale limit, the energy constraints (151) and (303) in the matter-dominated era, yields

$$\frac{\delta^{(1)}\rho_m}{\rho_m} = -2\psi^{(1)} \tag{319}$$

and

$$\frac{\delta^{(2)}\rho_m}{\rho_m} = -2\phi^{(2)} + 8(\psi^{(1)})^2 . \tag{320}$$

We finally obtain the fundamental relation

$$\frac{\Delta T}{T} = \frac{1}{3}\left[\psi^{(1)}_{\mathscr{E}} + \frac{1}{2}\left(\phi^{(2)}_{\mathscr{E}} - \frac{5}{3}(\psi^{(1)}_{\mathscr{E}})^2\right)\right] . \tag{321}$$

From Eq. (321), it is clear that the expression for the second-order temperature fluctuations is *not* a simple extension of the first-order Sachs–Wolfe effect $\Delta T^{(1)}/T = \psi^{(1)}_{\mathscr{E}}/3$ to second order since it receives a correction provided by the term $-(5/3)(\psi^{(1)}_{\mathscr{E}})^2$.

We can express the lapse function at second order as in Eq. (296), or equivalently as a general convolution (see, e.g., Ref. [3])

$$\phi = \phi^{(1)} + \tfrac{1}{2}\phi^{(2)} = \psi^{(1)} + f^{\phi}_{\rm NL} * (\psi^{(1)})^2 , \tag{322}$$

up to a constant offset. In order to connect the inflationary predictions with the definition of $f_{\rm NL}$ which has become standard in the CMB-related literature (see, e.g., Ref. [143]) we remind the standard Sachs–Wolfe formula

$$\frac{\Delta T}{T}(\hat{\mathbf{n}}, \tau_0) = -\frac{1}{3}\Phi(\hat{\mathbf{n}}(\tau_0 - \tau_{\mathscr{E}})) , \tag{323}$$

where $\Phi \equiv -\phi$ is Bardeen's potential [20], which is conventionally expanded in a form analogous to Eq. (322), namely

$$\Phi = \Phi_{\rm L} + f_{\rm NL} * (\Phi_{\rm L})^2 \tag{324}$$

(up to a constant offset, which only affects the temperature monopole term), where $\Phi_{\rm L} = -\phi^{(1)}$.

8.4.1. Angular averaging in a perturbed Universe and the value of $f_{\rm NL}$

One more non-linear effect that one should take into account is provided by the angular averaging implicit in the definition of observables such as the harmonic amplitudes of the CMB temperature as defined by an observer. Restricting ourselves to the pure Sachs–Wolfe effect, this amounts to performing an angular average with the physical (perturbed) metric on null hypersurfaces at fixed radial distance from the receiver. One can easily show that only the first-order correction to the metric gives a contribution to second-order quantities like the bispectrum and the effect can be accounted for by multiplying the angular differential element $\mathrm{d}\Omega$ by a conformal factor $(1 - 2\psi^{(1)})$. This operation implies shifting the value of

$$f^{\phi}_{\rm NL} \to (f^{\phi}_{\rm NL} - 1) . \tag{325}$$

It is interesting at this point to consider the particular "squeezed" configuration considered in Ref. [182], which consists in taking one of the wave numbers to be much smaller than the other two in the bispectrum $\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\rangle$, for instance $k_3 \ll k_1, k_2$. It is immediate to verify that in such a limit, and taking into account the form of the function (302), the bispectrum vanishes in the case of a scale-invariant power-spectrum within the standard scenario, where cosmological perturbations are due to the inflaton field. This is in full agreement with the general argument given in Ref. [182], where it was shown that in the squeezed limit the effect of the perturbation with the lowest momentum is only to rescale the other momenta in the corresponding fluctuations. For perturbations generated during inflation, this amounts to saying that fluctuations leaving the horizon at much earlier times act as a classical background for the evolution of the other modes.

From Eqs. (325), (321) and (324) we can now immediately derive the *true* non-linearity parameter $f_{\rm NL}$ which is the quantity actually measurable by high-resolution CMB experiments, after properly subtracting instrumental noise, foreground contributions and small-scale second-order terms. We find

$$f_{\rm NL} = -f_{\rm NL}^{\phi} + \tfrac{5}{6} + 1 = -f_{\rm NL}^{\phi} + \tfrac{11}{6} \ . \tag{326}$$

We warn the reader that this is the quantity which enters in the determination of higher-order statistics (such as the bispectrum of the temperature anisotropies) and to which the phenomenological study performed e.g. in Ref. [143] applies. A number of present and future CMB experiments, such as WMAP [139] and Planck, have enough resolution to either constrain or detect non-Gaussianity of CMB anisotropy data parametrized by $f_{\rm NL}$ with high precision [143].

Notice that the CMB temperature bispectrum does not vanish in the so-called squeezed limit discussed in Ref. [182], owing to the presence of second-order Sachs–Wolfe-like terms which give the extra term $\frac{5}{6}$ in Eq. (326). This statement applies to all scenarios of generation of cosmological perturbations and contrasts with the results in Ref. [73] (in the limit in which the two calculations can be compared), where the CMB bispectrum has been inferred directly from that of the gravitational potential (so that second-order Sachs–Wolfe-like terms were not included) and there is no matching to the primordial non-Gaussianity.

In Refs. [31,33] the matching among different cosmological eras has been obtained using the gauge-invariant curvature perturbation $\zeta^{(2)}$ defined in Eq. (134). If one goes to the uniform energy-density gauge defined by $\delta\rho = 0$ at any order, one recovers (at second order) the Salopek–Bond curvature perturbation defined in Ref. [249] $\zeta_{\rm SB}$ through the metric $\mathrm{d}s^2 = a^2(\tau)[-\mathrm{d}\tau^2 + e^{2\zeta_{\rm SB}}\,\mathrm{d}\mathbf{x}^2]$. Indeed, expanding $\zeta_{\rm SB}$ as $\zeta_{\rm SB} = \zeta_{\rm SB}^{(1)} + \frac{1}{2}\zeta_{\rm SB}^{(2)}$ and comparing to the metric (103), one immediately finds (on super-horizon scales)

$$\begin{aligned} \zeta^{(1)} &= -\psi^{(1)}|_{\rho} = \zeta_{\rm SB}^{(1)} \ , \\ \zeta^{(2)} &= -\psi^{(2)}|_{\rho} = \zeta_{\rm SB}^{(2)} + 2(\zeta_{\rm SB}^{(1)})^2 \ . \end{aligned} \tag{327}$$

The extra-term $2(\zeta_{\rm SB}^{(1)})^2$ beautifully matches the last term in the R.H.S. of Eq. (242) immediately explaining why $\zeta_{\rm SB}^{(2)} = \mathcal{O}(\epsilon, \eta)$ during inflation, as found in Ref. [182]. Notice, however, that the computation of the second-order temperature anisotropy in any gauge requires the use of the full gauge-invariant quantity $\zeta^{(2)}$ in order to properly account for terms proportional to $(\psi^{(1)})^2$. This step seems to be missing in Ref. [73], where both the uniform energy-density and the longitudinal gauges have been used.

After showing how the large scales perturbations in the gravitational potentials produce corresponding fluctuations in the CMB temperatures, we are finally in the position to give the predictions for the level of the non-Gaussianity in the three scenarios considered so far.

A. Standard scenario: Using Eqs. (326) and (301), we conclude that in the standard scenario where cosmological perturbations are generated by the inflaton field, the value of the non-Gaussianity parameter is provided by

$$f_{\rm NL}(\mathbf{k}_1, \mathbf{k}_2) \simeq \tfrac{7}{3} - g(\mathbf{k}_1, \mathbf{k}_2) \ . \tag{328}$$

B. Curvaton scenario: From expression (326) and Eq. (305), we find that the level of non-Gaussianity in the curvaton scenario is given by

$$f_{\rm NL} = -\left[-\frac{2}{3} + \frac{5}{6}r - \frac{5}{4r}\right] - g(\mathbf{k}_1, \mathbf{k}_2) \ , \tag{329}$$

where we recall that $r \approx (\rho_\sigma/\rho)_{\rm decay}$ is the ratio of the curvaton energy density to the total energy density at the curvaton decay.

C. Inhomogeneous reheating scenario: Using expression (326) and Eq. (307), we find that the level of non-Gaussianity in the inhomogeneous reheating scenario where $\Gamma \propto \chi^2$ is provided by

$$f_{\rm NL} = \tfrac{13}{12} - g(\mathbf{k}_1, \mathbf{k}_2) \ . \tag{330}$$

As explained in Section 8.3 we can relax some conditions and obtain an extension of Eq. (330) for a decay rate Γ which is only partially controlled by a scalar field χ as in Eq. (309). Using (326) and Eq. (311) we find

$$f_{\rm NL} = \tfrac{13}{12} - I - g(\mathbf{k}_1, \mathbf{k}_2) \ , \tag{331}$$

where $I = -5/2 + (5/12)\bar{\Gamma}/(\alpha\Gamma_1)$, with the parameter $0 < \alpha < 1/6$ and $\bar{\Gamma}/\Gamma_1 > 1$.

8.4.2. A comment on primordial non-Gaussianity and the post-inflationary evolution

The expressions for the non-linearity parameter $f_{\rm NL}$ obtained in the previous section are the results of three physical processes. The first one is the generation of an intrinsic non-Gaussianity during a primordial epoch and is strictly dependent on the particular mechanism which gives rise to the cosmological perturbations. This contribution sets the initial conditions for the evolution of the second-order perturbations in the radiation-dominated epoch obtained using the conservation of the curvature fluctuation $\zeta^{(2)}$, namely Eqs. (244), (276) and (295). The initial contribution is then processed by the second-order gravitational dynamics in the post-inflationary evolution given by Eq. (303) and the constraint equation (155). Finally, the non-linearities thus produced in the gravitational potential are transferred to the temperature anisotropies on large scales, where new second-order corrections arise (see Eq. (321)). In fact, it has been shown in Ref. [34] that the initial contribution from the primordial epoch can be neatly disentangled from the other contributions coming from the post-inflationary evolution. The key point here is that the gauge-invariant comoving curvature perturbation $\zeta^{(2)}$ remains *constant* on super-horizon scales after it has been generated and possible isocurvature perturbations are no longer present. Therefore, $\zeta^{(2)}$ provides us with all the necessary information about the "primordial" level of non-Gaussianity generated either during inflation, as in the standard scenario, or immediately after inflation, as in the curvaton scenario. Different scenarios are characterized by different values of $\zeta^{(2)}$, while the post-inflationary non-linear evolution induced by gravity is common to all scenarios [31–33].[26] For example, in standard single-field inflation$\zeta^{(2)}$ is generated during inflation and its value is given by $\zeta^{(2)} = 2(\zeta^{(1)})^2 + \mathcal{O}(n_\zeta - 1)$ [3,31] (as it can be seen also from Eqs. (242) and (243)). Notice that such a disentanglement can be performed unambiguously only by expressing the temperature anisotropies in a gauge-invariant way and identifying the primordial content in the gauge-invariant curvature perturbation $\zeta^{(2)}$.[27]

In Ref. [34] it has been shown that the general expression for the second-order temperature anisotropies given in Refs. [234,199] is in fact gauge-invariant under a time-shift $\tau \to \tau - \alpha_{(1)} + \frac{1}{2}(\alpha'_{(1)}\alpha_{(1)} - \alpha_{(2)})$ and that it is indeed possible to express the temperature anisotropies by properly defining some gauge-invariant metric and density perturbations. Notice that the gauge-invariance refers to the contributions to the temperature anisotropies on all scales. The definition of the gauge-invariant quantities proceeds by

[26] Once the initial conditions are set, one uses the same equations (303), (155) and (321).

[27] The observable large-scale temperature anisotropies are then given by the different contributions, and in that respect we used the word "primordial" in Section 8.4 to define the overall non-Gaussianity which survives on large-scales.

choosing the shifts $\alpha^{(r)}$ such that $\omega^{(r)}=0$.[28] For example the gauge-invariant definition of the gravitational potential $\phi^{(2)}$ reads

$$\phi^{(2)}_{\rm GI} = \phi^{(2)} + \omega^{(1)}\left[2\left(\psi^{(1)'} + 2\frac{a'}{a}\psi^{(1)}\right) + \omega^{(1)''} + 5\frac{a'}{a}\omega^{(1)'} + (\mathscr{H}' + 2\mathscr{H}^2)\omega^{(1)}\right] + 2\omega^{(1)'}(2\psi^{(1)} + \omega^{(1)'}) + \frac{1}{a}(a\alpha^{(2)})' \,, \tag{332}$$

where

$$\alpha^{(2)} = \omega^{(2)} + \omega^{(1)}\omega^{(1)'} + \nabla^{-2}\partial^i[-4\psi^{(1)}\partial_i\omega^{(1)} - 2\omega^{(1)'}\partial_i\omega^{(1)}] \,. \tag{333}$$

In terms of gauge-invariant quantities, the large-scale limit brings the same expression as in Eq. (313) where now each quantity is given by the corresponding gauge-invariant definition [34]. In the large-scale limit one again drops all those terms which represent integrated contributions and other second-order small-scale effects that can be distinguished from the large-scale part through their peculiar scale dependence. Taking the explicit expression of the conserved curvature perturbation $\zeta^{(2)}$ from Eq. (134) for a matter-dominated epoch, and using the $(0-0)$ component together with the traceless part of the $(i-j)$ Einstein equations at second order, one finds that on large scales the gauge-invariant expression for the temperature anisotropies reads [34]

$$\frac{\Delta T^{(2)}_{\rm GI}}{T} = \frac{1}{18}(\psi^{(1)}_{\rm GI})^2 - \frac{\mathscr{K}}{10} - \frac{1}{10}[\zeta^{(2)}_{\rm GI} - 2(\zeta^{(1)}_{\rm GI})^2] \,, \tag{334}$$

where

$$\psi^{(1)}_{\rm GI} = \psi^{(1)} - \mathscr{H}\omega^{(1)} \tag{335}$$

is the gauge-invariant definition of the linear gravitational potential $\psi^{(1)}$, $\zeta^{(1)}_{\rm GI}$ and $\zeta^{(2)}_{\rm GI}$ are large-scale curvature perturbations $\zeta^{(1)}$ and $\zeta^{(2)}$ expressed in terms of our gauge-invariant quantities, e.g. $\zeta^{(1)}_{\rm GI} = -\psi^{(1)}_{\rm GI} - \mathscr{H}(\delta^{(1)}\rho_{\rm GI}/\rho')$, and we have introduced a kernel

$$\mathscr{K} = 10\nabla^{-4}\partial_i\partial^j(\partial^i\psi^{(1)}_{\rm GI}\partial_j\psi^{(1)}_{\rm GI}) - \nabla^{-2}(\tfrac{10}{3}\partial^i\psi^{(1)}_{\rm GI}\partial_i\psi^{(1)}_{\rm GI}) \,. \tag{336}$$

Eq. (334) clearly shows that there are two contributions to the final non-linearity in the large-scale temperature anisotropies. The contribution $[\zeta^{(2)}_{\rm GI} - 2(\zeta^{(1)}_{\rm GI})^2]$, comes from the "primordial" conditions set during or immediately after inflation. It is encoded in the curvature perturbation ζ which remains constant once it has been generated. The remaining part of Eq. (334) describes the post-inflationary processing of the primordial non-Gaussian signal due to the non-linear gravitational dynamics. Thus, the expression in Eq. (334) allows to neatly disentangle the primordial contribution to non-Gaussianity from the one coming from the post-inflationary evolution. While the non-linear evolution after inflation is the same in each scenario, the primordial content will be different and depending on the particular mechanism generating the cosmological perturbations. We parametrize the primordial non-Gaussianity in terms of the conserved curvature perturbation (in the radiation or matter-dominated epochs)

$$\zeta^{(2)} = 2a(\zeta^{(1)})^2 \,, \tag{337}$$

[28] In fact one could easily extend such a procedure by including general coordinate transformations such that the expression of the gauge-invariant perturbations actually turns out to coincide with that of the corresponding quantities in the Poisson gauge.

where a will depend on the physics of the given scenario. For example in the curvaton case $a=(3/4r)-r/2$, where $r \approx (\rho_\sigma/\rho)_{\rm D}$ is the relative curvaton contribution to the total energy density at the curvaton decay, as it follows from Eqs. (261) and (274). In the minimal picture for the inhomogeneous reheating scenario from Eq. (294) we find $a=\frac{1}{4}$. From Eq. (334) we can extract the non-linearity parameter $f_{\rm NL}$ which is usually adopted to parametrize in a phenomenological way the level of non-Gaussianity in the cosmological perturbations and has become the standard quantity to be observationally constrained by CMB experiments [143,139]. Using the parametrization (337) and $\zeta^{(1)}=-\frac{5}{3}\,\psi^{(1)}_{\rm GI}$ during matter domination, from Eqs. (323) and (324) we immediately read the non-linearity parameter in momentum space

$$f_{\rm NL}(\mathbf{k}_1,\mathbf{k}_2)=-[\tfrac{5}{3}\,(1-a)+\tfrac{1}{6}-\tfrac{3}{10}\,\mathscr{K}]+1\ , \tag{338}$$

where $\mathscr{K}=10(\mathbf{k}_1\cdot\mathbf{k}_3)(\mathbf{k}_2\cdot\mathbf{k}_3)/k^4-(10/3)(\mathbf{k}_1\cdot\mathbf{k}_2)/k^2$ with $\mathbf{k}_3+\mathbf{k}_1+\mathbf{k_2}=0$ and $k=|\mathbf{k}_3|$. In fact formula (338) already takes into account the additional non-linear effect entering in the angular three-point function of the CMB from the angular averaging performed with a perturbed line-element $(1-2\psi^{(1)}_{\rm GI})\,{\rm d}\Omega$, implying a $+1$ shift in $f_{\rm NL}$. Notice that the procedure to get Eq. (334) and (338) is the same that we have used to compute the final values of the non-linearity parameter in the previous section, the only difference being that instead of determining the non-linearity in the gravitational potential $\phi^{(2)}$ and from that deducing $f_{\rm NL}$ through Eq. (321) here we kept track of the curvature perturbation $\zeta^{(2)}$ in the final expression for the temperature anisotropies. In fact it is immediate to recover Eqs. (328)–(330) taking into account that

$$\mathscr{K}=5-\tfrac{10}{3}\,g(\mathbf{k}_1,\mathbf{k}_2)\ . \tag{339}$$

In particular, within the standard scenario where the cosmological perturbations are due to the inflaton field, the primordial contribution to the non-Gaussianity is given by $a=1-\frac{1}{4}(n_\zeta-1)$ [3,31] and the non-linearity parameter from inflation now reads

$$f^{\rm inf}_{\rm NL}=-\tfrac{5}{12}\,(n_\zeta-1)+\tfrac{5}{6}+\tfrac{3}{10}\,\mathscr{K}\ . \tag{340}$$

Therefore, the main contribution to non-Gaussianity comes from the post-inflationary evolution of the second-order cosmological perturbations which give rise to order-one coefficients, while the primordial contribution is proportional to $|n_\zeta-1|\ll 1$. This is true even in the "squeezed" limit first discussed by Maldacena [182], where one of the wave numbers is much smaller than the other two, e.g. $k_1\ll k_2$, k_3 and $\mathscr{K}\to 0$.

9. Other mechanisms generating non-Gaussian density perturbations

In this section we describe in some detail some scenarios to generate non-Gaussianities in the observed cosmological perturbations which represent a plausible alternative to the mechanisms already discussed. For these scenarios a complete analysis of the perturbation evolution and hence a precise determination of the level of the non-Gaussianity in the large-scale CMB anisotropies is still missing, but nevertheless they offer some general and interesting insight on the ways non-Gaussianities can be produced from an inflationary epoch at a higher level than predicted by the single-field models of slow-roll inflation.

9.1. Non-Gaussianities from multiple interacting scalar fields during inflation

The standard models of inflation are based on the simple assumption that only the inflaton field is relevant both for the background evolution and for the produced density perturbations. However, especially on particle physics grounds, it is hard to believe that only one single scalar field φ plays a role during the inflationary stage. On the contrary it is quite natural that during the inflationary dynamics several other scalar fields χ_I ($I = 1, \ldots, N$) are present. The contribution to the total energy density of the extra scalar fields χ_I might or might not be negligible, compared to the one provided by the scalar field φ. If the latter is the case, then the model of inflation is called a multiple-field model [164,274,231,166,129]. However, as soon as one considers more than one scalar field, one must also consider the role of the isocurvature perturbations produced by the relative fluctuations of the scalar fields [166,129,132] in addition to the usual adiabatic mode. It is well known that in such a framework non-Gaussian *isocurvature* perturbations can be produced [11,200,295–298,249,248,168,54,222–224]. The disadvantage of these scenarios is that in general the observed pattern of CMB anisotropies and the observations of LSS constrain the amount of isocurvature perturbations to contribute only a small fraction (see, e.g. Ref. [269]). However, until very recently the adiabatic and the isocurvature perturbation modes had been considered as statistically independent, without taking into account that there can be a non-vanishing correlation between the curvature and the entropy modes. The physical origin of this correlation is actually due to the fact that the entropy mode on large scales can feed the adiabatic curvature perturbation as described by Eqs. (144) and (147). It is just such a cross-correlation produced during an inflationary epoch when several scalar fields are present that can introduce non-Gaussianity in the adiabatic mode too a it has been first suggested in Ref. [30].

Before entering in some details let us here summarize the underlying idea of this mechanism. The starting point is the simple observation that it is quite natural to expect that the inflaton field is coupled to the extra scalar fields present during inflation. It has been shown that such a coupling gives rise to a new mechanism for generating quantum fluctuations in the scalar fields, which was dubbed the *oscillation mechanism* in Ref. [28]. In this case the quantum fluctuations of the scalar fields are not generated only because of gravitational amplification during the de Sitter epoch as described in Section 2.4, but also because—due to the interaction terms—the quantum fluctuations of a scalar field χ can oscillate (evolve) into fluctuations of the scalar field φ with a calculable probability, in a way similar to the phenomenon of neutrino oscillations. The probability of oscillation is resonantly amplified when the perturbations leave the horizon and the perturbations in the scalar field χ may disappear at horizon crossing giving rise to perturbations in the scalar field φ. Adiabatic and entropy perturbation are inevitably correlated at the end of inflation [28]. The crucial observation is that—since the degree of mixing is governed by the squared mass matrix of the scalar fields—the oscillations can take place even if the energy density of the extra scalar fields is much smaller than the energy density of the inflaton field. This is an important point. Gaussian perturbations are usually expected in inflationary models because the inflaton potential is required to be very flat. This amounts to saying that the interaction terms in the inflaton potential are present, but small and non-Gaussian features are suppressed since the non-linearities in the inflaton potential are suppressed too. On the other hand, nothing prevents the inflaton field from being coupled to another scalar degree of freedom whose energy density is much smaller than the one stored in the inflaton field. This extra scalar field will not be constrained by slow-roll conditions and it is natural to expect that the self-interactions of such an extra field or the interaction terms with the inflaton field are sizeable, thereby representing potential sources for non-Gaussianity. If during the inflationary epoch, oscillations

between the perturbation of the inflaton field and the perturbations of the other scalar degrees of freedom occur, the non-Gaussian features generated in the system of the extra field are efficiently communicated to the inflaton sector. As it has been shown in Ref. [30] these non-Gaussianities can be left imprinted in the CMB anisotropies.

9.1.1. The oscillation mechanism

We now briefly describe the oscillation mechanism with an illustrative example. Consider two scalar fields, φ and χ interacting through a generic potential $V(\varphi, \chi)$. We will dub φ the inflaton field, even if this might be a misnomer as the two fields might give a comparable contribution to the total energy density of the Universe. In Fourier space the Klein–Gordon equations read

$$\delta\ddot{\varphi} + 3H\delta\dot{\varphi} + \frac{k^2}{a^2}\,\delta\varphi + V_{\varphi\varphi}\delta\varphi + V_{\varphi\chi}\delta\chi = 0\ ,$$

$$\delta\ddot{\chi} + 3H\delta\dot{\chi} + \frac{k^2}{a^2}\,\delta\chi + V_{\chi\chi}\delta\chi + V_{\chi\varphi}\delta\varphi = 0\ , \tag{341}$$

where we have used the notation $V_{\varphi\varphi} \equiv (\partial^2 V/\partial\varphi\partial\varphi)$ and similarly for the other derivatives.

The interactions between the two scalar fields is manifest in that the squared mass matrix

$$\mathscr{M}^2 = \begin{pmatrix} V_{\varphi\varphi} & V_{\varphi\chi} \\ V_{\varphi\chi} & V_{\chi\chi} \end{pmatrix} \tag{342}$$

is in general non-diagonal. This introduces a mixing between the two scalar fields. To estimate such a mixing one can diagonalize the system of equations (341) by introducing a time-dependent 2×2 unitary matrix $\mathscr{U}$ such that $\mathscr{U}^\dagger\mathscr{M}^2\mathscr{U} = \mathrm{diag}(\omega_1^2, \omega_2^2) \equiv \omega^2$. In the following we will assume that all the entries of the squared mass matrix $\mathscr{M}^2$ are real, so that the unitary matrix $\mathscr{U}$ reduces to an orthogonal matrix

$$\mathscr{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{343}$$

where $\tan 2\theta = 2V_{\chi\varphi}/(V_{\varphi\varphi} - V_{\chi\chi})$ and the mass eigenvalues are given by

$$\omega^2_{1,2} = \tfrac{1}{2}[(V_{\varphi\varphi} + V_{\chi\chi}) \pm \sqrt{(V_{\varphi\varphi} - V_{\chi\chi})^2 + 4V^2_{\chi\varphi}}]\,. \tag{344}$$

If for simplicity we work in the slow-roll approximation and assume that the entries of the squared mass matrix are constant in time, we obtain for the states $\Psi = (\Psi_1, \Psi_2)^{\mathrm{T}} = \mathscr{U}^{\mathrm{T}}(\varphi, \chi)^{\mathrm{T}}$

$$\delta\widetilde{\Psi}_1'' + \left(k^2 - \frac{a''}{a} + \omega_1^2 a^2\right)\delta\widetilde{\Psi}_1 = 0\ ,$$

$$\delta\widetilde{\Psi}_2'' + \left(k^2 - \frac{a''}{a} + \omega_2^2 a^2\right)\delta\widetilde{\Psi}_2 = 0\ . \tag{345}$$

These equations are exactly of the form of Eq. (41) and if we suppose a pure de Sitter phase the solutions for $\delta\widetilde{\Psi}_i$ are then given by Eq. (53), with $\nu_i^2 = 9/4 - (\omega_i/H)^2$. The mass eigenvalues can be simply expressed in terms of the slow-roll parameters $\epsilon_I = (1/16\pi G_{\mathrm{N}})(V_{\chi_I}/V)^2$ and $\eta_{IJ} = (1/8\pi G_{\mathrm{N}})(V_{\chi_I}V_{\chi_J}/V^2)$.

Since at a given time τ—the scalar perturbations $\delta\varphi$ and $\delta\chi$ are a linear combination of the scalar perturbations mass eigenstates $\delta\Psi_1$ and $\delta\Psi_2$,

$$\delta\varphi = \sum_{\ell=1,2} \mathscr{U}_{1\ell}\delta\Psi_\ell, \quad \delta\chi = \sum_{\ell=1,2} \mathscr{U}_{2\ell}\delta\Psi_\ell \ , \tag{346}$$

it is possible to calculate the probability that a scalar perturbation $\delta\chi$ at the time τ_0 becomes a scalar perturbation in the "inflaton" field $\delta\varphi$ at the time τ by computing $P[\delta\varphi(\tau_0) \to \delta\chi(\tau)] = |\langle\delta\varphi(\tau_0)\delta\chi^*(\tau)\rangle|^2$. As it has been shown in Ref. [28] on subhorizon scales $k \gtrsim aH$

$$P[\delta\varphi(\tau_0) \to \delta\chi(\tau)] \simeq 0 \quad (k \gg aH) \ , \tag{347}$$

but on superhorizon scales and in the limit $\omega_{1,2}^2 \ll H^2$ the conversion probability is non-vanishing

$$P[\delta\varphi(\tau_0) \to \delta\chi(\tau)] \simeq \sin^2 2\theta \sin^2\left(\frac{\pi}{12}\frac{\Delta\omega^2}{H^2}\right) \ . \tag{348}$$

Such a formula reminds the well-known formula which describes the evolution in time of the probability of oscillation between two neutrino flavors (see, e.g. Ref. [50]). This result shows that at horizon crossing there is a mechanism of amplification for the perturbations of the scalar field φ due to a conversion of the χ fluctuations into the inflaton perturbations.[29] In fact such an analysis can be extended to fully account for metric perturbations, and to include time-dependent terms in the squared mass matrix [28]. Still, the results are of the same form as in Eq. (348).

Two important remarks are in order. First of all, we wish to stress that the oscillation mechanism operates even if the energy of the inflaton field φ is much larger than the energy stored in the other scalar field χ. This is because what is crucial for the oscillations to occur is the *relative* magnitude of the elements of the squared mass matrix $\mathscr{M}^2$. Secondly, the magnitude of the probability depends upon two quantities, $\sin^2 2\theta$ and $\Delta\omega^2/H^2$. Both can be readily expressed in terms of the slow-roll parameters. The first factor is not necessarily small, in fact it may be even of order unity for maximal mixing. If expanded in terms of the slow-roll parameters, it is $\mathscr{O}(\eta^0, \epsilon^0)$. The second term is naturally smaller than unity and is linear in the slow-roll parameters. This reflects the fact that during inflation only perturbations in scalar fields with masses smaller than the Hubble rate may be excited. However, $\Delta\omega^2/H^2$ is not necessarily much smaller than unity and the amplification of the conversion probability at horizon crossing may be sizeable.

9.1.2. Transfer of non-Gaussianities

The oscillation mechanism is responsible for the transfer of non-Gaussianities from the isocurvature perturbation mode to the adiabatic mode. In order to see that, we can follow the elegant treatment of Ref. [93] to study adiabatic and entropy perturbations in the case of multiple interacting scalar fields. The

[29] The phenomenon of resonant amplification is easily understood if one remembers that a given wavelength leaves the horizon when $k = aH$, i.e. when $k^2 = a''/a$ using the conformal time. As long as the wavelength is subhorizon, $k^2 \gg a''/a$, the presence of the mass terms in the equations of motion (345) is completely negligible compared to the factor $(k^2 - a''/a)$. On the other hand, when the wavelength leaves the horizon the term $(k^2 - a''/a)$ vanishes and the effect of the mixing in the mass squared matrix is magnified, giving rise to the resonant effect. Finally, when the wavelength is larger than the horizon, $k^2 \ll a''/a$, the term $(k^2 - a''/a)$ starts to dominate again over the mass terms and the oscillations get frozen.

adiabatic and the entropy parts of the perturbations are expressed in terms of the original field fluctuations as

$$Q_A = (\cos\beta)\,Q_\varphi + (\sin\beta)\,Q_\chi \ , \tag{349}$$

$$\delta s = (\cos\beta)\,Q_\chi - (\sin\beta)\,Q_\varphi \ , \tag{350}$$

where

$$\cos\beta \equiv c_\beta = \frac{\dot\varphi}{\sqrt{\dot\varphi^2+\dot\chi^2}}, \quad \sin\beta \equiv s_\beta = \frac{\dot\chi}{\sqrt{\dot\varphi^2+\dot\chi^2}} \ . \tag{351}$$

Here we have used the gauge-invariant Sasaki–Mukhanov variables

$$Q_I \equiv \delta\chi_I + \frac{\dot\chi_I}{H}\psi \tag{352}$$

in order to take into account also the metric perturbations. For these fields the evolution equations are like Eqs. (341), with the squared mass matrix given by $\mathscr{M}^2_{IJ} = V_{\varphi_I\varphi_J} - 1/M_{\rm P}^2 a^3(a^3/H\dot\varphi_I\dot\varphi_J)^{\cdot} \simeq V/M_{\rm P}^2[\eta_{IJ} - 2(\pm\sqrt{\epsilon_I})(\pm\sqrt{\epsilon_J})]$, and the mixing angle reads $\tan 2\theta = 2\mathscr{M}^2_{\chi\varphi}/(\mathscr{M}^2_{\varphi\varphi} - \mathscr{M}^2_{\chi\chi})$.

The cross-correlation between the adiabatic and the entropy perturbations is

$$\langle Q_A(\mathbf{k})\delta s^*(\mathbf{k}')\rangle \equiv \frac{2\pi^2}{k^3}\,\mathscr{C}_{Q_A\delta s}\,\delta^{(3)}(\mathbf{k}-\mathbf{k}') \ , \tag{353}$$

in analogy with definition (59) for the power-spectrum of a given perturbation. Therefore, the origin of the cross-correlation is due to a rather transparent physical behavior in terms of the oscillation mechanism. During the inflationary epoch, the gauge invariant perturbations Q_φ and Q_χ are generated with different wavelengths stretched by the superluminal expansion of the scale-factor. Since the squared mass matrix of Q_φ and Q_χ is not diagonal, oscillations between the two quantities are expected. As long as the wavelength remains subhorizon, Q_φ and Q_χ evolve independently and may be considered good mass eigenstates. However, as soon as the wavelength leaves the horizon, an amplification in the probability of oscillation between Q_φ and Q_χ occurs: a non-vanishing correlation between Q_φ and Q_χ is created on superhorizon scales because of the non-diagonal mass matrix $\mathscr{M}^2_{IJ}$. Since the adiabatic and the isocurvature modes are a linear combination of Q_φ and Q_χ, at horizon crossing a non-vanishing correlation between the adiabatic and the isocurvature modes is left imprinted in the spectrum in the form[30]

$$a^2\langle Q_A(k)\delta s^*(k')\rangle = (s_\beta c_\theta - c_\beta s_\theta)(c_\beta c_\theta + s_\beta s_\theta)[|Q_\chi|^2 - |Q_\varphi|^2] \ . \tag{354}$$

In Ref. [93] it has been shown that the cross-correlation between the adiabatic and the entropy perturbations arise when the trajectories of the scalar fields in the background bend in the field space (φ, χ), which amounts to saying that $\dot\beta \neq 0$. Actually such a bending is simply due to the interactions between the two scalar fields [29,289]. In particular if β is constant no correlation is produced at the end of inflation. It could be the case where $\beta = 0$, which corresponds to the case where the scalar field χ is approximately static, or when there is some kind of attractor solution with $\varphi \propto \chi$. Notice from Eqs. (349)–(351) that in the former case the entropy perturbation is due entirely to the scalar field χ (in agreement with the results of Ref. [28]).

[30] For simplicity we quote from Ref. [28] the expression which neglects the time dependence of the square mass matrix of the fields Q_φ and Q_χ.

During inflation the (comoving) curvature perturbation is given by [93]

$$\mathscr{R} = \frac{H}{\dot{A}}\, Q_A, \quad \dot{A} = (\cos\beta)\dot{\phi} + (\sin\beta)\dot{\chi}\ , \tag{355}$$

while the entropy perturbation between the two scalar fields can be defined as

$$S_{\varphi\chi} = H\,\frac{\delta s}{\dot{A}}\ . \tag{356}$$

In fact one needs to consider the evolution of the perturbations throughout the reheating stage, and after the end of inflation in order to link the curvature and the entropy perturbations to their corresponding quantities, defined in the large-scale limit deep in the radiation era, which are actually the quantities that can be constrained observationally. However as explained in Sections 3.6.1 and 3.6.2 the adiabatic perturbation is sourced on large scales by the entropy mode, while an entropy perturbation cannot be generated on large scales from an adiabatic perturbation. Therefore generically one can describe the time evolution of the curvature and entropy perturbation modes on large scales as

$$\dot{\mathscr{R}} = \alpha H \mathscr{S}, \quad \dot{\mathscr{S}} = \beta H \mathscr{S}\ , \tag{357}$$

where α and β are in general time-dependent dimensionless functions. The explicit form of the coupling between the curvature and entropy perturbations will depend on the particular model under consideration: it has been computed for the case of interacting scalar fields [93,116,29,97,289] and non-interacting fluids [117]. In particular, as we already mentioned, in Ref. [93] it has been shown that in the case of two scalar fields $\alpha = 2\dot{\beta}/H$. Integrating Eqs. (357) one can parametrize the evolution of the perturbations on large scales through some transfer functions $T(t_*, t)$ relating curvature and entropy perturbations generated when a given mode is stretched outside the Hubble scale during inflation ($k = aH$, denoted by an asterisk) to curvature and entropy perturbations at some later time [12,289]

$$\begin{pmatrix} \mathscr{R} \\ \mathscr{S} \end{pmatrix} = \begin{pmatrix} 1 & T_{\mathscr{R}\mathscr{S}} \\ 0 & T_{\mathscr{S}\mathscr{S}} \end{pmatrix} \begin{pmatrix} \mathscr{R} \\ \mathscr{S} \end{pmatrix}_* , \tag{358}$$

where

$$\begin{aligned} T_{\mathscr{R}\mathscr{S}}(t_*, t) &= \int_{t_*}^{t} \alpha(t') T_{\mathscr{S}\mathscr{S}}(t_*, t') H(t')\, \mathrm{d}t'\ , \\ T_{\mathscr{S}\mathscr{S}}(t_*, t) &= \exp\left(\int_{t_*}^{t} \beta(t') H(t')\, \mathrm{d}t'\right)\ . \end{aligned} \tag{359}$$

For example, if the decay products of the reheating completely thermalize, then after the reheating process, in the radiation dominated era, $T_{\mathscr{S}\mathscr{S}} = 0$. On the other hand if among the decay products a CDM species remains decoupled, than an isocurvature perturbation between this component and the photons will survive after inflation. The simplest possibility is that one of the scalar fields (or its decay products) is just identified with the CDM.

The gravitational potential ϕ (in the longitudinal gauge) is indeed related to the curvature perturbation so that (at least at linear order) one can write for example at the beginning of the radiation epoch

$$\phi \simeq \frac{2}{3}\,\mathscr{R} = \mathscr{R}_* + T_{\mathscr{R}\mathscr{S}}\mathscr{S}_* = g_*(Q_{A*} + T_{\mathscr{R}\mathscr{S}}\delta s_*)\ , \tag{360}$$

where $g_* = (H/\dot{A})_*$ and we have used Eqs. (356) and (355). Therefore we are now in the position to estimate, for example, the bispectrum of the gravitational potential. According to the considerations of Section 2.5 it is reasonable to consider that the dominant term in the bispectrum is given by the terms proportional to $T_{\mathcal{R}\mathcal{S}}\delta s|_*$, so that we want to estimate

$$\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\rangle \propto T^3_{\mathcal{R}\mathcal{S}}\langle\delta s_*(\mathbf{k}_1)\delta s_*(\mathbf{k}_2)\delta s_*(\mathbf{k}_3)\rangle \ . \tag{361}$$

Eq. (361) shows that the correlation between the adiabatic and the entropy perturbations during inflation, parametrized by $T_{\mathcal{R}\mathcal{S}}$ actually sources the bispectrum of the gravitational potential (the adiabatic mode). Notice that this remains valid even if at the end of inflation only an adiabatic mode perturbation is left imprinted on very large-scales deep in the radiation era, for example if all the decay products of the scalar fields thermalize after the reheating stage. If there is enough time during inflation for the conversion from isocurvature to adiabatic perturbations to occur then a non-Gaussian adiabatic perturbation mode is generated. On the other hand, a residual isocurvature perturbation might survive the reheating stage, for example if the inflaton field decays into ordinary matter (the present day photons neutrinos and baryons) while the additional scalar field decays into decoupled dark matter, or it does not decay at all (like the case of an axion). In this case, as it has been shown in Ref. [30], the bispectrum of the CMB anisotropies receives two additional contributions, one from the intrinsic bispectrum of the isocurvature mode $\langle\mathcal{S}(\mathbf{k}_1)\mathcal{S}(\mathbf{k}_2)\mathcal{S}(\mathbf{k}_3)\rangle$, and the other from the bispectrum of cross-correlation terms of the type $\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\mathcal{S}(\mathbf{k}_3)\rangle$, providing a characteristic signatures of these non-Gaussian inflationary perturbations.

In both cases one has to estimate terms like the one appearing in Eq. (361) with a cubic combination of the transfer functions $T_{\mathcal{S}\mathcal{S}}$ and $T_{\mathcal{R}\mathcal{S}}$. We can do that in terms of the original scalar fields by evaluating

$$\begin{aligned}
\widehat{T}^3\langle\delta s_*(\mathbf{k}_1)\delta s_*(\mathbf{k}_2)\delta s_*(\mathbf{k}_3) &= \widehat{T}^3\langle\delta s_*(\mathbf{k}_1)\delta s_*(\mathbf{k}_1)\delta s_*(\mathbf{k}_1)\rangle \\
&= \widehat{T}^3\langle(c_\beta Q_{\chi 1} - s_\beta Q_{\varphi 1})(c_\beta Q_{\chi 2} - s_\beta Q_{\varphi 2})(c_\beta Q_{\chi 3} - s_\beta Q_{\varphi 3})\rangle \\
&= \widehat{T}^3[c^3_\beta\langle Q_{\chi 1}Q_{\chi 2}Q_{\chi 3}\rangle - c^2_\beta s_\beta\langle Q_{\chi 1}Q_{\chi 2}Q_{\varphi 3}\rangle - c^2_\beta s_\beta\langle Q_{\chi 1}Q_{\varphi 2}Q_{\chi 3}\rangle \\
&\quad + c_\beta s^2_\beta\langle Q_{\chi 1}Q_{\varphi 2}Q_{\varphi 3}\rangle - s_\beta c^2_\beta\langle Q_{\varphi 1}Q_{\chi 2}Q_{\chi 3}\rangle + s^2_\beta c_\beta\langle Q_{\varphi 1}Q_{\chi 2}Q_{\varphi 3}\rangle \\
&\quad + s^2_\beta c_\beta\langle Q_{\varphi 1}Q_{\varphi 2}Q_{\chi 3}\rangle - s^3_\beta\langle Q_{\varphi 1}Q_{\varphi 2}Q_{\varphi 3}\rangle] \ ,
\end{aligned} \tag{362}$$

where, for example, $Q_{\varphi 1}$ stands for $Q_\varphi(\mathbf{k}_1)$ and we have used Eq. (349), and for simplicity of notation we will omit the asterisk from now on. Also, we have used the notation $\widehat{T}$ to indicate that actually the proper analysis should be performed by extending these results to second order in the perturbations, and therefore $\widehat{T}$ should be considered as an effective transfer function accounting for the second-order effects.

Note that the bispectrum is a sum of different three-point correlation functions. The coefficients in front of each correlation function involve mixing angles which parametrize the amount of mixing between the adiabatic and the isocurvature modes. If such mixing is sizeable, all coefficients are of order unity and one expects that non-linearities in the perturbation of the scalar field χ may be efficiently transferred to the inflaton sector, thus generating large non-Gaussian features. We shall now consider some specific examples which may help in understanding how such a mechanism acts during inflation.

9.1.3. Some worked examples

One can envisage different situations:

(i) Inflation is driven by the inflaton field φ and there is another scalar field χ which does not interact with the inflaton and has a simple polynomial potential $V(\chi) \propto \chi^n$ leading to zero vacuum expectation value, $\langle\chi\rangle = 0$. In such a case, $\sin\beta = \sin\theta = 0$ and there is no mixing between the inflaton field and the χ-field as well as no cross-correlation between the adiabatic and isocurvature modes. Non-vanishing non-Gaussianity will be present in the isocurvature mode. This is indeed a known result [11,168]. We stress here that in particular for a potential like $V(\chi) \propto \chi^2$ the non-Gaussianity is generated in the *energy density* of the scalar field χ, while the scalar field is intrinsically Gaussian, as we discussed in Section 2.5. Actually this is the scenario considered in Ref. [168]. Notice that the same authors suggested that if the χ-field decays late after inflation into the CMB photons, the non-Gaussianity in its energy density will then be transferred to the final adiabatic perturbations, which is at the basis of the curvaton mechanism for the generation of non-Gaussian adiabatic perturbations.

(ii) Inflation is driven by two scalar fields φ and χ with equal mass, $V = \frac{m^2}{2}(\varphi^2 + \chi^2)$. In such a case the mixing is maximal, $\beta = \theta = \pi/4$. Nevertheless, the cross-correlation is again vanishing [93,28,29] and the bispectrum gets contributions from adiabatic and isocurvature modes independently. A term $\mu/3!\,\chi^3$ in the Lagrangian would be a source of non-Gaussianity and at the same time it would switch on a cross correlation between the adiabatic and the isocurvature modes, thus producing non-zero cross terms in the bispectrum of the type $\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\mathcal{S}(\mathbf{k}_3)\rangle$. However, these non-Gaussianities would be small because of the slow-roll conditions. In fact in such a situation, since the two scalar fields have equal mass, their evolution mimicks a single-field slow-roll inflaton.

(iii) Let us now sketch a general way for the oscillation mechanism to be operative. Let us suppose that inflation is driven by an inflaton field φ and there is another scalar field χ whose vacuum expectation value depends on the inflaton field and—eventually—on the Hubble parameter H and some other mass scale μ, $\langle\chi\rangle = f(\varphi, H, \mu)$. Under these circumstances, $\langle\dot{\chi}\rangle = (\partial f/\partial\varphi)\dot{\varphi} + (\partial f/\partial H)\dot{H}$. As in illustrative case, let us restrict ourselves to the case in which $(\partial f/\partial\varphi)\dot{\varphi}$ is the dominant term and we can approximate $\langle\dot{\chi}\rangle = (\partial f/\partial\varphi)\dot{\varphi}$. We have therefore $\tan\beta \simeq \partial f/\partial\varphi$ and $\dot{\beta} \simeq (\partial f/\partial\varphi)^{\cdot}/[1 + (\partial f/\partial\varphi)^2]$. In such a case, the cross-correlation between the adiabatic and the isocurvature modes may be large and non-Gaussianity may be efficiently transferred from one mode to the other.

An implementation of the transfer of non-Gaussianities from an isocurvature perturbation in the scalar field χ to the inflaton field φ has been given in Refs. [44,45]. A key point to bear in mind is that some kind of coupling between the inflaton field and the extra scalar field is needed for such transfer to occur. On the other hand we must always require the scalar field χ to have an effective mass which is less than the Hubble rate during inflation in order for χ to develop non-negligible fluctuations. The two requirements seem to act somewhat in opposite directions [44,45]. However, a model where the transfer is efficient is the one with a potential of form [44]

$$V(\varphi, \chi) = U(\varphi) + m_\times^2(\varphi - \varphi_0)\chi + \frac{\lambda}{n!}\chi^n \ . \tag{363}$$

The transfer of non-linearities from the χ sector to the inflaton can be easily understood looking at the Klein–Gordon equation for the inflaton field fluctuations at second order. Such an equation will contain a source term deriving from the coupling, which will be of the form $V_{\varphi\chi}\delta^{(2)}\chi \sim m_\times^2\delta^{(2)}\chi$. Therefore the scalar field χ will first develop some non-Gaussianites as described in Section 2.5 during inflation. Such non-Gaussianities are of isocurvature type since the scalar field χ remains subdominant and for cubic self

interactions they will be of the order $\delta^{(2)}\chi \sim (\delta^{(1)}\chi)^2\lambda(t-t_k)/H$, where t_k is the time of horizon crossing and H is the Hubble parameter during inflation. Then, through the coupling the inflaton will acquire such non-linearities

$$\delta^{(2)}\varphi \sim \int \mathrm{d}t\, \frac{m_\times^2}{H}\,\delta^{(2)}\chi \sim \lambda\,\frac{m_\times^2}{H^2}\,\frac{N_k^2}{H^2}(\delta^{(1)}\chi)^2 \sim \lambda N_k^2\,\frac{m_\times^2}{H^2}\ . \tag{364}$$

The condition that the scalar field χ is subdominant and light enough that $\dot{\chi} < \dot{\varphi}$ implies that $N_k < (H/m_\times)^2$ and it gives $\delta^{(2)}\chi < \lambda N_k/H^2$, which can be of the order of the Gaussian component $\delta^{(1)}\chi$. In Ref. [44] it has been argued that only with a quartic self-coupling for the scalar field χ it is possible to develop non-Gaussianities on a long time scale without any severe fine tuning on the model.

Of course the potential in Eq. (363) can be considered only as a toy model or as an effective potential at some stage of the inflationary dynamics. In Ref. [45] some more realistic particle physics realizations of the transfer of non-Gaussianities have been discussed. An example is the hybrid-type model of inflation involving three scalar fields. Let us consider the potential [45]

$$V(\varphi,\chi,\sigma) = \frac{1}{2}m^2\varphi^2 + \frac{\lambda}{4!}\chi^4 + \frac{\mu}{2}(\sigma^2-\sigma_0^2)^2 + \frac{g}{2}\sigma^2(\varphi\cos\alpha + \chi\sin\alpha)^2\ . \tag{365}$$

Here φ is the inflaton field, the second field χ is a light and subdominant scalar field with a quartic coupling and the third field σ is coupled to the other two scalar fields so as to trigger the end of inflation by a phase transition. Here σ_0 is the final vacuum expectation value of σ and α parametrizes the couplings of φ and χ to σ. For large values of φ, the scalar field σ is anchored to the origin. When φ reaches the critical value φ_{end} for which the effective mass of σ vanishes, then inflation ends and the fields roll down towards their true minima $\sigma = \pm\sigma_0$, $\varphi = 0$ and $\chi = 0$. The effective mass of σ is $g(\varphi\cos\alpha + \chi\sin\alpha)^2 - 2\mu\sigma_0^2$ and the instability point is

$$\varphi_{\mathrm{end}} = \frac{\pm\sqrt{2\mu/g}\sigma_0 - \chi\sin\alpha}{\cos\alpha}\ . \tag{366}$$

In this model the transfer of non-Gaussianity proceeds as follows. For $\varphi > \varphi_{\mathrm{end}}$ the inflaton field φ and the scalar field χ evolve independently. The scalar field φ drives inflation, meanwhile χ develops non-Gaussianities. Notice in particular that χ is almost constant, $\dot{\chi} \approx 0$. From Eqs. (349) and (350) this means that the isocurvature perturbations are just given by the fluctuations of the subdominant scalar field χ, while the fluctuations of the inflaton are curvature perturbations. In Eq. (362) only the term $\widehat{T}^3 c_\beta^3\langle Q_{\chi 1}Q_{\chi 2}Q_{\chi 3}\rangle$ survives. However at this stage there is no correlation between the two perturbations, the mixing angle β defined in Eqs. (351) being constant and equal to zero. In fact since the transfer function which measures the degree of correlation $T_{\mathcal{RS}}$ is proportional to $\dot{\beta}/H$, in Eq. (362) the transfer of non-Gaussianites from χ to φ is suppressed. However, because of the coupling between φ and χ through their interactions to σ, near the critical point φ_{end} the trajectories in the field space (φ,χ) start bending, i.e. $\dot{\beta} \neq 0$. Now in Eq. (362) the transfer of non-Gaussianities from χ to the inflaton φ is acting since $\widehat{T}^3 c_\beta^3\langle Q_{\chi 1}Q_{\chi 2}Q_{\chi 3}\rangle$ is non-vanishing and $c_\beta \neq 0$. In order to give an estimate of the non-Gaussianities, we recall that they depend on the coupling λ and on the number of e-folds N_k, according to Eq. (96). In fact to calculate the bispectra in Eq. (362) one should first recast the scalar field fluctuations in terms of a combination of mass eigenstates $\delta\Psi_i$ with mixing angles θ as defined in Eq. (346). This is due to the fact that the Q_χ fluctuations correspond to interaction eigenstates, and not to mass eigenstate [30]. In this way one can actually borrow the expression in Eqs. (90)–(92). However notice that the term $\widehat{T}^3 c_\beta^3\langle Q_{\chi 1}Q_{\chi 2}Q_{\chi 3}\rangle$ will

not be given only by expression (90)–(92), but it will contain also contributions that might arise in the dynamics after the end of inflation, e.g. depending on the details of the reheating stage. Apart from that, and most importantly, a precise determination of the level of non-Gaussianity produced by the transfer mechanism should require a full second-order analysis of the perturbation evolution. Therefore in the transfer function $\widehat{T}^3$ we are hiding the effective non-Gaussianity which will follow from such an analysis.

As noticed in Ref. [45] a general condition for this mechanism to work is that the terms responsible for the bending of the trajectories in field space must be different from the non-linear coupling term, otherwise there are not enough e-foldings for non-linearities to develop; moreover a kind of attractor trajectory is established which tends to suppress the χ fluctuations and the transfer of non-Gaussianities. In fact this situation corresponds to the first example we gave. In Refs. [44,45] an analytical expression for the one-point probability distribution function (PDF) for the isocurvature mode was also obtained. It reads [44,45]

$$P(\chi)d\chi = \sqrt{\frac{1}{2\pi}\left|\frac{1-\chi^2 v_3}{(1+\chi^2 v_3/3)^3}\right|}\exp\left[-\frac{3\chi^2}{(6+2\chi^2 v_3)\sigma_\chi^2}\right] . \tag{367}$$

As it can be seen from Eq. (367) the PDF depends only on two parameters, the variance of χ (which is proportional to H during inflation) and the parameter v_3 which quantifies the amount of non-Gaussianity. Notice that this parameter just corresponds to the quantity in Eq. (92) for a quartic potential. In Refs. [44,45] the time evolution of the PDF was studied and compared with numerical results finding good agreement with the expression above. Finally, notice from Eq. (360) that the total curvature perturbation will be given by a Gaussian perturbation $\mathscr{R}_*$, which mainly corresponds to the curvature perturbations in the inflaton field φ, and by the non-Gaussian curvature perturbation $T_{\mathscr{R}\mathscr{S}}\mathscr{S}_*$, induced by the isocurvature mode of the scalar field χ.

9.1.4. Some estimates of the non-linearity parameter $f_{\rm NL}^{\phi}$

We can now give a general order-of-magnitude estimate of the non-linearity parameter $f_{\rm NL}^{\phi}$, in those inflationary models where the mixing between adiabatic and entropy perturbations is operative, as follows:

$$f_{\rm NL}^{\phi} \simeq \frac{\langle\phi^3\rangle}{6\langle\phi^2\rangle^2} . \tag{368}$$

To estimate the quantity on the R.H.S. of the last equation we look at the dominant contribution given in Eq. (361). In fact we can extend Eq. (360) to second order in the perturbations and parametrize the transfer of the entropy perturbations to the adiabatic mode $\mathscr{R}^{(2)}$ as[31]

$$\mathscr{R}^{(2)} = \mathscr{R}_*^{(2)} + T_{\mathscr{R}\mathscr{S}}^{(2)}\mathscr{S}_*^{(2)} . \tag{369}$$

[31] By merely extending the first-order expression (360) to second order in the perturbations, corrections of the form (first-order)^2 are not considered. However, we can account for them including them in the transfer functions we introduce. Moreover, these contributions are expected to give corrections of order unity to the overall non-linearity parameter, which on the other hand turns out to be generically much larger than unity, for the scenario we are considering.

If we express the second-order entropy perturbation during inflation as $S_*^{(2)} = g_* \delta^{(2)} s_*$ from Eq. (356), then we can write

$$T^{(2)}_{\mathcal{RS}} \mathcal{S}^{(2)}_* = g_* T^{(2)}_{\mathcal{RS}} \delta^{(2)} s_* = g_* \widehat{T}^{(2)}_{\mathcal{RS}} v_3(\delta s)(\delta^{(1)} s)^2 \ , \tag{370}$$

where $g_* = (H/\dot{A})_*$ and $v_3(\delta s)$ is the non-linearity parameter characterizing the non-Gaussianities in the entropy field δs, in a similar way to what discussed in Eqs. (96) and (92). Here $\widehat{T}^{(2)}_{\mathcal{RS}}$ should account for the transfer of these non-Gaussianities, additional second-order effects from the gravitational dynamics after inflation, and it should also account for second-order corrections of the form (first order)^2. The latter contributions are expected to be $\mathcal{O}(1)$. On the other hand using Eq. (360) we find

$$\langle \phi^2 \rangle = \frac{4}{9} g_*^2 [1 + (T^{(1)}_{\mathcal{RS}})^2] \langle (\delta^{(1)} s_*)^2 \rangle \ , \tag{371}$$

where we have used the fact that $\langle (\delta^{(1)} s_*)^2 \rangle \simeq \langle (Q^{(1)}_{A*})^2 \rangle$ and the approximation that the adiabatic and entropy fields are uncorrelated when they approach the horizon during inflation [289].

Thus, combining Eqs. (370) and (371) we find

$$f^{\phi}_{\rm NL} \simeq \frac{3}{2} \frac{(T^{(1)}_{\mathcal{RS}})^2 \widehat{T}^{(2)}_{\mathcal{RS}}}{[1 + (T^{(1)}_{\mathcal{RS}})^2]^2} g_*^{-1} v_3(\delta s) \ . \tag{372}$$

In fact the expression in Eq. (372) shows how the non-linearities in the entropy field are acquired by the gravitational potential.

In order to give an order-of-magnitude estimate of $f^{\phi}_{\rm NL}$ we notice that it is possible to introduce a dimensionless measure of the correlation in terms of the correlation angle [289]

$$\cos \Delta \equiv \frac{\mathcal{C}_{\mathcal{RS}}}{\mathcal{P}^{1/2}_{\mathcal{R}} \mathcal{P}^{1/2}_{\mathcal{S}}} \ , \tag{373}$$

where $\mathcal{C}_{\mathcal{RS}} = T^{(1)}_{\mathcal{R}} T^{(1)}_{\mathcal{RS}} \mathcal{P}_{\mathcal{R}*}$ is the cross-correlation between adiabatic and entropy perturbations defined as in Eq. (353), and $\mathcal{P}_{\mathcal{R}} = [1 + (T^{(1)}_{\mathcal{RS}})^2] \mathcal{P}_{\mathcal{R}*}$, $\mathcal{P}_{\mathcal{S}} = (T^{(1)}_{\mathcal{SS}})^2 \mathcal{P}_{\mathcal{S}*}$ are the linear adiabatic and entropy power-spectra [289]. Such a correlation angle is indeed a measurable quantity [12,289]. Moreover it is possible to argue on general grounds that the transfer $\widehat{T}^{(2)}_{\mathcal{RS}}$ function can be written as $\widehat{T}^{(2)}_{\mathcal{RS}} \simeq \alpha' N_k (T^{(1)}_{\mathcal{SS}})^2$, where α' is a model-dependent coefficient. Then the non-linearity parameter in Eq. (372) reads

$$f^{\phi}_{\rm NL} \simeq \frac{3}{2} \frac{\mathcal{P}_{\mathcal{S}}}{\mathcal{P}_{\mathcal{R}}} \cos^2 \Delta \alpha' N_k g_*^{-1} v_3(\delta s) \ . \tag{374}$$

For $v_3(\delta s)$ we can use expression (92) if we completely identify the entropy field with the scalar field $\delta\chi$ with cubic self-interactions. Therefore we find

$$g_*^{-1} v_3(\delta s) = \sqrt{2} \sqrt{\epsilon} M_{\rm P} \frac{\lambda N_k}{3H^2} \ , \tag{375}$$

where we have used the fact that $(\dot{A}/H)^2 = 2\epsilon M_{\rm P}^2$. In order for the generation and the transfer of non-Gaussianities to be effective, some constraints are to be satisfied. One of these requirements is that the

effective mass of the entropy field is sufficiently small ($< H$ during inflation), to generate the primordial entropy perturbations. We now turn back to the toy model defined by the potential in Eq. (363) and, under some approximations, we specifically impose such constraints. First of all notice that we are able to recover an expression similar to Eqs. (374)–(375) starting from the second-order curvature perturbation

$$\zeta_{\varphi}^{(2)} \sim \frac{H}{\dot{\varphi}}\,\delta^{(2)}\varphi \sim \frac{H}{\dot{\varphi}}\,\lambda\,\frac{m_{\times}^2}{H^2}\,\frac{N_k^2}{3H^2}(\delta^{(1)}\chi)^2\ , \tag{376}$$

where we have used Eq. (364) in the case of cubic self-interactions for the scalar field χ. If we now say that $\delta^{(1)}\chi \sim \delta^{(1)}\varphi$ then $\zeta_{\varphi}^{(2)} \sim (\dot{\varphi}/H)(\lambda N_k^2 m_{\times}^2/H^4)(\zeta_{\varphi}^{(1)})^2$, from which we read the non-linearity parameter (up to factors of order unity)[32]

$$f_{\rm NL}^{\phi} \sim \frac{\dot{\varphi}}{H}\,\lambda\,\frac{N_k^2}{3H^2}\,\frac{m_{\times}^2}{H^2} \sim \sqrt{2}\sqrt{\epsilon}M_{\rm P}\lambda\frac{N_k^2}{3H^2}\,\frac{m_{\times}^2}{H^2}\ . \tag{377}$$

We now relate the third derivative of the potential with respect to the scalar field χ $V_{\chi\chi\chi}=\lambda$ to its effective mass $V_{\chi\chi}$ as $V_{\chi\chi}=V_{\chi\chi\chi}\chi$ so that

$$f_{\rm NL}^{\phi} \sim \sqrt{2}\sqrt{\epsilon}\frac{M_{\rm P}}{H}\,\frac{V_{\chi\chi}}{3H\chi}\,N_k^2\,\frac{m_{\times}^2}{H^2}\ . \tag{378}$$

Notice that we expect $(\sqrt{2}\sqrt{\epsilon}M_{\rm P}/H)^{-1}$ to be at most of the order of the amplitude of the produced density perturbations $2\pi\mathscr{P}_{\mathscr{R}}^{1/2} \sim 2.5\times 10^{-4}$ [174]. Therefore we conclude that

$$f_{\rm NL}^{\phi} \sim 4\times 10^3\,\frac{V_{\chi\chi}}{3H^2}\,\frac{H}{\chi}\,N_k^2\,\frac{m_{\times}^2}{H^2}\ . \tag{379}$$

If we require that the scalar field χ is light enough, $V_{\chi\chi}\ll H^2$, to acquire some fluctuations $\delta\chi$, then we can estimate $4\times 10^3(V_{\chi\chi}/3H^2) \sim 1$. This is actually a conservative estimate. In fact the mass of the scalar field $m_{\chi}^2=V_{\chi\chi}$ can be also of the order of the Hubble rate during inflation. The scalar field χ can be subdominant with respect to the (inflaton) field φ and thus the stringent slow-roll conditions are widely relaxed.[33] It is also worth to notice that in the inhomogeneous reheating scenario, even if the scalar field(s) χ determining the inflaton decay rate do not need to satisfy slow-roll conditions, the non-Gaussianities induced by the self-interactions of χ are constrained from the requirement that $2/3(V_{\chi\chi}/H^2) \sim 10^{-2}$ [300], to satisfy the observational limits on the spectral index of density perturbations which can be traced back directly to the fluctuations $\delta\chi$ (see Eq. (310)). This constraint does not apply to the case described in this subsection,

[32] We would like to warn the reader about some technical issues which appear in deriving the formula in Eq. (377). It coincides with Eqs. (374)–(375) in the limit where during inflation $(\mathscr{P}_{\mathscr{S}}/\mathscr{P}_{\mathscr{R}})\cos^2\Delta \approx 1$ (the factor $m_{\times}^2/H^2$, being contained in the α' coefficient for this particular model). In fact Eq. (377) has been derived under the approximation that the transfer of the entropy perturbation is extremely efficient and that the final curvature perturbation is essentially determined by such a transfer. This is of course a limiting case. Related to that, notice that in using Eq. (364) we have assumed that the transfer is operative during all the N_k e-folds from the time the mode leaves the horizon till the end of inflation. Actually the transfer could be efficient only for a shorter period of time. Therefore in the α' coefficient we should account for a fraction of the N_k e-folds. However, as we explain later, our estimate of the non-linearity parameter $f_{\rm NL}^{\phi}$ is, for other reasons, quite a conservative one.

[33] In the case of single-field slow-roll inflation the self-interactions of the inflaton field φ produce a non-linearity parameter $f_{\rm NL} \sim \mathscr{P}_{\mathscr{R}}^{-1/2}(V_{\varphi\varphi\varphi}/H)N_k$. The slow-roll condition on the third derivative of the inflaton potential imposes $\mathscr{P}_{\mathscr{R}}^{-1/2}(V_{\varphi\varphi\varphi}/H)\ll 1$.

since the fluctuations of the additional scalar field χ may give only a subdominant contribution to the total amplitude of the density perturbations. Another condition that we impose onthe model is that $\dot{\chi} < \dot{\phi}$ during inflation, in order to allow for non-Gaussianities in the χ sector to develop. This condition implies that $N_k < H^2/m_\chi^2$. Therefore, with our conservative estimate from Eq. (379) we find that

$$f_{\rm NL}^{\phi} \lesssim \frac{H}{\chi} N_k \sim 60\,\frac{H}{\chi}\ . \tag{380}$$

We see that if the χ field is not much larger than H then significant non-Gaussianities can be produced very close to the limits set by WMAP.

We conclude this section by introducing a simple parametrization for the non-Gaussianity generated in the gravitational potential by the transfer mechanism. Such a parametrization is needed for practical purposes when confronting with observations. Our results for the non-linearity parameter of the gravitational potential indicate that we expect the gravitational potential to be of the form

$$\phi = \phi_1 + f_{\rm NL}^{\phi}(\phi_2^2 - \langle\phi_2^2\rangle) + \mathcal{O}(f_{\rm NL}^{\phi 2})\ , \tag{381}$$

as it derives from Eqs. (360), (369) and (370). Here ϕ_1 and ϕ_2 are zero-mean Gaussian fields with non-vanishing cross-correlation $\langle\phi_1\phi_2\rangle \neq 0$, with the field ϕ_2 corresponding to that part of the gravitational potential induced by the evolution of the entropy perturbations. Such a parametrization has been introduced in Refs. [83,30] and also envisaged in Ref. [139] as a possible extension of the formula (296), to look for specific non-Gaussian signatures from two-field inflationary models. If an isocurvature perturbation mode survives after inflation it is reasonable to parametrize its non-linearities as

$$\mathcal{S} = \mathcal{S}^{(1)} + f_{\rm NL}^{\rm iso}(\mathcal{S}^{(1)2} - \langle\mathcal{S}^{(1)2}\rangle)\ . \tag{382}$$

Therefore, one can use also this parametrization to search for non-Gaussianities in primordial (correlated) adiabatic perturbations ϕ and isocurvature perturbations $\mathcal{S}$, as explained in detail in Ref. [144] (see also Ref. [30] for more generic cases).

9.2. Non-Gaussianity in unconventional inflation set-ups

While inflation driven by a scalar field with a very flat potential provides an early de Sitter phase of the Universe and elegantly solves the horizon and flatness problems, it is certainly worthwhile to look for alternatives whose predictions might be discriminated with present and future observations.Since a de Sitter phase of expansion obliterates the horizon and flatness problems, any alternative to the standard slow-roll inflation has to preserve this property. It is legitimate to ask whether the slow-roll picture is really necessary or if slow-roll can be obtained in some unconventional way. In the following, we review some possibilities which may predict a large amount of non-Gaussianity.

9.2.1. Warm inflation

Warm inflation [40,41] is an alternative to the standard scenario of supercooled inflation, where dissipative effects are assumed to play a dynamical role during inflation so that radiation production occurs simultaneously with the inflationary expansion. The warm inflation picture is a comprehensive set of possible interactions between fields during inflation; no a priori assumptions about multi-field interactions, thus particle production, during the inflationary epoch are made. As such, the warm inflation picture

makes explicit that the thermodynamic state of the Universe during inflation is a dynamical question. Supercooled inflation then emerges as the limiting case in which interactions are negligible.

The evolution of the (minimally coupled) inflaton in warm inflation is described by the phenomonological equation

$$\ddot{\varphi} + (3H + \Gamma)\dot{\varphi} - \frac{\nabla^2\varphi}{a^2} + V'(\varphi) = 0 \ , \tag{383}$$

where the dissipation rate Γ may generally depend upon φ.

The presence of radiation during inflation influences the seeds of density perturbations. It is therefore natural to ask whether in such a model the level of non-Gaussianity in the primordial perturbations might be sensibly different from that of the standard slow-roll scenario. This problem was analyzed in Ref. [99] (see also Ref. [98]), where the bispectrum of the gravitational field fluctuations was calculated through a simple generalization of the stochastic approach adopted in Ref. [86]. In analyzing the dynamics of inflaton fluctuations, metric fluctuations were disregarded for simplicity, as in Ref. [82], so that the bispectrum is non-zero only due to the presence of inflaton self-interactions. Requiring the slow-roll condition $|\ddot{\varphi}| \ll (3H + \Gamma)|\dot{\varphi}|$ and imposing a near-thermal-equilibrium, Markovian approximation, the equation of motion for the inflaton field emerges as

$$\frac{\mathrm{d}\varphi(\mathbf{x}, t)}{\mathrm{d}t} = \frac{1}{\Gamma}\left[\mathrm{e}^{-2Ht}\nabla^2\varphi(\mathbf{x}, t) - V'(\varphi(\mathbf{x}, t)) + \eta(\mathbf{x}, t)\right] \ . \tag{384}$$

Implementing the *fluctuation–dissipation* theorem determines the properties of the noise, which read

$$\langle\eta\rangle = 0 \ , \tag{385}$$

$$\langle\eta(\mathbf{k}, t)\eta(\mathbf{k}', t')\rangle = 2\Gamma T(2\pi)^3\delta^{(3)}(\mathbf{k} - \mathbf{k}')\delta(t - t') \ , \tag{386}$$

where T is the temperature and $\mathbf{k}$ and $\mathbf{k}'$ denote physical momenta.

By splitting as usual the inflaton field into a homogeneous background $\varphi_0(t)$ and a fluctuation field $\delta\varphi(\mathbf{x}, t)$, one can expand the equation of motion in powers of the fluctuations around the background. The bispectrum is then immediately obtained from the second-order contribution. The corresponding non-linearity strength f_{NL}^{ϕ} in the strong-dissipative regime, $\Gamma/H \gg 1$, is found [99]

$$f_{\mathrm{NL}}^{\phi} = \frac{5}{6}\left(\frac{\dot{\varphi}_0}{H^2}\right)\left[\ln\left(\frac{\Gamma}{H}\right)\frac{V'''}{\Gamma}\right] \ , \tag{387}$$

Applying this formalism to the $\lambda\phi^4$ model, and imposing that the amplitude of density fluctuations matches the COBE normalization one finds $f_{\mathrm{NL}}^{\phi} \approx 3.7 \times 10^{-2}$. A similar analysis in the weak-dissipative limit, $\Gamma/H \ll 1$, leads to a value of f_{NL}^{ϕ} smaller by an order of magnitude [98].

9.2.2. Ghost inflation

A new possibility of having a de Sitter phase in the Universe in a way differing from a cosmological constant has been proposed in Ref. [15]. It can be thought of as arising from a derivatively coupled "ghost" scalar field φ which "condenses" in a background where it has non-zero velocity

$$\langle\dot{\varphi}\rangle = M^2 \to \langle\varphi\rangle = M^2 t \ . \tag{388}$$

Unlike other scalar fields, the velocity $\dot{\varphi}$ does *not* redshift to zero as the Universe expands, it stays constant, and indeed the energy momentum tensor is identical to that of a cosmological constant. However, the ghost condensate is *not* a cosmological constant, it is a physical fluid with a physical fluctuation π defined as

$$\varphi = M^2 t + \pi \ . \tag{389}$$

The ghost condensate then gives an alternative way of realizing a de Sitter phase in the Universe. Furthermore, it can be shown that the symmetries of the theory allow to construct a systematic and reliable effective Lagrangian for π and gravity at energies lower than the ghost cut-off M. Neglecting the interactions with gravity, the effective Lagrangian for π (around flat space) has the form [14]

$$S = \int \mathrm{d}^4x \left(\frac{1}{2}\dot{\pi}^2 - \frac{\alpha^2}{2M^2}(\nabla^2\pi)^2 - \frac{\beta}{2M^2}\dot{\pi}(\nabla\pi)^2 + \cdots \right) \ , \tag{390}$$

where α and β are order one coefficients. The Lagrangian is non-Lorentz invariant, as it should be expected, since the background $\dot{\varphi} = M^2$ breaks Lorentz invariance spontaneously (the π field can be thought as the Goldstone boson for this symmetry breaking). The low-energy dispersion relation for π is of the unusual form

$$\omega^2 = \alpha^2 \frac{k^4}{M^2} \ . \tag{391}$$

The main motivation for such an approach is that coupling this sector to gravity leads to a variety of interesting modifications of gravity in the infrared, including antigravity and oscillatory modulation of the Newtonian potential at late times and large distances [14].

As pointed out in Ref. [15], two are the important differences here from ordinary slow-roll inflation. First, there is no slow-roll. Even in the approximation where the potential is exactly flat, $\dot{\varphi} = M^2$ is non-zero. The second important difference with standard slow-roll inflation concerns the size of the fluctuations in φ (or equivalently π). Since the effective Lagrangian for π is non-relativistic, and in particular there are no k^2 spatial kinetic terms, the fluctuations of π are less suppressed than in a relativistic theory. In a relativistic theory, the fact that scalar fields have scaling dimension one tells that the size of the fluctuations of a scalar field inside a region of size R is given by $\sim 1/R$, similarly, at a frequency E it is given by E. In ordinary inflation, the inflaton fluctuations freeze when they have a typical energy $E \sim H$, so that their typical size is $\delta\varphi \sim H$. We can determine the size of the fluctuations in the ghost inflation case by a simple scaling argument familiar from power-counting for non-relativistic effective theories. Suppose one scales energies by a factor of s, $E \to sE$, or alternatively $t \to s^{-1}t$. Clearly, because of the $\omega^2 \propto k^4$ dispersion relation, one has to scale k differently, $k \to s^{1/2}k$ or $x \to s^{-1/2}x$. We then determine the scaling dimension of π by requiring the quadratic action to be invariant, and one finds that π has scaling dimension $\frac{1}{4}$

$$\pi \to s^{1/4}\pi \ . \tag{392}$$

Now, π has *mass* dimension one, so the fluctuations at frequencies of order the cutoff M is $\delta\pi_M \sim M$. But the fact that π has scaling dimension $\frac{1}{4}$ tells that the fluctuation at a lower energy E is $\delta\pi_E \sim (EM^3)^{1/4}$. In particular, the size of ghost fluctuations that freeze, as usual, by Hubble friction when its frequency is of order $E \sim H$, is

$$\delta\pi_H \sim (HM^3)^{1/4} \ . \tag{393}$$

Of course, for consistency of the effective theory, one has to require $H \ll M$. The field φ fluctuates and is stretched out until the Hubble damping becomes important at frequency $E \sim H$; this does *not* correspond to $k \sim H$ but rather, from the dispersion relation, $k \sim \sqrt{HM}$. The fluctuation $\delta\pi_H$ causes inflation to end at slightly different times in different places, and so one has the estimate

$$\frac{\delta\rho}{\rho} \sim H\delta t = \frac{H\delta\pi_H}{\dot{\varphi}} \ . \tag{394}$$

Notice that $\dot{\varphi} = M^2$, has nothing to do with slow-roll parameters. Furthermore, $\delta\pi_H$ is much larger. One then finds

$$\frac{\delta\rho}{\rho} \sim \left(\frac{H}{M}\right)^{5/4} \ , \tag{395}$$

which can be compared with the standard inflationary case

$$\frac{\delta\rho}{\rho} \sim \frac{H}{M_{\rm P}\sqrt{\epsilon}} \ . \tag{396}$$

A detailed computation confirms expectation (395) and gives [15] the primordial curvature spectrum

$$P_{\mathscr{R}}^{1/2} = \frac{1}{\sqrt{\pi}\Gamma(1/4)} \frac{(H^5M^3\alpha^{-3})^{1/4}}{\dot{\varphi}} = \frac{1}{\sqrt{\pi}\Gamma(1/4)} \left(\frac{H}{M}\right)^{5/4} \alpha^{-3/4} \ . \tag{397}$$

As for the non-Gaussianity since the mass scales are all much smaller than the Planck scale, there is no effect coming from gravitational interactions. The dominant effect comes from the trilinear interaction in the π effective Lagrangian

$$\frac{\beta}{M^2}\dot{\pi}(\nabla\pi)^2 \ . \tag{398}$$

This leads to a non-zero three-point function for the density perturbations. To get an idea of the dimensionless size of this effect, one needs to find the dimensionless size of this coupling at an energy of order H. Since one knows the scaling dimension of t, x and π, the coefficient of the operator (398) has scaling dimension $\frac{1}{4}$. The dimensionless size of this interaction at the cutoff M is ~ 1; scaling it down to an energy of order H one gets an estimate for the non-Gaussianity of the perturbations [15]

$$\mathrm{NG} \sim \left(\frac{H}{M}\right)^{1/4} \sim \left(\frac{\delta\rho}{\rho}\right)^{1/5} \ . \tag{399}$$

These are much larger than in standard inflation, where one expects $\mathrm{NG} \sim \epsilon \cdot (\delta\rho/\rho)$ during inflation. A detailed calculation of the level of non-Gaussianity has been performed in Ref. [15] (see also Ref. [262]) and shows that $f_{\rm NL}$ has a non-trivial momentum dependence. However, for an equilateral configuration, defined by setting $k_1 = k_2 = k_3$, one can define an "effective" $f_{\rm NL}$

$$f_{\rm NL}^{\rm eff} \simeq -140 \cdot \beta \cdot \alpha^{-8/5} \ . \tag{400}$$

This value can be much larger than unity depending upon the parameters α and β. However, if the resulting $f_{\rm NL}^{\rm eff}$ turns out to be of order unity, one expects relevant corrections coming from the post–de Sitter phase.

9.2.3. "D-cceleration" mechanism of inflation

In Ref. [264] an unconventional mechanism for slow-roll inflation was introduced, motivated by the behavior of rolling scalar fields in strongly interacting theories, analyzed using the AdS/CFT correspondence [181]. The key feature of this model is that the inflaton field φ is naturally slowed as it approaches a point where many light degrees of freedom χ emerge, the slow-down arising from the virtual effects of the light particles. From the stringy perspective, this scenario translates to a probe $D3$-brane travelling down a five-dimensional warped throat geometry. The ultraviolet end of the throat joins smoothly onto a compactification in the manner of Randall–Sundrum [235], ensuring that gravity in four dimensions is dynamical while coupling the field theory to sectors in the compactification which can generate corrections to the effective action for φ. These corrections generically produce a non-trivial potential energy for φ including a mass term $m^2\varphi^2$ and a corresponding closing up of the throat in the infrared region at a scale φ_{IR}.

The dynamics of the probe $D3$-brane is captured by the Dirac–Born–Infeld (DBI) action coupled to gravity,

$$S = \int \mathrm{d}^4x\sqrt{-g}\left(\frac{1}{2}M_{\mathrm{P}}^2\mathscr{R} + L_{\mathrm{eff}} + \cdots\right) \tag{401}$$

with

$$L_{\mathrm{eff}} = -\frac{1}{g_s}\left(f(\varphi)^{-1}\sqrt{1 + f(\varphi)g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi} + V(\varphi)\right) . \tag{402}$$

Here g_s is the string coupling constant, the function $f(\varphi)$ is the (squared) warp factor of the AdS-like throat. For example, for a pure AdS_5 of radius R, it is simply $f(\varphi) = \lambda/\varphi^4$ with $\lambda \equiv R^4/\alpha'^2$ and α' is the string tension. The non-analytic behavior of the square-root in (402) gives rise to a speed limit restricting how fast the scalar field may roll. When $f(\varphi) = \lambda/\varphi^4$, the speed limit is given by $|\dot{\varphi}| \leqslant \varphi^2/\sqrt{\lambda}$ and a useful measure of how close we are to the limit is given by

$$\gamma = \frac{1}{\sqrt{1 - f(\varphi)\dot{\varphi}^2}} \tag{403}$$

which is analogous to the Lorentz contraction factor defined in special relativity and grows without bound as the speed limit is approached.

The late time behavior of the scale-factor a and the scalar field φ were determined in Ref. [264] for a variety of potentials $V(\varphi)$ and for $f(\varphi) = \lambda/\varphi^4$ using the Hamilton-Jacobi approach. This method elevates the scalar field φ to the role of cosmological time, so that the Hubble parameter $H = \dot{a}/a$ is considered as a function $H = H(\varphi)$ determined in terms of the potential by

$$V(\varphi) = 3g_sM_{\mathrm{P}}^2H(\varphi)^2 - \gamma(\varphi)/f(\varphi) , \tag{404}$$

where $\gamma(\varphi)$ is given by

$$\gamma(\varphi) = (1 + 4g_s^2M_{\mathrm{P}}^4f(\varphi)H'(\varphi)^2)^{1/2} . \tag{405}$$

The evolution of $\varphi(t)$ is then fixed by the first-order Friedmann equation,

$$\dot{\varphi} = -2g_sM_{\mathrm{P}}^2\frac{H'(\varphi)}{\gamma(\varphi)} . \tag{406}$$

As in the standard inflationary scenarios, it is useful to introduce a slow-roll parameter

$$\epsilon = \frac{2g_s M_{\rm P}^2}{\gamma}\left(\frac{H'}{H}\right)^2 \tag{407}$$

which parameterises the deviation from a pure de Sitter phase. In particular one has $\ddot{a}/a = H^2(1-\epsilon)$.

In Ref. [264] the first-order equations (404) and (406) were studied for a variety of potentials $V(\varphi)$. For the massive scalar field case, $V(\varphi) \sim m^2\varphi^2$, the late time dynamics was shown to be a power-law inflation given by $a(t) \to a_0 t^{1/\epsilon}$ and

$$\varphi \to \frac{\sqrt{\lambda}}{t}, \quad \gamma \to \sqrt{\frac{4g_s}{3\lambda}} M_{\rm P} m t^2, \quad H \to \frac{1}{\epsilon t}\ . \tag{408}$$

The coefficient of the (time-dependent) Hubble parameter is given by the slow-roll parameter which is a constant when evaluated on this background,

$$\frac{1}{\epsilon} = \frac{1}{3}\left(1 + \sqrt{1 + \frac{3m^2\lambda}{g_s M_{\rm P}^2}}\right) \approx \sqrt{\frac{\lambda}{3g_s}}\,\frac{m}{M_{\rm P}}\ . \tag{409}$$

For $\epsilon < 1$, one obtains a phase of power-law inflation. The exponential de Sitter phase can be thought of as the limit as $\epsilon \to 0$. Note that in contrast to usual single field slow-roll inflation, the accelerated expansion occurs only if the mass of inflaton m is suitably large. This is one novel aspect of this model.

It is clear from action (402) that expanding φ in fluctuations $\varphi \to \varphi + \delta\varphi$ involves expanding the square root in (402). This produces powers of γ accompanying the powers of the fluctuation $\delta\varphi$. The origin of the strong non-Gaussianities can be therefore understood rather simply. The Lagrangian is proportional to $\sqrt{1-v_p^2}$, where $v_p = \sqrt{\lambda}\dot{\varphi}/\varphi^2$ is the proper velocity of the brane probe whose position collective coordinate is the inflaton φ. The inflationary solution involves a proper velocity approaching the speed of light as φ approaches the origin (with both φ and $\dot{\varphi}$ decreasing towards zero). Expanding the action in fluctuations of φ involves expanding the square root in the Lagrangian, which produces powers of $\gamma = 1/\sqrt{1-v_p^2}$ accompanying powers of the fluctuations of the inflaton. Since γ is relatively large, this produces a large contribution to non-Gaussianities in the model. It is also easy to understand why linear perturbations with momentum k freeze when they cross the sound horizon at $aH = k/\gamma$ rather than at $aH = k$. The resulting power-spectrum reads [9] $\mathscr{P}_k^{\zeta} \sim g_s H^4/\dot{\varphi}^2 \sim g_s/\epsilon^4\lambda$, where we have used the inflationary background(408); and the string coupling g_s enters because the canonically normalized scalar field is $\varphi/\sqrt{g_s}$.

In the large γ, small ϵ regime, the leading terms in the interaction Hamiltonian are given by

$$\mathscr{H}_{\rm int} = -\frac{a^3\gamma^5}{g_s}\left[\frac{1}{2\dot{\varphi}}(\delta\dot{\varphi})^3 + \frac{k^2}{a^2\gamma^2}\frac{1}{2\dot{\varphi}}\,\delta\dot{\varphi}(\delta\varphi)^2\right]\ . \tag{410}$$

The computation of the three-point correlation function goes along the lines described in Section 2.5 and gives an $f_{\rm NL}$ with a non-trivial momentum dependence [9]. However, for an equilateral configuration defined by setting for $k_1 = k_2 = k_3$ one can define an "effective" $f_{\rm NL}$

$$f_{\rm NL}^{\rm eff} \simeq -10^{-1}\gamma^2\ , \tag{411}$$

which can be rather large. Again, if $\gamma \sim 20$ and $f_{\rm NL}^{\rm eff}$ is of order unity, one needs a detailed study of the post-inflationary evolution of the non-Gaussianities.

In Table 1 we summarize the level of primordial non-Gaussianity as generated in different cosmological scenarios. Actually for the first three cases the values displayed have been obtained by a complete study of the perturbation evolution from an early inflationary phase, through reheating till the radiation and the matter dominated epoch (to which the results are referred). For the other scenarios the values of $f_{\rm NL}$ actually do not refer to the non-Gaussianities generated in the CMB temperature anisotropies, rather to non-Gaussianities in the curvature perturbation ζ (and for the unconventional inflationary scenarios they refer only to the inflationary phase). In this case when the values shown are $f_{\rm NL} \gg 1$, then they should be a good approximation for the actual observable quantity, but when they are of order unity one expects significant corrections from a detailed study of the post-inflationary evolution of the non-Gaussianites. The function $g(\mathbf{k}_1, \mathbf{k}_2)$ is the same in the first three cases and the infrared behavior of $f_{\rm NL}(\mathbf{k}_1, \mathbf{k}_2)$ is automatically regularized once the monopole term is subtracted by requiring that $\langle \Delta T/T \rangle = 0$. In the curvaton scenario the parameter r corresponds to the ratio of the curvaton energy density to the total energy density at the epoch of the curvaton decay. In the inhomogeneous reheating scenario we have accounted for a fraction of the total decay rate to be dependent on a light scalar field, $\Gamma_1/\bar{\Gamma}$, and for a ratio $\bar{\Gamma}/H$ at the end of inflation not much smaller than one, allowing the (positive) α parameter to be $\alpha < 1/6$, as discussed in Section 8.3. The "minimal case" corresponds to an inflaton decay rate Γ that is fully controlled by a light scalar field χ, $\Gamma \propto \chi^2$, and such that $\Gamma \ll H$ during inflation ($\alpha = 1/6$). In the multiple-field case $V_{\chi\chi}/H^2$ represents the ratio of the mass of the additional self-interacting light scalar field χ and the Hubble parameter during inflation. The value shown is an estimate obtained from extending some linear relations to second order in the perturbations, and some corrections are expected when the estimate shown is of the order one. Here we must make a remark. Notice that the resulting $f_{\rm NL}$ depends on the "isocurvature fraction" $\mathcal{P}_{\mathcal{S}}/\mathcal{P}_{\mathcal{R}}$, determining the relative amplitude of the isocurvature perturbation to the adiabatic one, and on the correlation between them. Actually such a formula is valid not only if after inflation an isocurvature perturbation survives, but also if no entropy mode is left over. In the latter case one should interpret the combination $(\mathcal{P}_{\mathcal{S}}/\mathcal{P}_{\mathcal{R}})\cos^2\Delta$ evaluated at the end of inflation. If an entropy mode survives for the data analysis one must also consider a corresponding non-linearity parameter $f_{\rm NL}^{\rm iso}$ as shown in Eqs. (382) and (381). In the warm scenario, Γ indicates the inflaton dissipation rate. Finally, notice that in the ghost and D-cceleration scenarios, the values reported refer to an equilateral configuration of the wave numbers, $k_1 = k_2 = k_3$.

10. Observational constraints on non-Gaussianity

In this section, we discuss how to test observationally Gaussianity of primordial fluctuations and constrain the non-linearity parameter, $f_{\rm NL}$. Currently, the best constraint comes from WMAP's measurements of CMB anisotropy [139]; thus, we focus on testing Gaussianity of the CMB. This section is organized as follows: In Section 10.1 we study the statistical properties of the angular n-point harmonic spectra for $n = 2$ (power-spectrum), 3 (bispectrum), and 4 (trispectrum). In Section 10.2 we make theoretical predictions for the CMB angular bispectrum from inflation. In Section 10.3 we calculate the secondary bispectrum contribution from the Sunyaev–Zel'dovich effect and the weak-lensing effect, and foreground contribution from extragalactic radio and infrared astronomical sources. In Section 10.4 we estimate how well one can constrain $f_{\rm NL}$ with observations, and discuss how to distinguish between primordial,

secondary, and foreground bispectra. In Section 10.5 we present a practical way to determine $f_{\rm NL}$ as well as the point-source contribution from nearly full-sky CMB experiments. In Section 10.6 we review the observational constraints on $f_{\rm NL}$ from the WMAP experiment using the method described here.

10.1. Angular n-point harmonic spectrum on the sky

As we have mentioned in the introduction, the angular *n*-point correlation function,

$$\langle f(\hat{\mathbf{n}}_1) f(\hat{\mathbf{n}}_2) \dots f(\hat{\mathbf{n}}_n) \rangle \ , \tag{412}$$

is a simple statistic characterizing a clustering pattern of fluctuations on the sky, $f(\hat{\mathbf{n}})$. Here, the bracket denotes the ensemble average, and Fig. 1 sketches its meaning. If the fluctuation is Gaussian, then the two-point correlation function specifies all the statistical properties of $f(\hat{\mathbf{n}})$, for the two-point correlation function is the only parameter in a Gaussian distribution. If it is not Gaussian, then we need higher-order correlation functions to determine the statistical properties.

Yet simple, one disadvantage of the angular correlation function is that data points of the correlation function at different angular scales are generally not independent of each other, but correlated: the two-point correlation at 1° is correlated with that at 2°, and so on. This property makes a detailed statistical analysis and interpretation of the data rather complicated.

Hence, one finds it more convenient to expand $f(\hat{\mathbf{n}})$ into spherical harmonics, the orthonormal basis on the sphere, as

$$f(\hat{\mathbf{n}}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\hat{\mathbf{n}}) \ , \tag{413}$$

and then to consider the angular *n*-point harmonic spectrum, $\langle a_{l_1 m_1} a_{l_2 m_2} \dots a_{l_n m_n} \rangle$. While a_{lm} for $m \neq 0$ is complex, reality of $f(\hat{\mathbf{n}})$ gives $a_{l-m} = a_{lm}^*(-1)^m$, and thus the number of independent modes is not $4l+1$, but $2l+1$.

In particular, the angular two-, three-, and four-point harmonic spectra are called the angular *power-spectrum*, *bispectrum*, and *trispectrum*, respectively. By isotropy the angular power-spectra at different angular scales, or at different l's, are uncorrelated. Moreover, since the spherical harmonics are orthogonal for different l's, they highlight characteristic structures on the sky at a given l. In other words, even if the angular correlation function is featureless, the angular spectrum may have a distinct structure, for inflation predicts a prominent peak in the angular power-spectrum, not in the angular correlation function. In this section, we study statistical properties of the angular *n*-point harmonic spectra.

10.1.1. Statistical isotropy of the Universe

In reality, we cannot measure the ensemble average of the angular harmonic spectrum, but one realization such as $a_{l_1 m_1} a_{l_2 m_2} \dots a_{l_n m_n}$, which is so noisy that we want to average it somehow to reduce the noise.

We assume *statistical isotropy* of the Universe from which it follows that our sky is isotropic and has no preferred direction. Isotropy of the CMB justifies the assumption. The assumption readily implies that one can average the spectrum over m_i with an appropriate weight, as m_i represent an azimuthal orientation on the sky. The average over m_i enables us to reduce the statistical error of the measured harmonic spectra.

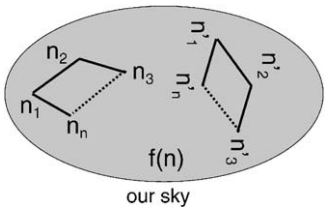

Fig. 4. Statistical isotropy of angular correlation function. A schematic view of statistical isotropy of the angular correlation function. As long as its configuration is preserved, we can average $f(\hat{\mathbf{n}}_1)\dots f(\hat{\mathbf{n}}_n)$ over all possible orientations and positions on the sky.

How can we find the weight? One finds it as a solution to statistical isotropy, or *rotational invariance* of the angular correlation function on the sky,

$$\langle Df(\hat{\mathbf{n}}_1)Df(\hat{\mathbf{n}}_2)\dots Df(\hat{\mathbf{n}}_n)\rangle = \langle f(\hat{\mathbf{n}}_1)f(\hat{\mathbf{n}}_2)\dots f(\hat{\mathbf{n}}_n)\rangle \ , \tag{414}$$

where $D = D(\alpha, \beta, \gamma)$ is a rotation matrix for the Euler angles α, β, and γ. Fig. 4 sketches the meaning of statistical isotropy. Substituting Eq. (413) for $f(\hat{\mathbf{n}})$ in Eq. (414), we then need rotation of the spherical harmonic, $DY_{lm}(\hat{\mathbf{n}})$. It is formally represented by the rotation matrix element, $D^{(l)}_{m'm}(\alpha, \beta, \gamma)$, as [245]

$$DY_{lm}(\hat{\mathbf{n}}) = \sum_{m'=-l}^{l} D^{(l)}_{m'm} Y_{lm'}(\hat{\mathbf{n}}) \ . \tag{415}$$

The matrix element, $D^{(l)}_{m'm} = \langle l, m'|D|l, m\rangle$, describes finite rotation of an initial state whose orbital angular momentum is represented by l and m into a final state represented by l and m'. Finally, we obtain the statistical isotropy condition on the angular n-point harmonic spectrum:

$$\langle a_{l_1m_1}a_{l_2m_2}\dots a_{l_nm_n}\rangle = \sum_{\text{all } m'} \langle a_{l_1m'_1}a_{l_2m'_2}\dots a_{l_nm'_n}\rangle D^{(l_1)}_{m'_1m_1} D^{(l_2)}_{m'_2m_2}\dots D^{(l_n)}_{m'_nm_n} \ . \tag{416}$$

Using this equation, in Ref. [114] appropriate weights for averaging the angular power-spectrum ($n = 2$), bispectrum ($n = 3$), and trispectrum ($n = 4$), over azimuthal angles have been systematically evaluated. Some of those may be found more intuitively; however, this method allows us to find the weight for any higher-order harmonic spectrum. In the following sections, we derive rotationally invariant, azimuthally averaged harmonic spectra for $n = 2$, 3, and 4, and study their statistical properties.

10.1.2. Angular power-spectrum

The angular power-spectrum measures how much fluctuations exist on a given angular scale. For example, the variance of a_{lm} for $l \geqslant 1$, $\langle a_{lm}a^*_{lm}\rangle$, measures the amplitude of fluctuations at a given l.

Generally speaking, the covariance matrix of a_{lm}, $\langle a_{l_1m_1}a^*_{l_2m_2}\rangle$, is not necessarily diagonal. It is, however, actually diagonal once we assume full sky coverage and rotational invariance of the angular two-point correlation function, as we will show in this section. The variance of a_{lm} thus describes the two-point correlation completely.

Rotational invariance (Eq. (416)) requires

$$\langle a_{l_1 m_1} a^*_{l_2 m_2}\rangle = \sum_{m'_1 m'_2} \langle a_{l_1 m'_1} a^*_{l_2 m'_2}\rangle D^{(l_1)}_{m'_1 m_1} D^{(l_2)*}_{m'_2 m_2} \tag{417}$$

to be satisfied, where we have used the complex conjugate for simplifying calculations. From this equation, we seek for a rotationally invariant representation of the angular power-spectrum. Suppose that the covariance matrix of a_{lm} is diagonal, i.e., $\langle a_{l_1 m_1} a^*_{l_2 m_2}\rangle = \langle C_{l_1}\rangle \delta_{l_1 l_2}\delta_{m_1 m_2}$. Eq. (417) then reduces to

$$\langle a_{l_1 m_1} a^*_{l_2 m_2}\rangle = \langle C_{l_1}\rangle \delta_{l_1 l_2} \sum_{m'_1} D^{(l_1)}_{m'_1 m_1} D^{(l_1)*}_{m'_1 m_2} = \langle C_{l_1}\rangle \delta_{l_1 l_2}\delta_{m_1 m_2} \ . \tag{418}$$

Thus, we have proven that $\langle C_l\rangle$ is rotationally invariant. Rotational invariance implies that the covariance matrix is diagonal.

Observationally, the unbiased estimator of $\langle C_l\rangle$ should be

$$\begin{aligned} C_l &= \frac{1}{2l+1} \sum_{m=-l}^{l} a_{lm} a^*_{lm} = \frac{1}{2l+1}\left(a^2_{l0} + 2\sum_{m=1}^{l} a_{lm} a^*_{lm}\right) \\ &= \frac{1}{2l+1}\left\{a^2_{l0} + 2\sum_{m=1}^{l} [(\Re a_{lm})^2 + (\Im a_{lm})^2]\right\} . \end{aligned} \tag{419}$$

The second equality follows from $a_{l-m} = a^*_{lm}(-1)^m$, i.e., $a_{l-m}a^*_{l-m} = a_{lm}a^*_{lm}$, and hence we average $2l+1$ independent samples for a given l. It suggests that fractional statistical error of C_l is reduced by $\sqrt{1/(2l+1)}$. This property is the main motivation for considering the azimuthally averaged harmonic spectrum.

We find it useful to define an azimuthally averaged harmonic transform, $e_l(\hat{\mathbf{n}})$, as

$$e_l(\hat{\mathbf{n}}) \equiv \sqrt{\frac{4\pi}{2l+1}} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\hat{\mathbf{n}}) \ , \tag{420}$$

which is interpreted as a square-root of C_l at a given position of the sky,

$$\int \frac{\mathrm{d}^2\hat{\mathbf{n}}}{4\pi} e^2_l(\hat{\mathbf{n}}) = C_l \ . \tag{421}$$

$e_l(\hat{\mathbf{n}})$ is particularly useful for measuring the angular bispectrum [268,145], trispectrum (Chapter 6 of Ref. [137]), and probably any higher-order harmonic spectra, because of being computationally very fast to calculate. This is very important, as the new satellite experiments, WMAP and Planck, have more than millions of pixels, for which we will crucially need a fast algorithm for measuring these higher-order harmonic spectra.

We derive the covariance matrix of C_l, $\langle C_l C_{l'}\rangle - \langle C_l\rangle\langle C_{l'}\rangle$, with the four-point function, the trispectrum. Starting with

$$\langle C_l C_{l'}\rangle = \frac{1}{(2l+1)(2l'+1)} \sum_{mm'} \langle a_{lm} a^*_{lm} a_{l'm'} a^*_{l'm'}\rangle \ , \tag{422}$$

we obtain the power-spectrum covariance matrix

$$\langle C_l C_{l'}\rangle - \langle C_l\rangle\langle C_{l'}\rangle = \frac{2\langle C_l\rangle^2}{2l+1}\,\delta_{ll'} + \frac{1}{(2l+1)(2l'+1)}\sum_{mm'}\langle a_{lm}a^*_{lm}a_{l'm'}a^*_{l'm'}\rangle_{\rm c}$$
$$= \frac{2\langle C_l\rangle^2}{2l+1}\,\delta_{ll'} + \frac{(-1)^{l+l'}}{\sqrt{(2l+1)(2l'+1)}}\,\langle T^{ll}_{l'l'}(0)\rangle_{\rm c}\ , \tag{423}$$

where $\langle a_{lm}a^*_{lm}a_{l'm'}a^*_{l'm'}\rangle_{\rm c}$ is the connected four-point harmonic spectrum, the connected trispectrum, which is exactly zero for a Gaussian field. It follows from this equation that the covariance matrix of C_l is exactly diagonal only when a_{lm} is Gaussian. $\langle T^{l_1l_2}_{l_3l_4}(L)\rangle_{\rm c}$ is the ensemble average of the angular averaged connected trispectrum, which we will define in Section 10.1.4 (Eq. (435)).

Unfortunately, we cannot measure the connected $T^{ll}_{l'l'}(0)$ directly from the angular trispectrum (see Section 10.1.4). We will thus never be sure if the power-spectrum covariance is precisely diagonal, as long as we use the angular trispectrum. We need other statistics able to pick up information on the connected $T^{ll}_{l'l'}(0)$, even though they are indirect. Otherwise, we need a model for the connected trispectrum, and use the model to constrain the connected $T^{ll}_{l'l'}(0)$ from the other trispectrum configurations.

There is no reason to assume the connected $T^{ll}_{l'l'}(0)$ is small. It is produced on large angular scales, if the topology of the Universe is closed hyperbolic [118]. An analytic prediction for the connected trispectrum produced in a closed hyperbolic Universe is derived in Appendix D of Ref. [137]. On small angular scales, several authors have shown that the weak gravitational lensing effect produces a non-zero connected trispectrum or four-point correlation function [43,301,299]; in Ref. [114] the induced off-diagonal terms are found to be negligible compared with the diagonal terms out to $l\sim 2000$.

If the connected trispectrum is negligible, then we obtain

$$\langle C_l C_{l'}\rangle - \langle C_l\rangle\langle C_{l'}\rangle \approx \frac{2\langle C_l\rangle^2}{2l+1}\delta_{ll'}\ . \tag{424}$$

The fractional error of C_l is thus proportional to $\sqrt{1/(2l+1)}$, as expected from our having $2l+1$ independent samples to average for a given l. The exact form follows from C_l being χ^2 distributed with $2l+1$ degrees of freedom when a_{lm} is Gaussian. If a_{lm} is Gaussian, then its probability density distribution is

$$P(a_{lm}) = \frac{\exp[-a^2_{lm}/(2\langle C_l\rangle)]}{\sqrt{2\pi\langle C_l\rangle}}\ . \tag{425}$$

One can use this distribution to generate Gaussian random realizations of a_{lm} for a given $\langle C_l\rangle$. First, calculate $\langle C_l\rangle$ with the CMBFAST code [260] for a set of cosmological parameters. Next, generate a realization of a_{lm}, $a_{lm}=\epsilon\langle C_l\rangle^{1/2}$, where ϵ is a Gaussian random variable with unit variance.

10.1.3. Angular bispectrum

The angular bispectrum consists of three harmonic transforms, $a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}$. For Gaussian a_{lm}, the expectation value is exactly zero. By imposing statistical isotropy upon the angular three-point correlation function, one finds that the angular averaged bispectrum, $B_{l_1l_2l_3}$, given by

$$\langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle = \langle B_{l_1l_2l_3}\rangle \begin{pmatrix} l_1 & l_2 & l_3\\ m_1 & m_2 & m_3\end{pmatrix} \tag{426}$$

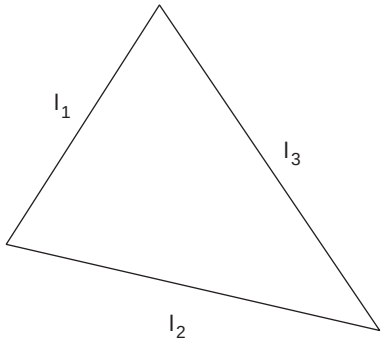

Fig. 5. Angular bispectrum configuration.

satisfies rotational invariance (Eq. (416)). Here, the matrix denotes the Wigner-3j symbol (see Appendix C). Since l_1, l_2, and l_3 form a triangle, $B_{l_1 l_2 l_3}$ satisfies the triangle condition, $|l_i - l_j| \leqslant l_k \leqslant l_i + l_j$ for all permutations of indices. Parity invariance of the angular correlation function demands $l_1 + l_2 + l_3 = \text{even}$. Fig. 5 sketches a configuration of the angular bispectrum.

The Wigner-3j symbol, which describes the coupling of two angular momenta, represents the azimuthal angle dependence of the angular bispectrum, since the bispectrum forms a triangle. Suppose that two "states" with (l_1, m_1) and (l_2, m_2) angular momenta form a coupled state with (l_3, m_3). They form a triangle whose orientation is represented by m_1, m_2, and m_3, with satisfying $m_1 + m_2 + m_3 = 0$. As we rotate the system, the Wigner-3j symbol transforms m's, yet preserving the configuration of the triangle. Similarly, rotational invariance of the angular bispectrum demands that the same triangle configuration gives the same amplitude of the bispectrum regardless of its orientation, and thus the Wigner-3j symbol describes the azimuthal angle dependence.

The proof of $\langle B_{l_1 l_2 l_3} \rangle$ to be rotationally invariant is as follows. Substituting Eq. (426) for the statistical isotropy condition (Eq. (416)) for $n = 3$, we obtain

$$
\begin{aligned}
&\langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle \\
&\quad = \sum_{\text{all } m'} \langle a_{l_1 m_1'} a_{l_2 m_2'} a_{l_3 m_3'} \rangle D^{(l_1)}_{m_1' m_1} D^{(l_2)}_{m_2' m_2} D^{(l_3)}_{m_3' m_3} \\
&\quad = \langle B_{l_1 l_2 l_3} \rangle \sum_{\text{all } m'} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1' & m_2' & m_3' \end{pmatrix} \\
&\qquad \times \sum_{LMM'} \begin{pmatrix} l_1 & l_2 & L \\ m_1' & m_2' & M' \end{pmatrix} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M \end{pmatrix} (2L+1) D^{(L)*}_{M'M} D^{(l_3)}_{m_3' m_3} \\
&\quad = \langle B_{l_1 l_2 l_3} \rangle \sum_{m_3'} \sum_{LMM'} \delta_{l_3 L} \delta_{m_3' M'} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M \end{pmatrix} D^{(L)*}_{M'M} D^{(l_3)}_{m_3' m_3} \\
&\quad = \langle B_{l_1 l_2 l_3} \rangle \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} . \qquad (427)
\end{aligned}
$$

In the second equality, we have reduced $D^{(l_1)}_{m'_1m_1}D^{(l_2)}_{m'_2m_2}$ to $D^{(L)*}_{M'M}$, using Eq. (C.17). In the third equality, we have used the identity [245],

$$\sum_{m'_1m'_2}\begin{pmatrix} l_1 & l_2 & l_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & L \\ m'_1 & m'_2 & M' \end{pmatrix} = \frac{\delta_{l_3L}\delta_{m'_3M'}}{2L+1} \ . \tag{428}$$

To obtain the unbiased estimator of the angular averaged bispectrum, $B_{l_1l_2l_3}$, we invert Eq. (426) with the identity Eq. (428), and obtain

$$B_{l_1l_2l_3} = \sum_{\text{all } m}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} a_{l_1m_1}a_{l_2m_2}a_{l_3m_3} \ . \tag{429}$$

We can rewrite this expression into a more computationally useful form. Using the azimuthally averaged harmonic transform, $e_l(\hat{\mathbf{n}})$ (Eq. (420)), and the identity [245],

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^{-1}\sqrt{\frac{(4\pi)^3}{(2l_1+1)(2l_2+1)(2l_3+1)}} \times \int \frac{\mathrm{d}^2\hat{\mathbf{n}}}{4\pi} Y_{l_1m_1}(\hat{\mathbf{n}})Y_{l_2m_2}(\hat{\mathbf{n}})Y_{l_3m_3}(\hat{\mathbf{n}}) \ , \tag{430}$$

we rewrite Eq. (429) as

$$B_{l_1l_2l_3} = \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^{-1}\int \frac{\mathrm{d}^2\hat{\mathbf{n}}}{4\pi} e_{l_1}(\hat{\mathbf{n}})e_{l_2}(\hat{\mathbf{n}})e_{l_3}(\hat{\mathbf{n}}) \ . \tag{431}$$

This expression is computationally efficient; we can quickly calculate $e_l(\hat{\mathbf{n}})$ with the spherical harmonic transform. Then, the average over the full sky, $\int \mathrm{d}^2\hat{\mathbf{n}}/(4\pi)$, is done by the sum over all pixels divided by the total number of pixels, $N^{-1}\sum_i^N$, if all the pixels have equal area. Note that the integral over $\hat{\mathbf{n}}$ must be done over the full sky even when a sky-cut is applied, as $e_l(\hat{\mathbf{n}})$ already encapsulates information on partial sky coverage through the a_{lm}, which may be measured on an incomplete sky.

We calculate the covariance matrix of $B_{l_1l_2l_3}$, provided that non-Gaussianity is weak, $\langle B_{l_1l_2l_3}\rangle \approx 0$. Since the covariance matrix is a product of six a_{lm}'s, we have ${}_6C_2\cdot{}_4C_2/3! = 15$ terms to evaluate, according to the Wick's theorem; however, using the identity [245],

$$\sum_m (-1)^m\begin{pmatrix} l & l & l' \\ m & -m & 0 \end{pmatrix} = \delta_{l'0} \ , \tag{432}$$

and assuming that none of the l's zero, we find only $3! = 6$ terms that do not include $\langle a_{l_im_i}a_{l_jm_j}\rangle$ but include only $\langle a_{l_im_i}a^*_{l_jm_j}\rangle$ non-vanishing. Evaluating these six terms, we obtain [171,108,268,88]

$$\begin{aligned}\langle B_{l_1l_2l_3}B_{l'_1l'_2l'_3}\rangle &= \sum_{\text{all } mm'}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} l'_1 & l'_2 & l'_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix}\langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}a^*_{l'_1m'_1}a^*_{l'_2m'_2}a^*_{l'_3m'_3}\rangle \\ &= \langle C_{l_1}\rangle\langle C_{l_2}\rangle\langle C_{l_3}\rangle[\delta^{l'_1l'_2l'_3}_{l_1l_2l_3} + \delta^{l'_3l'_1l'_2}_{l_1l_2l_3} + \delta^{l'_2l'_3l'_1}_{l_1l_2l_3} \\ &\quad + (-1)^{l_1+l_2+l_3}(\delta^{l'_1l'_3l'_2}_{l_1l_2l_3} + \delta^{l'_2l'_1l'_3}_{l_1l_2l_3} + \delta^{l'_3l'_2l'_1}_{l_1l_2l_3})] \ ,\end{aligned} \tag{433}$$

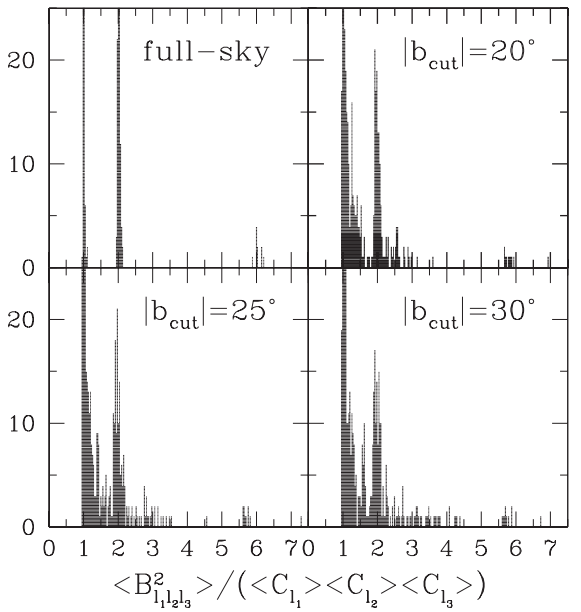

Fig. 6. Variance of angular bispectrum. Histograms of variance of the angular bispectrum for $l_1 \leqslant l_2 \leqslant l_3$ up to a maximum multipole of 20. There are 466 modes. These are derived from simulated realizations of a Gaussian sky. The top-left panel shows the case of full sky coverage, while the rest of panels show the cases of incomplete sky coverage. The top-right, bottom-left, and bottom-right panels use the 20, 25, and 30° Galactic sky-cuts, respectively.

where $\delta^{l_1' l_2' l_3'}_{l_1 l_2 l_3} \equiv \delta_{l_1 l_1'} \delta_{l_2 l_2'} \delta_{l_3 l_3'}$, and so on. Hence, the covariance matrix is diagonal in the weak non-Gaussian limit. The diagonal terms for $l_i \neq 0$ and $l_1 + l_2 + l_3 =$ even are

$$\langle B^2_{l_1 l_2 l_3} \rangle = \langle C_{l_1} \rangle \langle C_{l_2} \rangle \langle C_{l_3} \rangle (1 + 2\delta_{l_1 l_2} \delta_{l_2 l_3} + \delta_{l_1 l_2} + \delta_{l_2 l_3} + \delta_{l_3 l_1}) \ . \tag{434}$$

The variance is amplified by a factor of 2 or 6, when two or all l's are the same, respectively.

One finds that Eq. (434) becomes non-exact on the incomplete sky, where the variance distribution becomes more scattered. Using simulated realizations of a Gaussian sky, the authors of Ref. [139] measured the variance on the full sky as well as on the incomplete sky for three different Galactic sky-cuts, 20, 25 and 30°. Fig. 6 plots the results; one finds that Eq. (434) holds only approximately on the incomplete sky.

10.1.4. Angular trispectrum

The angular trispectrum consists of four harmonic transforms, $a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} a_{l_4 m_4}$. In Ref. [114] a rotationally invariant solution for the angular trispectrum was found as

$$\langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} a_{l_4 m_4} \rangle = \sum_{LM} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_3 & l_4 & L \\ m_3 & m_4 & M \end{pmatrix} (-1)^M \langle T^{l_1 l_2}_{l_3 l_4}(L) \rangle \ . \tag{435}$$

One can prove that this solution, $\langle T^{l_1 l_2}_{l_3 l_4}(L) \rangle$, is rotationally invariant by similar calculations to those proving the angular bispectrum to be so. By construction, l_1, l_2, and L form one triangle, while l_3, l_4, and L form the other triangle in a quadrilateral with sides l_1, l_2, l_3, and l_4. L represents a diagonal of the quadrilateral. Fig. 7 sketches a configuration of the angular trispectrum. When we arrange l_1, l_2, l_3, and

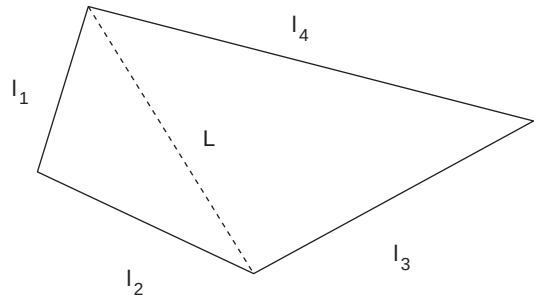

Fig. 7. Angular trispectrum configuration.

l_4 in order of $l_1 \leqslant l_2 \leqslant l_3 \leqslant l_4$, L lies in $\max(l_2 - l_1, l_4 - l_3) \leqslant L \leqslant \min(l_1 + l_2, l_3 + l_4)$. Parity invariance of the angular four-point correlation function demands $l_1 + l_2 + L =$ even and $l_3 + l_4 + L =$ even.

The angular trispectrum generically consists of two parts. One is the disconnected part, the contribution from Gaussian fields, which is given by the angular power-spectra [114],

$$\langle T^{l_1 l_2}_{l_3 l_4}(L)\rangle_{\text{disconnected}} = (-1)^{l_1+l_3}\sqrt{(2l_1+1)(2l_3+1)}\langle C_{l_1}\rangle\langle C_{l_3}\rangle\delta_{l_1 l_2}\delta_{l_3 l_4}\delta_{L0} + (2L+1)\langle C_{l_1}\rangle\langle C_{l_2}\rangle[(-1)^{l_2+l_3+L}\delta_{l_1 l_3}\delta_{l_2 l_4} + \delta_{l_1 l_4}\delta_{l_2 l_3}] \,. \tag{436}$$

For $l_1 \leqslant l_2 \leqslant l_3 \leqslant l_4$, the disconnected terms are non-zero only when $L = 0$ or $l_1 = l_2 = l_3 = l_4$. We have numerically confirmed that the estimator given below (Eq. (441)) accurately reproduces the disconnected terms (Eq. (436)) on a simulated Gaussian sky.

The other is the connected part whose expectation value is exactly zero for Gaussian fields; thus, the connected part is sensitive to non-Gaussianity. When none of the l's are the same in $T^{l_1 l_2}_{l_3 l_4}(L)$, one might expect the trispectrum to comprise the connected part only; however, this is true only on the full sky. The disconnected terms on the incomplete sky, which are often much bigger than the connected terms, leak the power to the other modes for which all l's are different. One should take this effect into account in the analysis.

Inverting Eq. (435), we obtain the unbiased estimator of $T^{l_1 l_2}_{l_3 l_4}(L)$ [114],

$$T^{l_1 l_2}_{l_3 l_4}(L) = (2L+1)\sum_{\text{all } m}\sum_{M}(-1)^M \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M \end{pmatrix}\begin{pmatrix} l_3 & l_4 & L \\ m_3 & m_4 & -M \end{pmatrix} \times a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} a_{l_4 m_4} \,. \tag{437}$$

Note that this expression includes both the connected and the disconnected terms.

We find that this estimator has a special property for $L = 0$ (which demands $l_1 = l_2$ and $l_3 = l_4$). The trispectrum estimator for these configurations, $T^{l_1 l_1}_{l_3 l_3}(0)$, reduces to a product of two power-spectrum estimators, $C_{l_1} C_{l_3}$,

$$T^{l_1 l_1}_{l_3 l_3}(0) = \sum_{m_1 m_3}\begin{pmatrix} l_1 & l_1 & 0 \\ m_1 & -m_1 & 0 \end{pmatrix}\begin{pmatrix} l_3 & l_3 & 0 \\ m_3 & -m_3 & 0 \end{pmatrix} a_{l_1 m_1} a_{l_1 -m_1} a_{l_3 m_3} a_{l_3 -m_3} = (-1)^{l_1+l_3}\sqrt{(2l_1+1)(2l_3+1)}C_{l_1}C_{l_3} \,, \tag{438}$$

where $C_l = (2l+1)^{-1}\sum_m a_{lm}a_{lm}^*$. We have used the identity, Eq. (432), and $a_{l-m} = (-1)^m a_{lm}^*$ in the second equality. From this equation, one may assume that $T^{l_1 l_1}_{l_3 l_3}(0)$ coincides with the disconnected terms for $L = 0$ (see Eq. (436)),

$$\langle T^{l_1 l_1}_{l_3 l_3}(0)\rangle_{\text{disconnected}} = (-1)^{l_1+l_3}\sqrt{(2l_1+1)(2l_3+1)}\langle C_{l_1}\rangle\langle C_{l_3}\rangle + 2\langle C_{l_1}\rangle^2\delta_{l_1 l_3} \ . \tag{439}$$

They are, however, different for non-Gaussian fields, because of the power-spectrum covariance, Eq. (423). By taking the ensemble average of $T^{l_1 l_1}_{l_3 l_3}(0)$, and substituting Eq. (423) for $\langle C_{l_1} C_{l_3}\rangle$, we find a rather trivial result:

$$\begin{aligned}\langle T^{l_1 l_1}_{l_3 l_3}(0)\rangle &= (-1)^{l_1+l_3}\sqrt{(2l_1+1)(2l_3+1)}\langle C_{l_1} C_{l_3}\rangle \\ &= (-1)^{l_1+l_3}\sqrt{(2l_1+1)(2l_3+1)}\langle C_{l_1}\rangle\langle C_{l_3}\rangle + 2\langle C_{l_1}\rangle^2\delta_{l_1 l_3} + \langle T^{l_1 l_1}_{l_3 l_3}(0)\rangle_{\text{c}} \\ &= \langle T^{l_1 l_1}_{l_3 l_3}(0)\rangle_{\text{disconnected}} + \langle T^{l_1 l_1}_{l_3 l_3}(0)\rangle_{\text{c}} \ . \end{aligned} \tag{440}$$

Hence, $T^{l_1 l_1}_{l_3 l_3}(0)$ contains information not only on the disconnected trispectrum, but also of the connected one.

Unfortunately, we cannot measure the connected part of $T^{l_1 l_1}_{l_3 l_3}(0)$ directly from the angular trispectrum because of the following reason. To measure the connected terms, we have to subtract the disconnected terms from the measured trispectrum first. Since we are never able to measure the ensemble average of the disconnected terms (Eq. (439)), we estimate them by using the estimated power-spectrum, C_l. If we subtract the estimated disconnected terms, $\propto C_{l_1} C_{l_3}$, from measured $T^{l_1 l_1}_{l_3 l_3}(0)$, then it follows from Eq. (438) that $T^{l_1 l_1}_{l_3 l_3}(0)$ vanishes *exactly*: $T^{l_1 l_1}_{l_3 l_3}(0) = 0$; thus, $T^{l_1 l_1}_{l_3 l_3}(0)$ has no statistical power of measuring the connected terms.

For practical measurement of the angular trispectrum, we rewrite the trispectrum estimator given by Eq. (437) with the azimuthally averaged harmonic transform, $e_l(\hat{\mathbf{n}})$ (Eq. (420)). We find that the following form is particularly computationally efficient:

$$T^{l_1 l_2}_{l_3 l_4}(L) = \frac{1}{2L+1}\sum_{M=-L}^{L} t^{l_1 l_2 *}_{LM} t^{l_3 l_4}_{LM} \ , \tag{441}$$

where $t^{l_1 l_2}_{LM}$ is given by

$$t^{l_1 l_2}_{LM} \equiv \sqrt{\frac{2L+1}{4\pi}}\begin{pmatrix} l_1 & l_2 & L \\ 0 & 0 & 0 \end{pmatrix}^{-1}\int \mathrm{d}^2\hat{\mathbf{n}}[e_{l_1}(\hat{\mathbf{n}})e_{l_2}(\hat{\mathbf{n}})]Y^*_{LM}(\hat{\mathbf{n}}) \ . \tag{442}$$

Since $t^{l_1 l_2}_{LM}$ is the harmonic transform on the full sky, we can calculate it quickly. This method makes measurement of the angular trispectrum computationally feasible even for the WMAP data in which we have more than millions of pixels; thus, the methods developed here can be applied not only to the COBE DMR data, but also to the WMAP data.

Let us calculate the covariance of the trispectrum in the weakly non-Gaussian limit. Since the trispectrum covariance comprises eight a_{lm}'s, the total number of terms is ${}_8C_2\cdot{}_6C_2\cdot{}_4C_2/4! = 105$ according to the Wick's theorem. The full calculation will be a nightmare, for we have to deal with

$$\begin{aligned}\langle T^{l_1l_2}_{l_3l_4}(L)T^{l'_1l'_2}_{l'_3l'_4}(L')\rangle = (2L+1)(2L'+1)\sum_{\text{all } mm'}\sum_{MM'}(-1)^{M+M'}\\ \times\begin{pmatrix} l_1 & l_2 & L\\ m_1 & m_2 & M\end{pmatrix}\begin{pmatrix} l_3 & l_4 & L\\ m_3 & m_4 & -M\end{pmatrix}\\ \times\begin{pmatrix} l'_1 & l'_2 & L'\\ m'_1 & m'_2 & M'\end{pmatrix}\begin{pmatrix} l'_3 & l'_4 & L'\\ m'_3 & m'_4 & -M'\end{pmatrix}\\ \times\langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}a_{l_4m_4}a^*_{l'_1m'_1}a^*_{l'_2m'_2}a^*_{l'_3m'_3}a^*_{l'_4m'_4}\rangle\ .\end{aligned}\tag{443}$$

One can reduce this intricate expression to much a simpler form for some particular configurations. For $L, L' \neq 0$ terms, thanks to the identity (432), only $4! = 24$ terms that do not include $\langle a_{l_im_i}a_{l_jm_j}\rangle$ but include only $\langle a_{l_im_i}a^*_{l_jm_j}\rangle$ are non-vanishing. For $L = L' = 0$ terms, the triangle conditions in a quadrilateral demand $l_1 = l_2$ and $l_3 = l_4$ (see Fig. 7). As we have shown, these configurations have no statistical power of measuring the connected trispectrum of interest. Hence, we evaluate $L, L' \neq 0$ terms in the following.

Evaluating the 24 $L, L' \neq 0$ terms is still a headache; however, for $l_1 \leqslant l_2 < l_3 \leqslant l_4$, we have only eight terms left:

$$\begin{aligned}&\frac{\langle T^{l_1l_2}_{l_3l_4}(L)T^{l'_1l'_2}_{l'_3l'_4}(L')\rangle}{(2L+1)\langle C_{l_1}\rangle\langle C_{l_2}\rangle\langle C_{l_3}\rangle\langle C_{l_4}\rangle}\\ &\quad=\delta_{LL'}[\delta^{l_1l_2l_3l_4}_{l'_1l'_2l'_3l'_4}+\delta^{l_1l_2l_3l_4}_{l'_3l'_4l'_1l'_2}+(-1)^{l_1+l_2+l_3+l_4}(\delta^{l_1l_2l_3l_4}_{l'_2l'_1l'_4l'_3}+\delta^{l_1l_2l_3l_4}_{l'_4l'_3l'_2l'_1})\\ &\qquad+(-1)^{l_1+l_2+L}(\delta^{l_1l_2l_3l_4}_{l'_2l'_1l'_3l'_4}+\delta^{l_1l_2l_3l_4}_{l'_4l'_3l'_1l'_2})+(-1)^{l_3+l_4+L}(\delta^{l_1l_2l_3l_4}_{l'_1l'_2l'_4l'_3}+\delta^{l_1l_2l_3l_4}_{l'_3l'_4l'_2l'_1})]\ ,\end{aligned}\tag{444}$$

where $\delta^{l'_1l'_2l'_3l'_4}_{l_1l_2l_3l_4} \equiv \delta_{l_1l'_1}\delta_{l_2l'_2}\delta_{l_3l'_3}\delta_{l_4l'_4}$, and so on. Using parity invariance, $l_1+l_2+L =$ even and $l_3+l_4+L =$ even, one finds the covariance matrix diagonal. Thus, the diagonal terms for $L \neq 0$ and $l_1 \leqslant l_2 < l_3 \leqslant l_4$ are simplified very much as

$$\langle[T^{l_1l_2}_{l_3l_4}(L)]^2\rangle = (2L+1)\langle C_{l_1}\rangle\langle C_{l_2}\rangle\langle C_{l_3}\rangle\langle C_{l_4}\rangle(1+\delta_{l_1l_2}+\delta_{l_3l_4}+\delta_{l_1l_2}\delta_{l_3l_4})\ .\tag{445}$$

This result is strictly correct only on the full sky; the incomplete sky makes the variance distribution much more scattered. Fig. 8 plots the variance on the full sky as well as on the incomplete sky.

For the rest of configurations for which the disconnected terms vanish, $L \neq 0$, $l_2 = l_3$, and $l_1 \neq l_4$, the covariance matrix is no longer diagonal in L, L' [114]. Fig. 9 plots the numerically evaluated variance on the full sky as well as on the incomplete sky. The variance divided by $(2L+1)\langle C_{l_1}\rangle\langle C_{l_2}\rangle\langle C_{l_3}\rangle\langle C_{l_4}\rangle$ is no longer an integer, but more scattered than that for $L \neq 0$ and $l_1 \leqslant l_2 < l_3 \leqslant l_4$ even on the full sky.

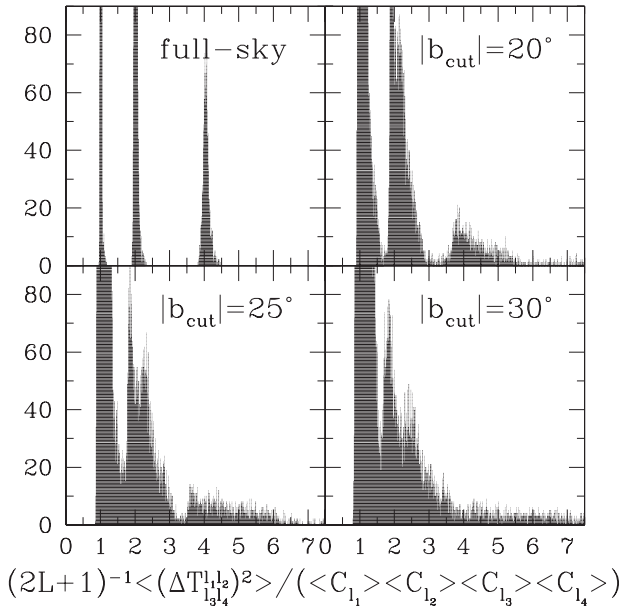

Fig. 8. Variance of angular trispectrum I. Histograms of variance of the angular trispectrum for $L \neq 0$ and $l_1 \leqslant l_2 < l_3 \leqslant l_4$, for which the disconnected terms vanish on the full sky. There are 16,554 modes, up to a maximum multipole of 20. The meaning of the panels is the same as in Fig. 6.

10.1.5. Power-spectrum and bispectrum on the incomplete sky

Incomplete sky coverage destroys the orthonormality of the spherical harmonics on the sky. The degree to which orthonormality is broken is often characterized by the coupling integral [221],

$$W_{ll'mm'} \equiv \int \mathrm{d}^2\hat{\mathbf{n}} W(\hat{\mathbf{n}}) Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}) = \int_{\Omega_{\rm obs}} \mathrm{d}^2\hat{\mathbf{n}} Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}) \ , \tag{446}$$

where $W(\hat{\mathbf{n}})$ is zero in a cut region otherwise 1, and $\Omega_{\rm obs}$ denotes the solid angle of the observed sky. When $W_{ll'mm'} \neq \delta_{ll'}\delta_{mm'}$, the measured harmonic transform of the temperature anisotropy field, a_{lm}, becomes a *biased* estimator of the true harmonic transform, $a^{\rm true}_{lm}$, through

$$a_{lm} = \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} a^{\rm true}_{l'm'} W_{ll'mm'} \ . \tag{447}$$

Hence, we must correct our estimators of the power-spectrum and the bispectrum for the bias arising from incomplete sky coverage.

First, we derive a relationship between the angular power-spectrum on the incomplete sky and that on the full sky. Taking the ensemble average of the estimator of the power-spectrum, the pseudo-C_l [288],

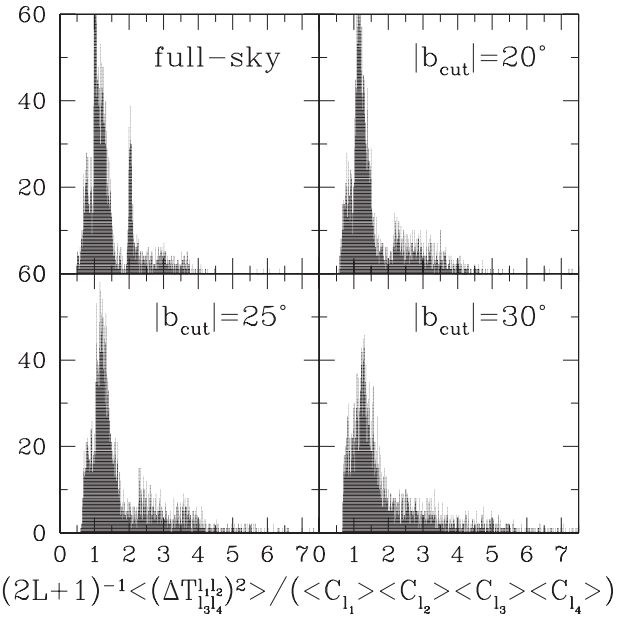

Fig. 9. Variance of angular trispectrum II. Histograms of variance of the angular trispectrum for $L \neq 0$, $l_2 = l_3$, and $l_1 \neq l_4$, for which the disconnected terms vanish on the full sky. There are 4059 modes, up to a maximum multipole of 20. The meaning of the panels is the same as in Fig. 6.

$C_l = (2l+1)^{-1}\sum_m |a_{lm}|^2$, we have

$$\begin{aligned}
\langle C_l \rangle &= \frac{1}{2l+1}\sum_{l'} C_{l'}^{\text{true}} \sum_{mm'} |W_{ll'mm'}|^2 \\
&\approx \frac{1}{2l+1} C_l^{\text{true}} \sum_m \sum_{l'm'} \int \mathrm{d}^2\hat{\mathbf{n}} W(\hat{\mathbf{n}}) Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}) \int \mathrm{d}^2\hat{\mathbf{m}} W(\hat{\mathbf{m}}) Y_{lm}(\hat{\mathbf{m}}) Y^*_{l'm'}(\hat{\mathbf{m}}) \\
&= \frac{1}{2l+1} C_l^{\text{true}} \sum_m \int \mathrm{d}^2\hat{\mathbf{n}} W(\hat{\mathbf{n}}) Y^*_{lm}(\hat{\mathbf{n}}) \int \mathrm{d}^2\hat{\mathbf{m}} W(\hat{\mathbf{m}}) Y_{lm}(\hat{\mathbf{m}}) \delta^{(2)}(\hat{\mathbf{n}} - \hat{\mathbf{m}}) \\
&= C_l^{\text{true}} \int \frac{\mathrm{d}^2\hat{\mathbf{n}}}{4\pi} W(\hat{\mathbf{n}}) P_l(1) \\
&= C_l^{\text{true}} \frac{\Omega_{\text{obs}}}{4\pi} \,. \qquad (448)
\end{aligned}$$

In the second equality, we have taken $C_{l'}^{\text{true}}$ out of the summation over l', as $|W_{ll'mm'}|^2$ peaks very sharply at $l = l'$, and $C_{l'}^{\text{true}}$ varies much more slowly than $|W_{ll'mm'}|^2$ in l'. This approximation is good for nearly full sky coverage. In the third equality, we have used $\sum_{l'm'} Y_{l'm'}(\hat{\mathbf{n}}) Y^*_{l'm'}(\hat{\mathbf{m}}) = \delta^{(2)}(\hat{\mathbf{n}} - \hat{\mathbf{m}})$. In the fourth equality, we have used $\sum_m Y^*_{lm}(\hat{\mathbf{n}}) Y_{lm}(\hat{\mathbf{m}}) = (2l+1)/4\pi\, P_l(\hat{\mathbf{n}} \cdot \hat{\mathbf{m}})$. The result indicates that the bias amounts approximately to a fraction of the sky covered by observations.

Next, we derive a relationship between the angular bispectrum on the incomplete sky and that on the full sky. We begin with

$$\langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle = \sum_{\text{all } l'm'} \langle a^{\text{true}}_{l'_1m'_1}a^{\text{true}}_{l'_2m'_2}a^{\text{true}}_{l'_3m'_3}\rangle W_{l_1l'_1m_1m'_1}W_{l_2l'_2m_2m'_2}W_{l_3l'_3m_3m'_3} \ . \tag{449}$$

Rotational and parity invariance of the bispectrum implies that the bispectrum is given by

$$\langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle = b_{l_1l_2l_3}\int \mathrm{d}^2\hat{\mathbf{n}}Y^*_{l_1m_1}(\hat{\mathbf{n}})Y^*_{l_2m_2}(\hat{\mathbf{n}})Y^*_{l_3m_3}(\hat{\mathbf{n}}) \ , \tag{450}$$

where $b_{l_1l_2l_3}$ is an arbitrary real symmetric function, which is related to the angular averaged bispectrum, $B_{l_1l_2l_3}$. When $b^{\text{true}}_{l_1l_2l_3}$ varies much more slowly than the coupling integral, we obtain

$$\begin{aligned}\langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle = &\sum_{\text{all } l'} b^{\text{true}}_{l'_1l'_2l'_3}\sum_{\text{all } m'}\int \mathrm{d}^2\hat{\mathbf{n}}Y^*_{l'_1m'_1}(\hat{\mathbf{n}})Y^*_{l'_2m'_2}(\hat{\mathbf{n}})Y^*_{l'_3m'_3}(\hat{\mathbf{n}})\\ &\times\int \mathrm{d}^2\hat{\mathbf{n}}_1 W(\hat{\mathbf{n}}_1)Y_{l'_1m'_1}(\hat{\mathbf{n}}_1)Y^*_{l_1m_1}(\hat{\mathbf{n}}_1)\\ &\times\int \mathrm{d}^2\hat{\mathbf{n}}_2 W(\hat{\mathbf{n}}_2)Y_{l'_2m'_2}(\hat{\mathbf{n}}_2)Y^*_{l_2m_2}(\hat{\mathbf{n}}_2)\\ &\times\int \mathrm{d}^2\hat{\mathbf{n}}_3 W(\hat{\mathbf{n}}_3)Y_{l'_3m'_3}(\hat{\mathbf{n}}_3)Y^*_{l_3m_3}(\hat{\mathbf{n}}_3)\\ \approx\ & b^{\text{true}}_{l_1l_2l_3}\int \mathrm{d}^2\hat{\mathbf{n}}W(\hat{\mathbf{n}})Y^*_{l_1m_1}(\hat{\mathbf{n}})Y^*_{l_2m_2}(\hat{\mathbf{n}})Y^*_{l_3m_3}(\hat{\mathbf{n}}) \ .\end{aligned} \tag{451}$$

Then, we calculate the angular averaged bispectrum, $B_{l_1l_2l_3}$ (Eq. (426)). By convolving Eq. (429) with the Wigner-3j symbol and using the identity Eq. (430), we obtain

$$\begin{aligned}\langle B_{l_1l_2l_3}\rangle \approx\ & b^{\text{true}}_{l_1l_2l_3}\sqrt{\frac{4\pi}{(2l_1+1)(2l_2+1)(2l_3+1)}}\begin{pmatrix}l_1 & l_2 & l_3\\ 0 & 0 & 0\end{pmatrix}^{-1}\\ &\times\sum_{\text{all } m}\int \mathrm{d}^2\hat{\mathbf{m}}Y_{l_1m_1}(\hat{\mathbf{m}})Y_{l_2m_2}(\hat{\mathbf{m}})Y_{l_3m_3}(\hat{\mathbf{m}})\\ &\times\int \mathrm{d}^2\hat{\mathbf{n}}W(\hat{\mathbf{n}})Y^*_{l_1m_1}(\hat{\mathbf{n}})Y^*_{l_2m_2}(\hat{\mathbf{n}})Y^*_{l_3m_3}(\hat{\mathbf{n}})\\ =\ & b^{\text{true}}_{l_1l_2l_3}\sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}}\begin{pmatrix}l_1 & l_2 & l_3\\ 0 & 0 & 0\end{pmatrix}^{-1}\\ &\times\int\frac{\mathrm{d}^2\hat{\mathbf{m}}}{4\pi}\int\frac{\mathrm{d}^2\hat{\mathbf{n}}}{4\pi}W(\hat{\mathbf{n}})P_{l_1}(\hat{\mathbf{m}}\cdot\hat{\mathbf{n}})P_{l_2}(\hat{\mathbf{m}}\cdot\hat{\mathbf{n}})P_{l_3}(\hat{\mathbf{m}}\cdot\hat{\mathbf{n}})\\ =\ & b^{\text{true}}_{l_1l_2l_3}\sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}}\begin{pmatrix}l_1 & l_2 & l_3\\ 0 & 0 & 0\end{pmatrix}\frac{\Omega_{\text{obs}}}{4\pi}\\ =\ & B^{\text{true}}_{l_1l_2l_3}\ \frac{\Omega_{\text{obs}}}{4\pi} \ ,\end{aligned} \tag{452}$$

where we have used the identity,

$$\int_{-1}^{1} \frac{\mathrm{d}x}{2} P_{l_1}(x) P_{l_2}(x) P_{l_3}(x) = \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 . \tag{453}$$

Thus, the bias for the angular bispectrum on the incomplete sky is also approximately given by a fraction of the sky covered by observations.

10.2. Theoretical predictions for the CMB bispectrum from inflation

In this section, we derive analytical predictions for the angular bispectrum from inflation. We expand the observed CMB temperature fluctuation field, $\Delta T(\hat{\mathbf{n}})/T$, into the spherical harmonics,

$$a_{lm} = \int \mathrm{d}^2\hat{\mathbf{n}} \, \frac{\Delta T(\hat{\mathbf{n}})}{T} Y^*_{lm}(\hat{\mathbf{n}}) \ , \tag{454}$$

where the hats denote unit vectors. The CMB angular bispectrum is given by

$$B^{m_1 m_2 m_3}_{l_1 l_2 l_3} \equiv \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle \ , \tag{455}$$

and the angular averaged bispectrum is (Eq. (429))

$$B_{l_1 l_2 l_3} = \sum_{\text{all } m} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} B^{m_1 m_2 m_3}_{l_1 l_2 l_3} \ , \tag{456}$$

where the matrix is the Wigner-3j symbol (see Appendix C). The bispectrum, $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$, satisfies the triangle conditions and parity invariance: $m_1 + m_2 + m_3 = 0$, $l_1 + l_2 + l_3 = \text{even}$, and $|l_i - l_j| \leqslant l_k \leqslant l_i + l_j$ for all permutations of indices. It implies that $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ consists of the Gaunt integral, $\mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3}$, defined by

$$\begin{aligned} \mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} &\equiv \int \mathrm{d}^2\hat{\mathbf{n}} Y_{l_1 m_1}(\hat{\mathbf{n}}) Y_{l_2 m_2}(\hat{\mathbf{n}}) Y_{l_3 m_3}(\hat{\mathbf{n}}) \\ &= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} . \end{aligned} \tag{457}$$

$\mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ is real, and satisfies all the conditions mentioned above.

Rotational invariance of the angular three-point correlation function implies that $B_{l_1 l_2 l_3}$ is written as

$$B^{m_1 m_2 m_3}_{l_1 l_2 l_3} = \mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} b_{l_1 l_2 l_3} \ , \tag{458}$$

where $b_{l_1 l_2 l_3}$ is an arbitrary real symmetric function of l_1, l_2, and l_3. This form, Eq. (458), is necessary and sufficient to construct generic $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ under rotational invariance; thus, we will use $b_{l_1 l_2 l_3}$ more frequently than $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ in this section, and call this function the *reduced* bispectrum, as $b_{l_1 l_2 l_3}$ contains all physical information in $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$. Since the reduced bispectrum does not contain the Wigner-3j symbol, which merely ensures the triangle conditions and parity invariance, it is easier to calculate physical properties of the bispectrum.

We calculate the angular averaged bispectrum, $B_{l_1l_2l_3}$, by substituting Eq. (458) into Eq. (456),

$$B_{l_1l_2l_3} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} b_{l_1l_2l_3} \ , \tag{459}$$

where we have used the identity,

$$\sum_{\text{all } m} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} . \tag{460}$$

Alternatively, one can define the bispectrum in the flat-sky approximation,

$$\langle a(\mathbf{l}_1)a(\mathbf{l}_1)a(\mathbf{l}_3)\rangle = (2\pi)^2\delta^{(2)}(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3)B(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3) \ , \tag{461}$$

where $\mathbf{l}$ is a two-dimensional wave vector on the sky. This definition of $B(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3)$ reduces to Eq. (458) with the correspondence

$$\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3} \to (2\pi)^2\delta^{(2)}(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3) \ , \tag{462}$$

in the flat-sky limit [113]. Thus, we have

$$b_{l_1l_2l_3} \approx B(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3) \quad \text{(flat-sky approximation)} \ . \tag{463}$$

This fact motivates our use of the reduced bispectrum, $b_{l_1l_2l_3}$, rather than the angular averaged bispectrum, $B_{l_1l_2l_3}$. Note that $b_{l_1l_2l_3}$ is similar to $\hat{B}_{l_1l_2l_3}$ defined in Ref. [179]; the relation is $b_{l_1l_2l_3} = \sqrt{4\pi}\hat{B}_{l_1l_2l_3}$.

If primordial fluctuations are adiabatic scalar fluctuations, then

$$a_{lm} = 4\pi(-i)^l \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \Phi(\mathbf{k}) g_{\mathrm{T}l}(k) Y^*_{lm}(\hat{\mathbf{k}}) \ , \tag{464}$$

where $\Phi(\mathbf{k})$ is the primordial curvature perturbation in Fourier space, and $g_{\mathrm{T}l}(k)$ is the radiation transfer function. a_{lm} takes over the non-Gaussianity, if any, from $\Phi(\mathbf{k})$. Although Eq. (464) is valid only if the Universe is flat, it is straightforward to extend this to an arbitrary geometry. We can calculate the isocurvature fluctuations similarly by using the entropy perturbation and the proper transfer function.

As it has been shown in Section 8, the primordial non-Gaussianity may be parameterized as a linear plus quadratic term in the gravitational potential in the general form of Eq. (296), where the non-linearity parameter f_{NL} appears as a kernel in Fourier space, rather than a constant. This gives rise to an angular modulation of the quadratic non-linearity, which might be used to search for specific signatures of inflationary non-Gaussianity in the CMB [161]. In this section, however, we restrict ourselves to the simplest weak non-linear coupling case, assuming that f_{NL} is merely a multiplicative constant, as done in data analyses so far. Hence we write

$$\Phi(\mathbf{x}) = \Phi_{\mathrm{L}}(\mathbf{x}) + f_{\mathrm{NL}}[\Phi^2_{\mathrm{L}}(\mathbf{x}) - \langle\Phi^2_{\mathrm{L}}(\mathbf{x})\rangle] \ , \tag{465}$$

in real space, where $\Phi_{\mathrm{L}}(\mathbf{x})$ denotes the linear Gaussian part of the perturbation, and $\langle\Phi(\mathbf{x})\rangle = 0$ is guaranteed.

In Fourier space, we decompose $\Phi(\mathbf{k})$ into two parts,

$$\Phi(\mathbf{k}) = \Phi_{\mathrm{L}}(\mathbf{k}) + \Phi_{\mathrm{NL}}(\mathbf{k}) \ , \tag{466}$$

and accordingly we have

$$a_{lm} = a_{lm}^{\rm L} + a_{lm}^{\rm NL} \ , \tag{467}$$

where $\Phi_{\rm NL}(\mathbf{k})$ is a non-linear curvature perturbation defined by

$$\Phi_{\rm NL}(\mathbf{k}) \equiv f_{\rm NL}\left[\int \frac{{\rm d}^3\mathbf{p}}{(2\pi)^3}\Phi_{\rm L}(\mathbf{k}+\mathbf{p})\Phi_{\rm L}^*(\mathbf{p}) - (2\pi)^3\delta^{(3)}(\mathbf{k})\langle\Phi_{\rm L}^2(\mathbf{x})\rangle\right] \ . \tag{468}$$

One can immediately check that $\langle\Phi(\mathbf{k})\rangle = 0$ is satisfied. In this model, a non-vanishing component of the $\Phi(\mathbf{k})$-field bispectrum is

$$\langle\Phi_{\rm L}(\mathbf{k}_1)\Phi_{\rm L}(\mathbf{k}_2)\Phi_{\rm NL}(\mathbf{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)2f_{\rm NL}P_\Phi(k_1)P_\Phi(k_2) \ , \tag{469}$$

where $P_\Phi(k)$ is Bardeen's potential linear power-spectrum given by

$$\langle\Phi_{\rm L}(\mathbf{k}_1)\Phi_{\rm L}(\mathbf{k}_2)\rangle = (2\pi)^3 P_\Phi(k_1)\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2) \ . \tag{470}$$

We have also used

$$\langle\Phi_{\rm L}(\mathbf{k}+\mathbf{p})\Phi_{\rm L}^*(\mathbf{p})\rangle = (2\pi)^3 P_\Phi(p)\delta^{(3)}(\mathbf{k}) \ , \tag{471}$$

and

$$\langle\Phi_{\rm L}^2(\mathbf{x})\rangle = (2\pi)^{-3}\int {\rm d}^3\mathbf{k}P_\Phi(k) \ . \tag{472}$$

Substituting Eq. (464) into Eq. (455), using Eq. (469) for the $\Phi(\mathbf{k})$-field bispectrum, and then integrating over angles $\hat{\mathbf{k}}_1$, $\hat{\mathbf{k}}_3$, and $\hat{\mathbf{k}}_3$, we obtain the primordial CMB angular bispectrum,

$$\begin{aligned} B_{l_1l_2l_3}^{m_1m_2m_3} &= \langle a_{l_1m_1}^{\rm L}a_{l_2m_2}^{\rm L}a_{l_3m_3}^{\rm NL}\rangle + \langle a_{l_1m_1}^{\rm L}a_{l_2m_2}^{\rm NL}a_{l_3m_3}^{\rm L}\rangle + \langle a_{l_1m_1}^{\rm NL}a_{l_2m_2}^{\rm L}a_{l_3m_3}^{\rm L}\rangle \\ &= 2\mathscr{G}_{l_1l_2l_3}^{m_1m_2m_3}\int_0^\infty r^2\,{\rm d}r[b_{l_1}^{\rm L}(r)b_{l_2}^{\rm L}(r)b_{l_3}^{\rm NL}(r) + b_{l_1}^{\rm L}(r)b_{l_2}^{\rm NL}(r)b_{l_3}^{\rm L}(r) \\ &\quad + b_{l_1}^{\rm NL}(r)b_{l_2}^{\rm L}(r)b_{l_3}^{\rm L}(r)] \ , \end{aligned} \tag{473}$$

where

$$b_l^{\rm L}(r) \equiv \frac{2}{\pi}\int_0^\infty k^2\,{\rm d}k\,P_\Phi(k)g_{{\rm T}l}(k)\,j_l(kr) \ , \tag{474}$$

$$b_l^{\rm NL}(r) \equiv \frac{2}{\pi}\int_0^\infty k^2\,{\rm d}k f_{\rm NL}g_{{\rm T}l}(k)\,j_l(kr) \ . \tag{475}$$

Note that $b_l^{\rm L}(r)$ is dimensionless, while $b_l^{\rm NL}(r)$ has a dimension of L^{-3}.

One can immediately check that Eq. (458) holds; thus, the reduced bispectrum, $b_{l_1l_2l_3}$ (Eq. (458)), for the primordial non-Gaussianity reads

$$b_{l_1l_2l_3}^{\rm prim} = 2\int_0^\infty r^2\,{\rm d}r[b_{l_1}^{\rm L}(r)b_{l_2}^{\rm L}(r)b_{l_3}^{\rm NL}(r) + b_{l_1}^{\rm L}(r)b_{l_2}^{\rm NL}(r)b_{l_3}^{\rm L}(r) + b_{l_1}^{\rm NL}(r)b_{l_2}^{\rm L}(r)b_{l_3}^{\rm L}(r)] \ . \tag{476}$$

We can fully specify $b_{l_1l_2l_3}^{\rm prim}$ by a single constant parameter, $f_{\rm NL}$, as the CMB angular power-spectrum, C_l, will precisely measure cosmological parameters. We stress again that this formula is valid only when

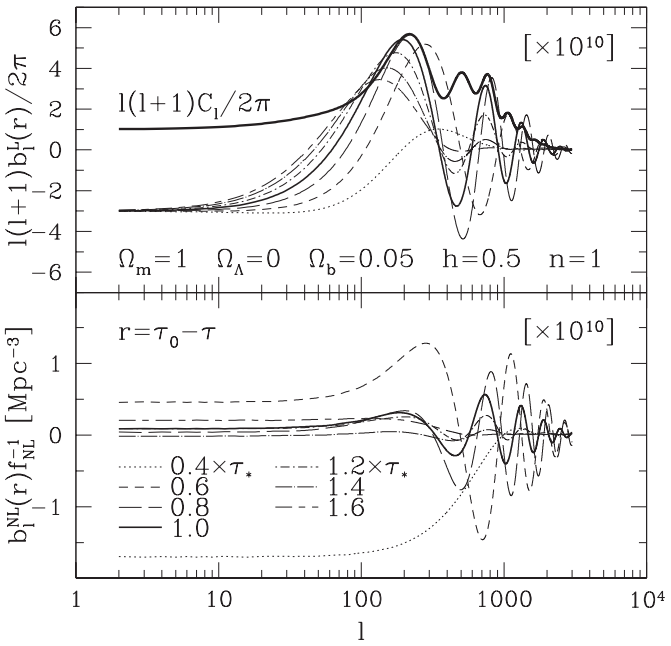

Fig. 10. Components of primordial CMB bispectrum [143]. This figure shows $b_l^{\rm L}(r)$ (Eq. (474)) and $b_l^{\rm NL}(r)$ (Eq. (475)), the two terms in our calculation of the primordial CMB angular bispectrum, as a function of r. Various lines in the top panel show $[l(l+1)b_l^{\rm L}(r)/2\pi]\times 10^{10}$, where $r=c(\tau_0-\tau)$, at $\tau=0.4, 0.6, 0.8, 1.0, 1.2, 1.4$, and $1.6\times\tau_*$ (decoupling time); $[b_l^{\rm NL}(r)\,f_{\rm NL}^{-1}]\times 10^{10}$ are shown in the bottom panel. τ_0 is the present-day conformal time. Note that $c\tau_0=11.8$ Gpc, and $c\tau_*=235$ Mpc in our cosmological model chosen here. The thickest solid line in the top panel is the CMB angular power-spectrum, $[l(l+1)C_l/2\pi]\times 10^{10}$. C_l is shown for comparison.

the scale-dependence of $f_{\rm NL}$ is weak, which is a good approximation if the momentum-independent part of $f_{\rm NL}$ is larger than unity.

One can calculate the primordial CMB bispectrum (Eqs. (473)–(476)) numerically as follows. One computes the full radiation transfer function, $g_{{\rm T}l}(k)$, with the CMBFAST code [260], assuming a single power-law spectrum, $P_\Phi(k)\propto k^{n-4}$, for the primordial curvature fluctuations. After doing the integration over k (Eqs. (474) and (475)) with the same algorithm of CMBFAST , one performs the integration over r (Eq. (476)), $r=c(\tau_0-\tau)$, where τ is the conformal time. τ_0 is the present-day value. In our model, $c\tau_0=11.8$ Gpc, and the decoupling occurs at $c\tau_*=235$ Mpc at which the differential visibility has a maximum. Our $c\tau_0$ includes radiation effects on the expansion of the Universe; otherwise, $c\tau_0=12.0$ Gpc. Since most of the primordial signal is generated at τ_*, we choose the r integration boundary as $c(\tau_0-2\tau_*)\leqslant r\leqslant c(\tau_0-0.1\tau_*)$. We use a step-size of $0.1c\tau_*$, as we have found that a step size of $0.01c\tau_*$ gives very similar results. As cosmological model, let us assume a scale-invariant standard cold dark matter (SCDM) model with $\Omega_{\rm m}=1$, $\Omega_\Lambda=0$, $\Omega_{\rm b}=0.05$, $h=0.5$, and $n=1$, and with power-spectrum $P_\Phi(k)$ normalized to COBE [56]. Although this model is almost excluded by current observations, it is still useful to depict the basic effects of the transfer function on the bispectrum (see also Ref. [161]).

Fig. 10 shows $b_l^{\rm L}(r)$ (Eq. (474)) and $b_l^{\rm NL}(r)$ (Eq. (475)) for several different values of r. We find that $b_l^{\rm L}(r)$ and C_l look very similar to each other in shape and amplitude at $l\gtrsim 100$, although the amplitude

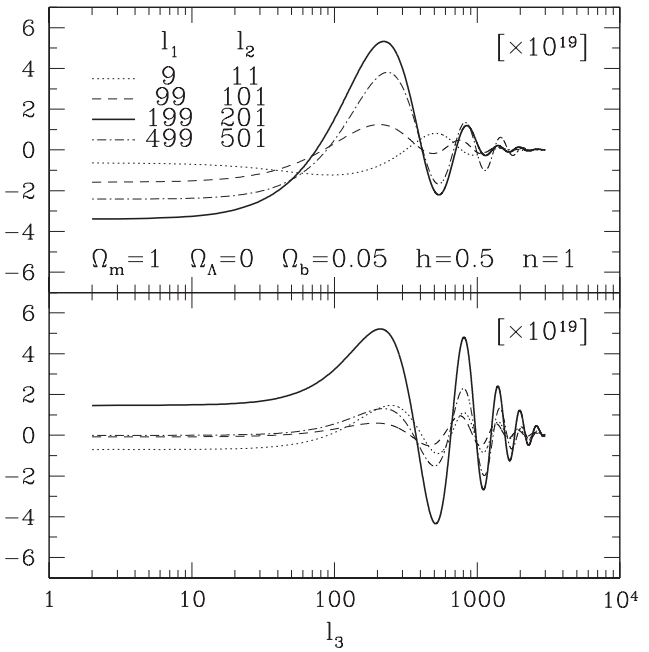

Fig. 11. Primordial CMB bispectrum [143]. The primordial angular bispectrum (Eq. (473)), divided by the Gaunt integral, $\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3}$ (Eq. (457)). The bispectrum is plotted as a function of l_3 for (l_1, l_2)=(9,11), (99,101), (199,201), and (499,501). Each panel plots a different 1-dimensional slice of the bispectrum. The top panel plots $l_2(l_2+1)l_3(l_3+1)\langle a^{\rm NL}_{l_1m_1}a^{\rm L}_{l_2m_2}a^{\rm L}_{l_3m_3}\rangle f^{-1}_{\rm NL}(\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3})^{-1}/(2\pi)^2$, while the bottom panel plots $l_1(l_1+1)l_2(l_2+1)\langle a^{\rm L}_{l_1m_1}a^{\rm L}_{l_2m_2}a^{\rm NL}_{l_3m_3}\rangle f^{-1}_{\rm NL}(\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3})^{-1}/(2\pi)^2$. Note that we have multiplied the bispectrum in each panel by a factor of 10^{19}.

in the Sachs–Wolfe regime is different by a factor of -3. This is because $C_l \propto P_\Phi(k)g^2_{{\rm T}l}(k)$, while $b^{\rm L}_l(r) \propto P_\Phi(k)g_{{\rm T}l}(k)$, where $g_{{\rm T}l} = -\frac{1}{3}$. We also find that $b^{\rm L}_l(r)$ has a good phase coherence over a wide range of r, while the phase of $b^{\rm NL}_l(r)$ in the high-l regime oscillates rapidly as a function of r. This strongly damps the integrated result (Eq. (473)) in the high-l regime. The main difference between C_l and $b_l(r)$ is that $b_l(r)$ changes the sign, while C_l does not.

Looking at Fig. 10, we find $l^2b^{\rm L}_l \sim 2\times10^{-9}$ and $b^{\rm NL}_l f^{-1}_{\rm NL} \sim 10^{-10}\,{\rm Mpc}^{-3}$. As most of the signal is coming from the decoupling epoch, the volume element at τ_* is $r^2_*\Delta r_* \sim (10^4)^2\times10^2\,{\rm Mpc}^3$; thus, we can give an order-of-magnitude estimate of the primordial reduced bispectrum (Eq. (476)) as

$$b^{\rm prim}_{lll} \sim l^{-4}[2r^2_*\Delta r_*(l^2b^{\rm L}_l)^2b^{\rm NL}_l\times3] \sim l^{-4}\times2\times10^{-17}f_{\rm NL}\ . \tag{477}$$

Since $b^{\rm NL}_l f^{-1}_{\rm NL} \sim r^{-2}_*\delta(r-r_*)$ (see Eq. (480)), $r^2_*\Delta r_* b^{\rm NL}_l f^{-1}_{\rm NL} \sim 1$. This rough estimate agrees with the numerical result below (Fig. 11).

Fig. 11 shows the integrated bispectrum (Eq. (473)) divided by the Gaunt integral, $\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3}$, which is the reduced bispectrum, $b^{\rm prim}_{l_1l_2l_3}$. While the bispectrum is a 3-d function, we show different 1-d slices of the bispectrum in this figure. We plot

$$l_2(l_2+1)l_3(l_3+1)\langle a^{\rm NL}_{l_1m_1}a^{\rm L}_{l_2m_2}a^{\rm L}_{l_3m_3}\rangle(\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3})^{-1}/(2\pi)^2$$

as a function of l_3 in the top panel, while we plot

$$l_1(l_1+1)l_2(l_2+1)\langle a^{\rm L}_{l_1m_1}a^{\rm L}_{l_2m_2}a^{\rm NL}_{l_3m_3}\rangle(\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3})^{-1}/(2\pi)^2$$

in the bottom panel. We have multiplied each $b^{\rm L}_l(r)$ which contains $P_\Phi(k)$ by $l(l+1)/(2\pi)$ so that the Sachs–Wolfe plateau at $l_3 \lesssim 10$ is easily seen. We have chosen l_1 and l_2 so as $(l_1, l_2) = (9, 11)$, $(99, 101)$, $(199, 201)$, and $(499, 501)$. We find that the $(l_1, l_2) = (199, 201)$ mode, the first acoustic peak mode, has the largest signal in this family of parameters. The top panel has a prominent first acoustic peak, and strongly damped oscillations in the high-l regime; the bottom panel also has a first peak, but damps more slowly. The typical amplitude of the reduced bispectrum is $l^4 b^{\rm prim}_{lll} f^{-1}_{\rm NL} \sim 10^{-17}$, which agrees with the order-of-magnitude estimate of Eq. (477).

The formula in Eq. (476) and numerical results agree with [86] in the Sachs–Wolfe regime, where $g_{{\rm T}l}(k) \approx -j_l(kr_*)/3$, and

$$b^{\rm prim}_{l_1l_2l_3} \approx -6 f_{\rm NL}(C^{\rm SW}_{l_1}C^{\rm SW}_{l_2} + C^{\rm SW}_{l_1}C^{\rm SW}_{l_3} + C^{\rm SW}_{l_2}C^{\rm SW}_{l_3}) \quad \text{(SW approximation)} \ . \tag{478}$$

Each term is of the same order as Eq. (476). Here, $C^{\rm SW}_l$ is the CMB angular power-spectrum in the Sachs–Wolfe approximation,

$$C^{\rm SW}_l \equiv \frac{2}{9\pi}\int_0^\infty k^2\,{\rm d}k\,P_\Phi(k)\,j_l^2(kr_*) \ . \tag{479}$$

In deriving Eq. (478) from Eq. (476), we have approximated $b^{\rm NL}_l(r)$ (Eq. (475)) with

$$b^{\rm NL}_l(r) \approx \left(-\frac{f_{\rm NL}}{3}\right)\frac{2}{\pi}\int_0^\infty k^2\,{\rm d}k j_l(kr_*)\,j_l(kr) = -\frac{f_{\rm NL}}{3}\,r_*^{-2}\delta(r-r_*) \ . \tag{480}$$

The Sachs–Wolfe approximation (Eq. (478)) is valid only when l_1, l_2, and l_3 are all smaller than ~ 10, for which the authors of Ref. [86] give $\sim -6\times 10^{-20}$ in Fig. 11. We stress again that the Sachs–Wolfe approximation gives a qualitatively different result from our full calculation (Eq. (476)) at $l_i \gtrsim 10$. The full bispectrum changes sign, while the approximation never changes sign because of the use of $C^{\rm SW}_l$. The acoustic oscillation and the sign-change are actually great advantages when we try to separate the primordial bispectrum from various secondary bispectra. We will analyze this point later.

As we have calculated the full bispectrum at all scales, it is now possible to calculate the three-point function in real space. Unlike the bispectrum, however, the form of the full three-point function is fairly complicated; nevertheless, one can obtain a simple form for the skewness, S_3, given by

$$S_3 \equiv \left\langle\left(\frac{\Delta T(\hat{\mathbf{n}})}{T}\right)^3\right\rangle \ , \tag{481}$$

which is perhaps the simplest (but less powerful) statistic characterizing non-Gaussianity. We expand S_3 in terms of $B_{l_1l_2l_3}$ (Eq. (456)), or $b_{l_1l_2l_3}$ (Eq. (458)), as

$$\begin{aligned} S_3 &= \frac{1}{4\pi}\sum_{l_1l_2l_3}\sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}}\begin{pmatrix} l_1 & l_2 & l_3\\ 0 & 0 & 0\end{pmatrix}B_{l_1l_2l_3}W_{l_1}W_{l_2}W_{l_3} \\ &= \frac{1}{2\pi^2}\sum_{2\leqslant l_1l_2l_3}\left(l_1+\frac{1}{2}\right)\left(l_2+\frac{1}{2}\right)\left(l_3+\frac{1}{2}\right)\begin{pmatrix} l_1 & l_2 & l_3\\ 0 & 0 & 0\end{pmatrix}^2 b_{l_1l_2l_3}W_{l_1}W_{l_2}W_{l_3} \ , \end{aligned} \tag{482}$$

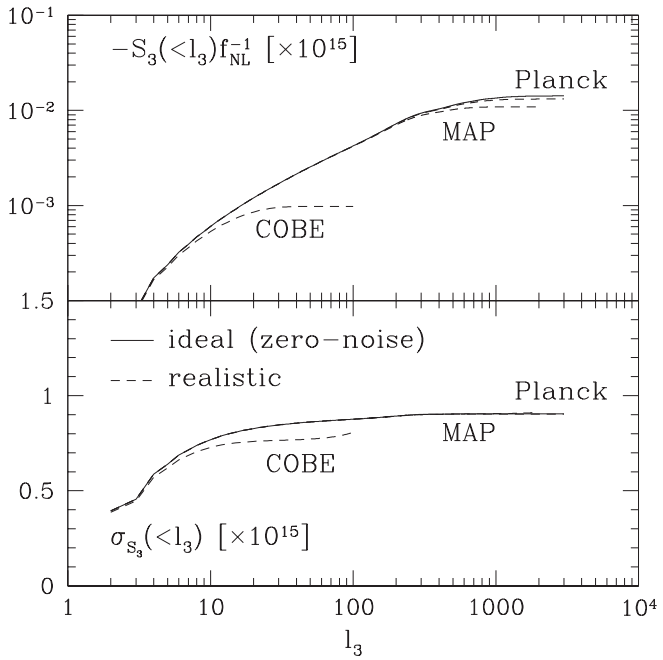

Fig. 12. Primordial skewness [143]. The top panel shows the primordial CMB skewness (Eq. (482)) summed up to a certain l_3, $-S_3(<l_3)f_{\rm NL}^{-1}\times 10^{15}$. The bottom panel shows the error of S_3 (Eq. (517)) summed up to l_3, $\sigma_{S_3}(<l_3)\times 10^{15}$. The solid line represents the zero-noise ideal experiment, while the dotted lines show COBE, WMAP, and Planck experiments.

where W_l is the experimental window function. We have used Eq. (459) to replace $B_{l_1l_2l_3}$ by the reduced bispectrum, $b_{l_1l_2l_3}$, in the last equality. Since $l=0$ and 1 modes are not observable, we have excluded them from the summation. Throughout this section, we consider a single-beam window function, $W_l = \mathrm{e}^{-l(l+1)/(2\sigma_{\rm b}^2)}$, where $\sigma_{\rm b} = \mathrm{FWHM}/\sqrt{8\ln 2}$. Since $\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 b_{l_1l_2l_3}$ is symmetric under permutation of indices, we change the way of summation as

$$\sum_{2\leqslant l_1l_2l_3} \longrightarrow 6 \sum_{2\leqslant l_1\leqslant l_2\leqslant l_3} \quad . \tag{483}$$

This reduces the number of summations by a factor of $\simeq 6$. We will use this convention henceforth.

The top panel of Fig. 12 plots $S_3(<l_3)$, which is S_3 summed up to a certain l_3, for FWHM beam sizes of 7°, 13′, and 5′.5. These values correspond to COBE, WMAP, and Planck beam sizes, respectively. Fig. 12 also plots the infinitesimally thin beam case. We find that WMAP, Planck, and the ideal experiments measure very similar S_3 to one another, despite the fact that Planck and the ideal experiments can use many more modes than WMAP. The reason is as follows. Looking at Eq. (482), one finds that S_3 is a linear integral of $b_{l_1l_2l_3}$ over l_i; thus, integrating oscillations in $b^{\rm prim}_{l_1l_2l_3}$ around zero (see Fig. 11) damps the non-Gaussian signal on small angular scales, $l \gtrsim 300$. Since the Sachs–Wolfe effect, implying no oscillation, dominates the COBE-scale anisotropy, the cancellation on the COBE scale affects S_3 less significantly than on the WMAP and Planck scales. Planck suffers from severe cancellation in small angular scales:

Planck and the ideal experiments measure only the same amount of S_3 as WMAP does. As a result, the measured S_3 almost saturates at the WMAP resolution scale, $l \sim 500$.

We conclude this section by noting that when we can calculate the expected form of the bispectrum, then it becomes a "matched filter" for detecting non-Gaussianity in the data, and thus much more powerful a tool than the skewness in which the information is lost through the coarse-graining.

10.3. Secondary sources of CMB bispectrum

Even if the CMB bispectrum were significantly detected in the CMB map, its origin would not necessarily be primordial, but rather there would be various secondary sources such as the Sunyaev–Zel'dovich (SZ) effect [302], the weak lensing effect, and so on, or foreground sources such as extragalactic radio sources. To isolate the primordial origin from the others, we have to know the accurate form of bispectra produced by secondary and foreground sources.

10.3.1. Coupling between the weak lensing and the Sunyaev–Zel'dovich effects

The coupling between the SZ effect and the weak lensing effect produces an observable effect in the bispectrum [92,70]. We expand the CMB temperature field including the SZ and the lensing effect as

$$\begin{aligned}\frac{\Delta T(\hat{\mathbf{n}})}{T} &= \frac{\Delta T^{\mathrm{P}}(\hat{\mathbf{n}}+\nabla\Theta(\hat{\mathbf{n}}))}{T} + \frac{\Delta T^{\mathrm{SZ}}(\hat{\mathbf{n}})}{T} \\ &\approx \frac{\Delta T^{\mathrm{P}}(\hat{\mathbf{n}})}{T} + \nabla\left(\frac{\Delta T^{\mathrm{P}}(\hat{\mathbf{n}})}{T}\right)\cdot\nabla\Theta(\hat{\mathbf{n}}) + \frac{\Delta T^{\mathrm{SZ}}(\hat{\mathbf{n}})}{T}\ ,\end{aligned} \tag{484}$$

where P denotes the primordial anisotropy, $\Theta(\hat{\mathbf{n}})$ is the lensing potential,

$$\Theta(\hat{\mathbf{n}}) \equiv -2\int_0^{r_*} \mathrm{d}r\,\frac{r_*-r}{r r_*}\,\Phi(r,\hat{\mathbf{n}}r)\ , \tag{485}$$

and SZ denotes the SZ effect,

$$\frac{\Delta T^{\mathrm{SZ}}(\hat{\mathbf{n}})}{T} = y(\hat{\mathbf{n}})\, j_\nu, \tag{486}$$

where j_ν is a spectral function of the SZ effect [302]. $y(\hat{\mathbf{n}})$ is the Compton y-parameter given by

$$y(\hat{\mathbf{n}}) \equiv y_0 \int \frac{\mathrm{d}r}{r_*}\,\frac{T_\rho(r,\hat{\mathbf{n}}r)}{\overline{T}_{\rho 0}}\,a^{-2}(r)\ , \tag{487}$$

where

$$y_0 \equiv \frac{\sigma_T \overline{\rho}_{\mathrm{gas}0} k_{\mathrm{B}} \overline{T}_{\rho 0} r_*}{\mu_e m_p m_e c^2} = 4.3\times 10^{-4}\mu_e^{-1}(\Omega_{\mathrm{b}}h^2)\left(\frac{k_{\mathrm{B}}\overline{T}_{\rho 0}}{1\,\mathrm{keV}}\right)\left(\frac{r_*}{10\,\mathrm{Gpc}}\right)\ . \tag{488}$$

$T_\rho \equiv \rho_{\mathrm{gas}} T_e/\overline{\rho}_{\mathrm{gas}}$ is the electron temperature weighted by the gas mass density, the overline denotes the volume average, and the subscript 0 means the present epoch. We adopt $\mu_e^{-1} = 0.88$, where $\mu_e^{-1} \equiv n_e/(\rho_{\mathrm{gas}}/m_p)$ is the number of electrons per proton mass in the fully ionized medium. Other quantities have their usual meaning.

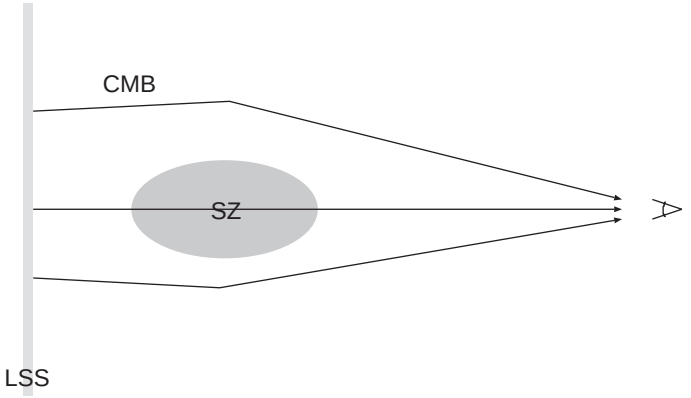

Fig. 13. SZ-lensing coupling. A schematic view of the SZ-lensing coupling bispectrum. One of the three CMB photons, which are decoupled at the last scattering surface (LSS), penetrates through a SZ cluster, changing its temperature, and coming toward us. As the other two photons pass near the SZ cluster, they are deflected by the lensing effect, changing their propagation directions, and coming toward us. As a result, the three photons are correlated, generating a three-point correlation, the bispectrum.

Transforming Eq. (484) into harmonic space, we obtain

$$\begin{aligned} a_{lm} &= a^{\mathrm{P}}_{lm} + \sum_{l'm'}\sum_{l''m''} (-1)^m \mathscr{G}^{-mm'm''}_{ll'l''} \frac{l'(l'+1) - l(l+1) + l''(l''+1)}{2} a^{\mathrm{P}}_{l'm'} \Theta_{l''m''} + a^{\mathrm{SZ}}_{lm} \\ &= a^{\mathrm{P}}_{lm} + \sum_{l'm'}\sum_{l''m''} (-1)^{m+m'+m''} \mathscr{G}^{-mm'm''}_{ll'l''} \\ &\quad \times \frac{l'(l'+1) - l(l+1) + l''(l''+1)}{2} a^{\mathrm{P}*}_{l'-m'} \Theta^*_{l''-m''} + a^{\mathrm{SZ}}_{lm} \ , \end{aligned} \tag{489}$$

where $\mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3}$ is the Gaunt integral (Eq. (457)). Substituting Eq. (489) into Eq. (455), and using the identity, $\mathscr{G}^{-m_1-m_2-m_3}_{l_1l_2l_3} = \mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3}$, we obtain the bispectrum,

$$\begin{aligned} B^{m_1m_2m_3}_{l_1l_2l_3} = \mathscr{G}^{m_1m_2m_3}_{l_1l_2l_3} \Bigg[& \frac{l_1(l_1+1) - l_2(l_2+1) + l_3(l_3+1)}{2} C^{\mathrm{P}}_{l_1} \langle \Theta^*_{l_3m_3} a^{\mathrm{SZ}}_{l_3m_3} \rangle \\ & + 5 \text{ permutations} \Bigg] \ . \end{aligned} \tag{490}$$

The form of Eq. (458) is confirmed; the reduced bispectrum $b^{\text{sz-lens}}_{l_1l_2l_3}$ includes the terms in square brackets.

While Eq. (490) is complicated, we can understand the physical effect producing the SZ-lensing bispectrum intuitively. Fig. 13 shows how the SZ-lensing coupling produces the three-point correlation. Suppose that there are three CMB photons decoupled at the last scattering surface (LSS), and one of these photons penetrates through a SZ cluster between the LSS and us; the energy of the photon changes because of the SZ effect. When the other two photons pass near the SZ cluster, they are deflected by the gravitational lensing effect, changing their propagation directions, and coming toward us. What do we see after all? We see that the three CMB photons are correlated; we then measure a non-zero angular bispectrum. The cross-correlation strength between the SZ and lensing effects, $\langle \Theta^*_{l_3m_3} a^{\mathrm{SZ}}_{l_3m_3} \rangle$, thus determines the bispectrum amplitude, as indicated by Eq. (490).

In Ref. [92] $\langle \Theta^*_{lm} a^{\rm SZ}_{lm}\rangle$ was derived assuming the linear pressure bias model [226], $T_\rho = \overline{T}_\rho b_{\rm gas}\delta$, and the mean temperature evolution, $\overline{T}_\rho \simeq \overline{T}_{\rho 0}(1+z)^{-1}$, for $z<2$, which is roughly suggested by recent hydrodynamic simulations [60,237,270]. They obtained

$$\langle \Theta^*_{lm} a^{\rm SZ}_{lm}\rangle \simeq -j_\nu \frac{4 y_0 b_{\rm gas} l^2}{3\Omega_{\rm m} H_0^2}\int_0^{z_*} {\rm d}z \frac{{\rm d}r}{{\rm d}z} D^2(z)(1+z)^2 \frac{r_* - r(z)}{r_*^2 r^5(z)} P_\Phi\left(k=\frac{l}{r(z)}\right) , \tag{491}$$

where $D(z)$ is the linear growth factor. Simulations without non-gravitational heating [237,270] suggest that $\overline{T}_{\rho 0} \sim 0.2$–$0.4\,{\rm keV}$ and $b_{\rm gas} \sim 5$–10; analytic estimations give similar numbers [237,303]. In the pressure bias model, the free parameters (except cosmological ones) are $\overline{T}_{\rho 0}$ and $b_{\rm gas}$; however, both actually depend upon the cosmological model [237]. Since $l^3\langle \Theta^*_{lm} a^{\rm SZ}_{lm}\rangle \sim 2\times 10^{-10} j_\nu \overline{T}_{\rho 0} b_{\rm gas}$ [92,70] and $l^2 C_l^{\rm P} \sim 6\times 10^{-10}$, we have

$$b^{\text{sz-lens}}_{lll} \sim l^{-3}[(l^2 C_l^{\rm P})(l^3\langle \Theta^*_{lm} a^{\rm SZ}_{lm}\rangle)\times 5/2] \sim l^{-3}\times 3\times 10^{-19} j_\nu \overline{T}_{\rho 0} b_{\rm gas} , \tag{492}$$

where $\overline{T}_{\rho 0}$ is in units of 1 keV, and $b_{l_1 l_2 l_3} = B^{m_1 m_2 m_3}_{l_1 l_2 l_3}(\mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3})^{-1}$ is the reduced bispectrum (Eq. (458)). Comparing this with Eq. (477), we obtain

$$\frac{b^{\rm prim}_{lll}}{b^{\text{sz-lens}}_{lll}} \sim l^{-1}\times 10\left(\frac{f_{\rm NL}}{j_\nu \overline{T}_{\rho 0} b_{\rm gas}}\right) . \tag{493}$$

This estimate suggests that the SZ-lensing bispectrum overwhelms the primordial bispectrum on small angular scales. This is why we have to separate the primordial from the SZ-lensing effect.

While the pressure bias model gives a rough estimate of the SZ power-spectrum, more accurate predictions exist. Several authors have predicted the SZ power-spectrum analytically using the Press–Schechter approach [63,180,16,140,68,202] or the hyper-extended perturbation theory [303]. The predictions agree with hydrodynamic simulations well [237,258,270,239]. While a big uncertainty in the predictions lies in phenomenological models which describe the SZ surface brightness profile of halos, the authors of Refs. [141,142] have proposed universal gas and temperature profiles and predicted the SZ profile relying on a more physical basis; they have then used the universal profiles to improve upon the analytic prediction for the SZ power-spectrum. The universal profiles should describe the SZ profile in the average sense; on the individual halo-to-halo basis, there could be significant deviation from the universal profile, owing to substructures in halos (see, e.g., Ref. [138]).

10.3.2. Extragalactic radio and infrared sources

The bispectrum from extragalactic radio and infrared sources whose fluxes, F, are smaller than a certain detection threshold, $F_{\rm d}$, is simple to estimate, when we assume the Poisson distribution. The authors of Ref. [279,13] have shown that the Poisson distribution is a good approximation at low frequencies ($\nu < 100\,{\rm GHz}$). The Poisson distribution has white-noise power-spectrum; thus, the reduced bispectrum (Eq. (458)) is constant, $b^{\rm src}_{l_1 l_2 l_3} = b^{\rm src} = $ constant, and we obtain

$$B^{m_1 m_2 m_3}_{l_1 l_2 l_3} = \mathscr{G}^{m_1 m_2 m_3}_{l_l 1 2 l_3} b^{\rm src} , \tag{494}$$

where

$$b^{\rm src}(<F_{\rm d}) \equiv g^3(x)\int_0^{F_{\rm d}} {\rm d}F F^3 \frac{{\rm d}n}{{\rm d}F} = g^3(x)\frac{\beta}{3-\beta} n(>F_{\rm d}) F_{\rm d}^3 . \tag{495}$$

Here, $\mathrm{d}n/\mathrm{d}F$ is the differential source count per unit solid angle, and $n(>F_{\rm d}) \equiv \int_{F_{\rm d}}^{\infty} \mathrm{d}F(\mathrm{d}n/\mathrm{d}F)$. We have assumed a power-law count, $\mathrm{d}n/\mathrm{d}F \propto F^{-\beta-1}$, for $\beta < 2$. The other symbols mean $x \equiv h\nu/k_{\rm B}T \simeq (\nu/56.80\,\mathrm{GHz})(T/2.726\,\mathrm{K})^{-1}$, and

$$g(x) \equiv 2\,\frac{(hc)^2}{(k_{\rm B}T)^3}\left(\frac{\sinh x/2}{x^2}\right)^2 \simeq \frac{1}{67.55\,\mathrm{MJy\,sr^{-1}}}\left(\frac{T}{2.726\,\mathrm{K}}\right)^{-3}\left(\frac{\sinh x/2}{x^2}\right)^2 . \tag{496}$$

Using the Poisson angular power-spectrum, $C^{\rm ps}$, given by

$$C^{\rm ps}(<F_{\rm d}) \equiv g^2(x)\int_0^{F_{\rm d}} \mathrm{d}F\,F^2\,\frac{\mathrm{d}n}{\mathrm{d}F} = g^2(x)\,\frac{\beta}{2-\beta}\,n(>F_{\rm d})F_{\rm d}^2 , \tag{497}$$

we can rewrite $b^{\rm src}$ into a different form,

$$b^{\rm src}(<F_{\rm d}) = \frac{(2-\beta)^{3/2}}{\beta^{1/2}(3-\beta)}\,[n(>F_{\rm d})]^{-1/2}[C^{\rm ps}(<F_{\rm d})]^{3/2} . \tag{498}$$

The authors of Ref. [279] have estimated $n(>F_{\rm d}) \sim 300\,\mathrm{sr^{-1}}$ for $F_{\rm d} \sim 0.2\,\mathrm{Jy}$ at 217 GHz. This $F_{\rm d}$ corresponds to 5σ detection threshold for the Planck experiment at 217 GHz. In Ref. [238] their estimation was extrapolated to 94 GHz, finding $n(>F_{\rm d}) \sim 7\,\mathrm{sr^{-1}}$ for $F_{\rm d} \sim 2\,\mathrm{Jy}$, which corresponds to the WMAP 5σ threshold. These values yield

$$C^{\rm ps}(90\,\mathrm{GHz}, <2\,\mathrm{Jy}) \sim 2\times 10^{-16} , \tag{499}$$

$$C^{\rm ps}(217\,\mathrm{GHz}, <0.2\,\mathrm{Jy}) \sim 1\times 10^{-17} . \tag{500}$$

Thus, rough estimates for $b^{\rm src}$ are

$$b^{\rm src}(90\,\mathrm{GHz}, <2\,\mathrm{Jy}) \sim 2\times 10^{-25} , \tag{501}$$

$$b^{\rm src}(217\,\mathrm{GHz}, <0.2\,\mathrm{Jy}) \sim 5\times 10^{-28} . \tag{502}$$

While we have assumed the Euclidean source count ($\beta = \frac{3}{2}$) for definiteness, this assumption does not affect order-of-magnitude estimates here.

As the primordial reduced bispectrum is $\propto l^{-4}$ (Eq. (477)), and the SZ-lensing reduced bispectrum is $\propto l^{-3}$ (Eq. (492)), the point-source bispectrum rapidly becomes to dominate the total bispectrum on small angular scales:

$$\frac{b_{lll}^{\rm prim}}{b^{\rm src}} \sim l^{-4}\times 10^7\left(\frac{f_{\rm NL}}{b^{\rm src}/10^{-25}}\right) , \tag{503}$$

$$\frac{b_{lll}^{\text{sz-lens}}}{b^{\rm src}} \sim l^{-3}\times 10^6\left(\frac{j_\nu \overline{T}_{\rho 0} b_{\rm gas}}{b^{\rm src}/10^{-25}}\right) . \tag{504}$$

For example, the point-sources overwhelm the SZ-lensing bispectrum measured by WMAP at $l \gtrsim 100$.

What do the SZ-lensing bispectrum and the point-source bispectrum look like? Fig. 14 plots the primordial, the SZ-lensing, and the point-source reduced bispectra for the equilateral configurations, $l \equiv l_1 = l_2 = l_3$. We have plotted $l^2(l+1)^2 b_{lll}/(2\pi)^2$. We find that these bispectra are very different from each other in shape on small angular scales. It thus suggests that we can separate these three contributions on the basis of shape difference. We study this point in the next section.

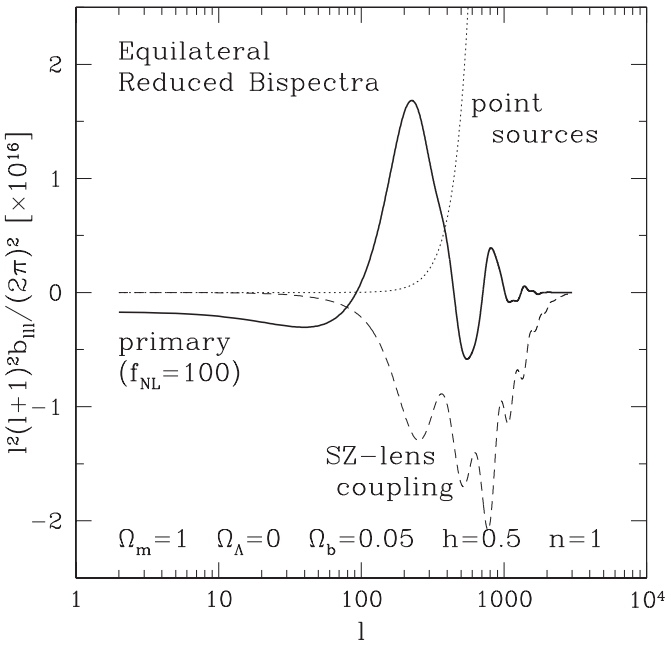

Fig. 14. Equilateral reduced bispectra. Comparison between the primordial (solid line), the SZ-lensing (dashed line), and the point-source (dotted line) reduced bispectra for the equilateral configurations, $l \equiv l_1 = l_2 = l_3$. We have plotted $[l^2(l+1)^2 b_{lll}/(2\pi)^2] \times 10^{16}$, which makes the Sachs–Wolfe plateau of the primordial reduced bispectrum on large angular scales, $l \lesssim 10$, easily seen.

10.4. Measuring bispectra: signal-to-noise estimation

In this section, we study how well we can measure the primordial bispectrum, and how well we can separate it from the secondary bispectra. Suppose that we fit the observed bispectrum, $B^{\rm obs}_{l_1 l_2 l_3}$, by theoretically calculated bispectra, which include both the primordial and secondary sources. We minimize χ^2 defined by

$$\chi^2 \equiv \sum_{2 \leqslant l_1 \leqslant l_2 \leqslant l_3} \frac{(B^{\rm obs}_{l_1 l_2 l_3} - \sum_i A_i B^{(i)}_{l_1 l_2 l_3})^2}{\sigma^2_{l_1 l_2 l_3}} \ , \tag{505}$$

where i denotes a component such as the primordial, the SZ and lensing effects, extragalactic sources, and so on. We have removed unobservable modes, $l = 0$ and 1.

As we have shown in Section 10.1, the variance of the bispectrum, $\sigma^2_{l_1 l_2 l_3}$, is the six-point function of a_{lm} [171,108]. When non-Gaussianity is weak, we calculate it as [268,88]

$$\sigma^2_{l_1 l_2 l_3} \equiv \langle B^2_{l_1 l_2 l_3} \rangle - \langle B_{l_1 l_2 l_3} \rangle^2 \approx \mathscr{C}_{l_1} \mathscr{C}_{l_2} \mathscr{C}_{l_3} \Delta_{l_1 l_2 l_3} \ , \tag{506}$$

where $\Delta_{l_1 l_2 l_3}$ takes values 1, 2, or 6 when all l's are different, two are the same, or all are the same, respectively. $\mathscr{C}_l \equiv C_l + C^{\rm N}_l$ is the total CMB angular power-spectrum, which includes the power-spectrum of the detector noise, $C^{\rm N}_l$. We calculate $C^{\rm N}_l$ analytically following [125] with the noise characteristics of relevant experiments. We do not include C_l from secondary sources, as they are subdominant compared

with the primordial C_l and $C_l^{\rm N}$ for relevant experiments. Including C_l from extragalactic sources (Eqs. (499) or (500)) changes our results by less than 10%.

Taking $\partial\chi^2/\partial A_i = 0$, we obtain the equation

$$\sum_j \left[\sum_{2\leqslant l_1\leqslant l_2\leqslant l_3} \frac{B^{(i)}_{l_1l_2l_3} B^{(j)}_{l_1l_2l_3}}{\sigma^2_{l_1l_2l_3}} \right] A_j = \sum_{2\leqslant l_1\leqslant l_2\leqslant l_3} \frac{B^{\rm obs}_{l_1l_2l_3} B^{(i)}_{l_1l_2l_3}}{\sigma^2_{l_1l_2l_3}} \ . \tag{507}$$

We then define the Fisher matrix, F_{ij}, as

$$\begin{aligned} F_{ij} &\equiv \sum_{2\leqslant l_1\leqslant l_2\leqslant l_3} \frac{B^{(i)}_{l_1l_2l_3} B^{(j)}_{l_1l_2l_3}}{\sigma^2_{l_1l_2l_3}} \\ &= \frac{2}{\pi} \sum_{2\leqslant l_1\leqslant l_2\leqslant l_3} \left(l_1+\frac{1}{2}\right)\left(l_2+\frac{1}{2}\right)\left(l_3+\frac{1}{2}\right) \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \frac{b^{(i)}_{l_1l_2l_3} b^{(j)}_{l_1l_2l_3}}{\sigma^2_{l_1l_2l_3}} \ , \end{aligned} \tag{508}$$

where we have used Eq. (459) to replace $B_{l_1l_2l_3}$ by the reduced bispectrum, $b_{l_1l_2l_3}$ (see Eq. (458) for definition). Since the covariance matrix of A_i is F^{-1}_{ij}, we define the signal-to-noise ratio, $(S/N)_i$, for a component i, the correlation coefficient, r_{ij}, between different components i and j, and the degradation parameter, d_i, of $(S/N)_i$ due to r_{ij}, as

$$\left(\frac{S}{N}\right)_i \equiv \frac{1}{\sqrt{F^{-1}_{ii}}} \ , \tag{509}$$

$$r_{ij} \equiv \frac{F^{-1}_{ij}}{\sqrt{F^{-1}_{ii} F^{-1}_{jj}}} \ , \tag{510}$$

$$d_i \equiv F_{ii} F^{-1}_{ii} \ . \tag{511}$$

Note that r_{ij} does not depend upon the amplitude of the bispectra, but on their shape. We have defined d_i so as $d_i = 1$ for zero degradation, while $d_i > 1$ for degraded $(S/N)_i$. In Refs. [268,70] the diagonal component of F^{-1}_{ij} has been considered. We study all the components to look at the separability between various bispectra.

We can give an order-of-magnitude estimate of S/N as a function of the angular resolution, l, as follows. Since the number of modes contributing to S/N increases as $l^{3/2}$, and $l^3 \begin{pmatrix} l & l & l \\ 0 & 0 & 0 \end{pmatrix}^2 \sim 0.36 \times l$, we estimate $(S/N)_i \sim (F_{ii})^{1/2}$ as

$$\left(\frac{S}{N}\right)_i \sim \frac{1}{3\pi} l^{3/2} \times l^{3/2} \left| \begin{pmatrix} l & l & l \\ 0 & 0 & 0 \end{pmatrix} \right| \times \frac{l^3 b^{(i)}_{lll}}{(l^2 C_l)^{3/2}} \sim l^5 b^{(i)}_{lll} \times 4 \times 10^{12} \ , \tag{512}$$

where we have used $l^2 C_l \sim 6 \times 10^{-10}$.

Table 2 tabulates F_{ij}, while Table 3 tabulates F^{-1}_{ij}; Table 4 tabulates $(S/N)_i$, while Table 5 tabulates d_i in the diagonal, and r_{ij} in the off-diagonal parts.

Table 2
Fisher matrix

COBE	Primordial	SZ-lensing	Point-sources
Primordial	$4.2\times10^{-6} f_{\rm NL}^2$	$-4.0\times10^{-7} f_{\rm NL} j_\nu \overline{T}_{\rho 0} b_{\rm gas}$	$-1.0\times10^{-9} f_{\rm NL} b_{25}^{\rm src}$
SZ-lensing		$1.3\times10^{-7}(j_\nu \overline{T}_{\rho 0} b_{\rm gas})^2$	$3.1\times10^{-10} j_\nu \overline{T}_{\rho 0} b_{\rm gas} b_{25}^{\rm src}$
Point-sources			$1.1\times10^{-12}(b_{25}^{\rm src})^2$
WMAP			
Primordial	$3.4\times10^{-3} f_{\rm NL}^2$	$2.6\times10^{-3} f_{\rm NL} j_\nu \overline{T}_{\rho 0} b_{\rm gas}$	$2.4\times10^{-3} f_{\rm NL} b_{25}^{\rm src}$
SZ-lensing		$0.14(j_\nu \overline{T}_{\rho 0} b_{\rm gas})^2$	$0.31 j_\nu \overline{T}_{\rho 0} b_{\rm gas} b_{25}^{\rm src}$
Point-sources			$5.6(b_{25}^{\rm src})^2$
Planck			
Primordial	$3.8\times10^{-2} f_{\rm NL}^2$	$7.2\times10^{-2} f_{\rm NL} j_\nu \overline{T}_{\rho 0} b_{\rm gas}$	$1.6\times10^{-2} f_{\rm NL} b_{27}^{\rm src}$
SZ-lensing		$39(j_\nu \overline{T}_{\rho 0} b_{\rm gas})^2$	$5.7 j_\nu \overline{T}_{\rho 0} b_{\rm gas} b_{27}^{\rm src}$
Point-sources			$2.7\times10^{3}(b_{27}^{\rm src})^2$

Fisher matrix, F_{ij} (see Eq. (508)): i denotes a component in the first row; j denotes a component in the first column. $\overline{T}_{\rho 0}$ is in units of 1 keV, $b_{25}^{\rm src} \equiv b^{\rm src}/10^{-25}$, and $b_{27}^{\rm src} \equiv b^{\rm src}/10^{-27}$.

Table 3
Inverted fisher matrix

COBE	Primordial	SZ-lensing	Point-sources
Primordial	$3.5\times10^{5} f_{\rm NL}^{-2}$	$1.1\times10^{6}(f_{\rm NL} j_\nu \overline{T}_{\rho 0} b_{\rm gas})^{-1}$	$1.3\times10^{7}(f_{\rm NL} b_{25}^{\rm src})^{-1}$
SZ-lensing		$3.1\times10^{7}(j_\nu \overline{T}_{\rho 0} b_{\rm gas})^{-2}$	$-7.8\times10^{9}(j_\nu \overline{T}_{\rho 0} b_{\rm gas} b_{25}^{\rm src})^{-1}$
Point sources			$3.1\times10^{12}(b_{25}^{\rm src})^{-2}$
WMAP			
Primordial	$3.0\times10^{2} f_{\rm NL}^{-2}$	$-6.1(f_{\rm NL} j_\nu \overline{T}_{\rho 0} b_{\rm gas})^{-1}$	$0.21(f_{\rm NL} b_{25}^{\rm src})^{-1}$
SZ-lensing		$8.4(j_\nu \overline{T}_{\rho 0} b_{\rm gas})^{-2}$	$-0.46(j_\nu \overline{T}_{\rho 0} b_{\rm gas} b_{25}^{\rm src})^{-1}$
Point-sources			$0.21(b_{25}^{\rm src})^{-2}$
Planck			
Primordial	$26 f_{\rm NL}^{-2}$	$-4.9\times10^{-2}(f_{\rm NL} j_\nu \overline{T}_{\rho 0} b_{\rm gas})^{-1}$	$-5.7\times10^{-5}(f_{\rm NL} b_{27}^{\rm src})^{-1}$
SZ-lensing		$2.6\times10^{-2}(j_\nu \overline{T}_{\rho 0} b_{\rm gas})^{-2}$	$-5.4\times10^{-5}(j_\nu \overline{T}_{\rho 0} b_{\rm gas} b_{27}^{\rm src})^{-1}$
Point-sources			$3.7\times10^{-4}(b_{27}^{\rm src})^{-2}$

Inverted fisher matrix, F_{ij}^{-1}. The meaning of the symbols is the same as in Table 2.

Table 4
Signal-to-noise ratio

	Primordial	SZ-lensing	Point-sources
COBE	$1.7\times10^{-3} f_{\rm NL}$	$1.8\times10^{-4} \lvert j_\nu \rvert \overline{T}_{\rho 0} b_{\rm gas}$	$5.7\times10^{-7} b_{25}^{\rm ps}$
WMAP	$5.8\times10^{-2} f_{\rm NL}$	$0.34 \lvert j_\nu \rvert \overline{T}_{\rho 0} b_{\rm gas}$	$2.2 b_{25}^{\rm ps}$
Planck	$0.19 f_{\rm NL}$	$6.2 \lvert j_\nu \rvert \overline{T}_{\rho 0} b_{\rm gas}$	$52 b_{27}^{\rm ps}$

Signal-to-noise ratio, $(S/N)_i$ (see Eq. (509)), of detecting the bispectrum. i denotes a component in the first row. The meaning of the symbols is the same as in Table 2.

Table 5
Signal degradation and correlation matrix

COBE	Primordial	SZ-lensing	Point-sources
Primordial	1.46	$0.33\,\mathrm{sgn}(j_\nu)$	1.6×10^{-2}
SZ-lensing		3.89	$-0.79\,\mathrm{sgn}(j_\nu)$
Point-sources			3.45
WMAP			
Primordial	1.01	$-0.12\,\mathrm{sgn}(j_\nu)$	2.7×10^{-2}
SZ-lensing		1.16	$-0.35\,\mathrm{sgn}(j_\nu)$
Point-sources			1.14
Planck			
Primordial	1.00	$-5.9\times10^{-2}\,\mathrm{sgn}(j_\nu)$	-5.8×10^{-4}
SZ-lensing		1.00	$-1.8\times10^{-2}\,\mathrm{sgn}(j_\nu)$
Point-sources			1.00

Signal degradation parameter, d_i (see Eq. (511)), and correlation coefficient, r_{ij} (see Eq. (510)), matrix. i denotes a component in the first row; j denotes a component in the first column. d_i for $i=j$, while r_{ij} for $i\neq j$.

10.4.1. Measuring the primordial bispectrum

Fig. 15 shows the signal-to-noise ratio, S/N. The top panel shows the differential S/N for the primordial bispectrum at $\ln l_3$ interval, $[\mathrm{d}(S/N)^2/\mathrm{d}\,\ln\,l_3]^{1/2} f_{\mathrm{NL}}^{-1}$, and the bottom panel shows the cumulative S/N, $(S/N)(<l_3) f_{\mathrm{NL}}^{-1}$, which is S/N summed up to a certain l_3. We have computed the detector noise power-spectrum, C_l^{N}, for COBE four-year map [36], WMAP 90 GHz channel, and Planck 217 GHz channel, and assumed full sky coverage. Fig. 15 also shows the ideal experiment with no noise: $C_l^{\mathrm{N}}=0$. Both $[\mathrm{d}(S/N)^2/\mathrm{d}\,\ln\,l_3]^{1/2}$ and $(S/N)(<l_3)$ increase monotonically with l_3, roughly $\propto l_3$, up to $l_3\sim 2000$ for the ideal experiment.

Beyond $l_3\sim 2000$, an enhancement of the damping tail in C_l because of the weak lensing effect [257] stops $[\mathrm{d}(S/N)^2/\mathrm{d}\,\ln\,l_3]^{1/2}$, and hence $(S/N)(<l_3)$, increasing. This leads to an important constraint on observations; even for the ideal noise-free, infinitesimally thin beam experiment, there is an upper limit on the value of $S/N\lesssim 0.3 f_{\mathrm{NL}}$. For a given realistic experiment, $[\mathrm{d}(S/N)^2/\mathrm{d}\,\ln\,l_3]^{1/2}$ has a maximum at a scale near the beam size.

For COBE, WMAP and Planck experiments, the total $(S/N) f_{\mathrm{NL}}^{-1}$ are 1.7×10^{-3}, 5.8×10^{-2}, and 0.19, respectively (see Table 4). To obtain $S/N>1$, we need $f_{\mathrm{NL}}>600$, 20, and 5, while the ideal experiment requires $f_{\mathrm{NL}}>3$ (see Table 6). We can also roughly obtain these values by substituting Eq. (477) into (512),

$$\left(\frac{S}{N}\right)_{\mathrm{prim}}\sim l\times10^{-4} f_{\mathrm{NL}}\ . \tag{513}$$

The degradation parameters, d_{prim}, are 1.46, 1.01, and 1.00 for COBE, WMAP, and Planck experiments, respectively (see Table 5), suggesting that WMAP and Planck experiments will separate the primordial bispectrum from the others with 1% or better accuracy; however, COBE cannot discriminate between them very well, as the primordial and the secondary sources change monotonically on the COBE angular scales. On the WMAP and Planck scales, the primordial bispectrum starts oscillating around zero, being well separated in shape from the secondaries that do not oscillate. This is good news for the forthcoming high angular resolution CMB experiments.

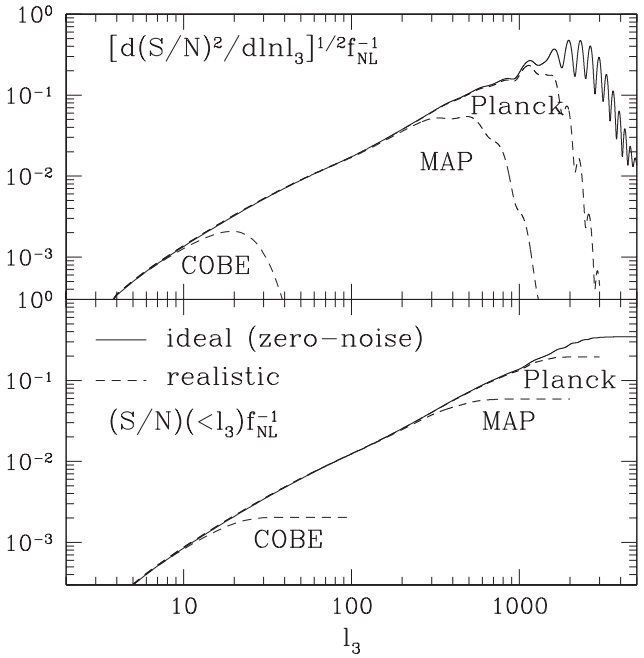

Fig. 15. Signal-to-noise ratio [143]. The predictions of the signal-to-noise ratio, S/N, for COBE, WMAP, and Planck experiments (see Eq. (509)). The differential S/N at $\ln l_3$ interval is shown in the upper panel, while the cumulative S/N up to a certain l_3 is shown in the bottom panel. Both are in units of $f_{\rm NL}$. Solid line represents the zero-noise ideal experiment, while dotted lines show the realistic experiments mentioned above. The total $(S/N) f_{\rm NL}^{-1}$ are 1.7×10^{-3}, 5.8×10^{-2}, and 0.19 for COBE, WMAP, and Planck experiments, respectively.

Table 6
Detection limit for the non-linearity parameter

Experiments	$f_{\rm NL}$ (Bispectrum)	$f_{\rm NL}$ (Skewness)
COBE	600	800
WMAP	20	80
Planck	5	70
Ideal	3	60

The minimum non-linearity parameter, $f_{\rm NL}$, needed for detecting the primordial non-Gaussianity by the bispectrum or the skewness with signal-to-noise ratio greater than 1. These estimates include the effects of cosmic variance, detector noise, and foreground sources.

10.4.2. Measuring secondary bispectra

Signal-to-noise ratios for detecting the SZ-lensing bispectrum, $(S/N)_{\text{sz-lens}}$, in units of $|j_\nu|\overline{T}_{\rho 0} b_{\rm gas}$ are 1.8×10^{-4}, 0.34, and 6.2 for COBE, WMAP, and Planck experiments, respectively (see Table 4), where $\overline{T}_{\rho 0}$ is in units of 1 keV. Using Eqs. (512) and (492), we can roughly estimate $(S/N)_{\text{sz-lens}}$ as

$$\left(\frac{S}{N}\right)_{\text{sz-lens}} \sim l^2 \times 10^{-6} |j_\nu| \overline{T}_{\rho 0} b_{\rm gas} \ . \tag{514}$$

Hence, $(S/N)_{\text{sz-lens}}$ increases with the angular resolution more rapidly than the primordial bispectrum (see Eq. (513)). Since $|j_\nu|\overline{T}_{\rho 0} b_{\text{gas}}$ should be of order unity, COBE and WMAP cannot detect the SZ-lensing bispectrum; however, Planck is sensitive enough to detect, depending on the frequency, i.e., a value of j_ν. For example, 217 GHz is insensitive to the SZ effect as $j_\nu \sim 0$, while $j_\nu = -2$ in the Rayleigh–Jeans regime.

The degradation parameters, $d_{\text{sz-lens}}$, are 3.89, 1.16, and 1.00 for COBE, WMAP, and Planck experiments, respectively (see Table 5); thus, Planck will separate the SZ-lensing bispectrum from the other effects. Note that the $(S/N)_{\text{sz-lens}}$ values must be understood as order-of-magnitude estimates, since our cosmological model is the COBE normalized SCDM that yields $\sigma_8 = 1.2$, which is a factor of 2 greater than the cluster normalization for $\Omega_{\text{m}} = 1$, and 20% greater than the normalization for $\Omega_{\text{m}} = 0.3$ [124]. Hence, this factor tends to overestimate $\langle \Theta^*_{lm} a^{\text{SZ}}_{lm} \rangle$ (Eq. (491)) by a factor of less than 10; on the other hand, using the linear $P_\Phi(k)$ power-spectrum rather than the non-linear power-spectrum tends to underestimate the effect by a factor of less than 10 at $l \sim 3000$ [70]. Yet, our main goal is to discriminate between the shapes of various bispectra, not to determine the amplitude, so that this factor does not affect our conclusion on the degradation parameters, d_i.

For the extragalactic radio and infrared sources, one can estimate the signal-to-noise ratios as $5.7 \times 10^{-7}(b^{\text{src}}/10^{-25})$, $2.2(b^{\text{src}}/10^{-25})$, and $52(b^{\text{src}}/10^{-27})$ for COBE, WMAP, and Planck experiments, respectively (see Table 4), and the degradation parameters, d_{ps}, as 3.45, 1.14, and 1.00 (see Table 5). This estimate is consistent with that of Ref. [238]. From Eq. (512), we find

$$\left(\frac{S}{N}\right)_{\text{ps}} \sim l^5 \times 10^{-13} \left(\frac{b^{\text{src}}}{10^{-25}}\right) \ ; \tag{515}$$

thus, S/N of the point-source bispectrum increases very rapidly with the angular resolution.

Although WMAP cannot separate the Poisson bispectrum from the SZ-lensing bispectrum very well (see r_{ij} in Table 5), the SZ-lensing bispectrum is too small to be measured by WMAP anyway. Planck will do an excellent job on separating all kinds of bispectra, at least including the primordial signal, SZ-lensing coupling, and extragalactic point-sources, on the basis of the shape difference.

10.4.3. Measuring primordial skewness

For the skewness, we define S/N as

$$\left(\frac{S}{N}\right)^2 \equiv \frac{S_3^2}{\sigma^2_{S_3}} \ , \tag{516}$$

where the variance is [271]

$$\begin{aligned}
\sigma^2_{S_3} &\equiv \langle (S_3)^2 \rangle = 6 \int_{-1}^{1} \frac{\mathrm{d}\cos\theta}{2} [\mathscr{C}(\theta)]^3 \\
&= 6 \sum_{l_1 l_2 l_3} \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{(4\pi)^3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \mathscr{C}_{l_1} \mathscr{C}_{l_2} \mathscr{C}_{l_3} W^2_{l_1} W^2_{l_2} W^2_{l_3} \\
&= \frac{9}{2\pi^3} \sum_{2 \leqslant l_1 \leqslant l_2 \leqslant l_3} \left(l_1 + \frac{1}{2}\right)\left(l_2 + \frac{1}{2}\right)\left(l_3 + \frac{1}{2}\right) \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \mathscr{C}_{l_1} \mathscr{C}_{l_2} \mathscr{C}_{l_3} W^2_{l_1} W^2_{l_2} W^2_{l_3} \ .
\end{aligned} \tag{517}$$

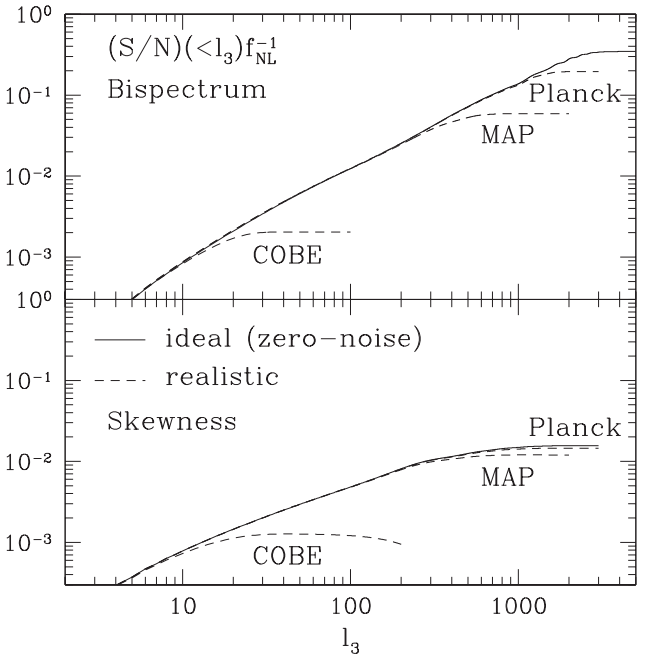

Fig. 16. Bispectrum vs. skewness [143]. Comparison of the signal-to-noise ratio summed up to a certain l_3, $S/N(<l_3)$, for the bispectrum (top panel; Eq. (509)) and the skewness (bottom panel; Eq. (516)). $S/N(<l_3)$ is in units of $f_{\rm NL}$. The dotted lines show COBE, WMAP, and Planck experiments (dotted lines), while the solid line shows the ideal experiment. See Table 6 for $f_{\rm NL}$ to obtain $S/N>1$.

In the last equality, we have used symmetry of the summed quantity with respect to indices (Eq. (483)), and removed unobservable modes, $l=0$ and 1. Typically $\sigma_{S_3}\sim 10^{-15}$, as $\sigma_{S_3}\sim[\mathscr{C}(0)]^{3/2}\sim 10^{-15}$, where $\mathscr{C}(\theta)$ is the temperature auto-correlation function including noise.

The bottom panel of Fig. 12 plots $\sigma_{S_3}(<l_3)$, which is σ_{S_3} summed up to a certain l_3, for COBE, WMAP, and Planck experiments as well as for the ideal experiment. Since $\mathscr{C}_l W_l^2 = C_l e^{-l(l+1)\sigma_b^2} + w^{-1}$, where w^{-1} is the white-noise power-spectrum of the detector noise [125], w^{-1} keeps $\sigma_{S_3}(<l_3)$ slightly increasing with l_3 beyond the experimental angular resolution scale, $l\sim\sigma_b^{-1}$. In contrast, $S_3(<l_3)$ becomes constant beyond $l\sim\sigma_b^{-1}$ (see the top panel of Fig. 12). As a result, S/N starts slightly decreasing beyond the resolution. We use the maximum S/N for calculating the minimum value of $f_{\rm NL}$ above which the primordial S_3 is detectable; we find that $f_{\rm NL}>800$, 80, 70, and 60 for COBE, WMAP, Planck, and the ideal experiments, respectively, assuming full sky coverage.

These $f_{\rm NL}$ values are systematically larger than those for detecting $B_{l_1l_2l_3}$ by a factor of 1.3, 4, 14, and 20, respectively (see Table 6). The higher the angular resolution is, the less sensitive the primordial S_3 is to non-Gaussianity than $B_{l_1l_2l_3}$. This is because of the cancellation effect on smaller angular scales caused by the oscillation of $B_{l_1l_2l_3}$ damps S_3.

Fig. 16 compares the expected signal-to-noise ratio of detecting the primordial non-Gaussianity based on the bispectrum (Eq. (509)) with that based on the skewness (Eq. (516)). It shows that the bispectrum is almost an order of magnitude more sensitive to the non-Gaussianity than the skewness. We conclude that

when we can compute the predicted form of the bispectrum, it becomes a "matched filter" for detecting the non-Gaussianity in data, and thus much more a powerful tool than the skewness. Table 6 summarizes the minimum $f_{\rm NL}$ for detecting the primordial non-Gaussianity using the bispectrum or the skewness for COBE, WMAP, Planck, and the ideal experiments. This shows that even the ideal experiment needs $f_{\rm NL} > 3$ to detect the primordial bispectrum.

10.5. Measuring primordial non-Gaussianity in the cosmic microwave background

Measuring $f_{\rm NL}$ from nearly full-sky experiments is challenging. The bispectrum analysis explained in Section 10.1 requires $N^{5/2}$ operations ($N^{3/2}$ for computing three l's and N for averaging over the sky) where N is the number of pixels. The brute-force analysis is possible for the COBE data for which $N \sim 3000$ [145], while it is quite challenging for mega-pixel experiments (e.g., $N \sim 3 \times 10^6$ for WMAP, 5×10^7 for Planck). In fact, just measuring all configurations of the bispectrum from the data is possible. What is challenging is to carry out many Monte Carlo simulations: in order to quantify the statistical significance of the measurements, one needs many simulations. It is the simulations that are computationally very expensive. Since the brute-force trispectrum analysis requires N^3, it is even more challenging.

Although we measure the individual triangle configurations of the bispectrum (or quadrilateral configurations of the trispectrum) at first, we eventually combine all of them to constrain model parameters such as $f_{\rm NL}$, as the signal-to-noise per configuration is nearly zero. This may sound inefficient. Measuring all configurations is enormously time consuming. Is there any statistic which *already* combines all the configurations optimally, and fast to compute? Yes [144]. A physical justification for our methodology is as follows. A model like Eq. (466) generates non-Gaussianity in real space, and the Central-Limit Theorem makes the Fourier modes nearly Gaussian; thus, real-space statistics should be more sensitive. On the other hand, real-space statistics are weighted sum of Fourier-space statistics, which are often easier to predict. Therefore, we need to understand the shape of Fourier-space statistics to find sensitive real-space statistics, and for this purpose it is useful to have a specific, physically motivated non-Gaussian model, compute Fourier statistics, and find optimal real-space statistics.

10.5.1. Reconstructing primordial fluctuations from temperature anisotropy

We begin with the primordial curvature perturbations $\Phi(\mathbf{x})$ and isocurvature perturbations $S(\mathbf{x})$. If we can reconstruct these primordial fluctuations from the observed CMB anisotropy, $\Delta T(\hat{\mathbf{n}})/T$, then we can improve the sensitivity to primordial non-Gaussianity. We find that the harmonic coefficients of the CMB anisotropy, $a_{lm} = T^{-1}\int \mathrm{d}^2\hat{\mathbf{n}}\Delta T(\hat{\mathbf{n}})Y^*_{lm}(\hat{\mathbf{n}})$, are related to the primordial fluctuations as

$$a_{lm} = W_l \int r^2\,\mathrm{d}r[\Phi_{lm}(r)\alpha_l^{\rm adi}(r) + S_{lm}(r)\alpha_l^{\rm iso}(r)] + n_{lm}\ , \tag{518}$$

where $\Phi_{lm}(r)$ and $S_{lm}(r)$ are the harmonic coefficients of the fluctuations at a given comoving distance, $r = |\mathbf{x}|$ from the observer. A beam function W_l and the harmonic coefficients of the noise n_{lm} represent instrumental effects. Since noise can be spatially inhomogeneous, the noise covariance matrix $\langle n_{lm}n^*_{l'm'}\rangle$ can be non-diagonal; however, we approximate it with $\simeq \sigma_0^2\delta_{ll'}\delta_{mm'}$. We thus assume the "mildly inhomogeneous" noise for which this approximation holds. The function $\alpha_l(r)$ is defined by

$$\alpha_l(r) \equiv \frac{2}{\pi}\int k^2\,\mathrm{d}k g_{{\rm T}l}(k)\,j_l(kr)\ , \tag{519}$$

where $g_{Tl}(k)$ is the radiation transfer function of either adiabatic (adi) or isocurvature (iso) perturbations. Note that this function is equal to $f_{\rm NL}^{-1} b_l^{\rm NL}(r)$ (see Eq. (474)).

Next, assuming that $\Phi(\mathbf{x})$ dominates, we try to reconstruct $\Phi(\mathbf{x})$ from the observed $\Delta T(\hat{\mathbf{n}})$. A linear filter, $\mathcal{O}_l(r)$, which reconstructs the underlying field, can be obtained by minimizing the variance of the difference between the filtered field $\mathcal{O}_l(r)a_{lm}$ and the underlying field $\Phi_{lm}(r)$. By evaluating

$$\frac{\partial}{\partial \mathcal{O}_l(r)} \langle |\mathcal{O}_l(r)a_{lm} - \Phi_{lm}(r)|^2 \rangle = 0 \ , \tag{520}$$

one obtains a solution for the filter as

$$\mathcal{O}_l(r) = \frac{\beta_l(r) W_l}{\tilde{C}_l} \ , \tag{521}$$

where the function $\beta_l(r)$ is given by

$$\beta_l(r) \equiv \frac{2}{\pi} \int k^2 \, {\rm d}k \, P(k) g_{Tl}(k) j_l(kr) \ , \tag{522}$$

and $P(k)$ is the power–spectrum of Φ. Of course, one can replace Φ with S when S dominates. This function is equal to $b_l^{\rm L}(r)$ (see Eq. (475)). Here, we put a tilde on a quantity that includes effects of W_l and noise such that $\tilde{C}_l \equiv C_l W_l^2 + \sigma_0^2$, where C_l is the theoretical power-spectrum that uses the same cosmological model as $g_{Tl}(k)$.

Finally, we transform the filtered field $\mathcal{O}_l(r)a_{lm}$ back to pixel space to obtain an Wiener-filtered, reconstructed map of $\Phi(r, \hat{\mathbf{n}})$ or $S(r, \hat{\mathbf{n}})$. We have assumed that there is no correlation between Φ and S. We will return to study the case of non-zero correlation later (Section 10.5.3).

Fig. 17 shows $\mathcal{O}_l(r)$ as a function of l and r for (a) an adiabatic SCDM ($\Omega_{\rm m} = 1$), (b) an adiabatic ΛCDM ($\Omega_{\rm m} = 0.3$), (c) an isocurvature SCDM, and (d) an isocurvature ΛCDM. While we have used $P(k) \propto k^{-3}$ for both adiabatic and isocurvature modes, the specific choice of $P(k)$ does not affect $\mathcal{O}_l$ very much as $P(k)$ in β_l in the numerator approximately cancels out $P(k)$ in C_l in the denominator. On large angular scales (smaller l) the Sachs–Wolfe (SW) effect makes $\mathcal{O}_l$ equal to -3 for adiabatic modes and $-\frac{5}{2}$ for isocurvature modes in SCDM. For the ΛCDM models the late-time decay of the gravitational potential makes this limit different. Adiabatic and isocurvature modes are out of phase in l.

The figure shows that $\mathcal{O}_l$ changes the sign of the fluctuations as a function of scales. This indicates that acoustic physics at the last scattering surface modulates fluctuations so that hot spots in the primordial fluctuations can be cold spots in the CMB, for example. Therefore, the shape of $\mathcal{O}_l$ "deconvolves" the sign change, recovering the phases of fluctuations. This is an intuitive reason why our cubic statistic derived below (Eq. (525)) works, and it proves more advantageous to measure primordial non-Gaussianity on a filtered map than on a temperature map.

This property should be compared to that of real-space statistics measured on a temperature map. As we have shown in Section 10.4 the skewness of a temperature map is much less sensitive to the primordial non-Gaussianity than the bispectrum, exactly because of the cancellation effect from the acoustic oscillations. The skewness of a filtered map, on the other hand, has a larger signal-to-noise ratio, and more optimal statistics like our cubic statistic derived below can be constructed. Other real-space statistics such as Minkowski functionals [194,96,255,57], peak–peak correlations [109] may also be more sensitive to the primordial non-Gaussianity, when measured on the filtered maps.

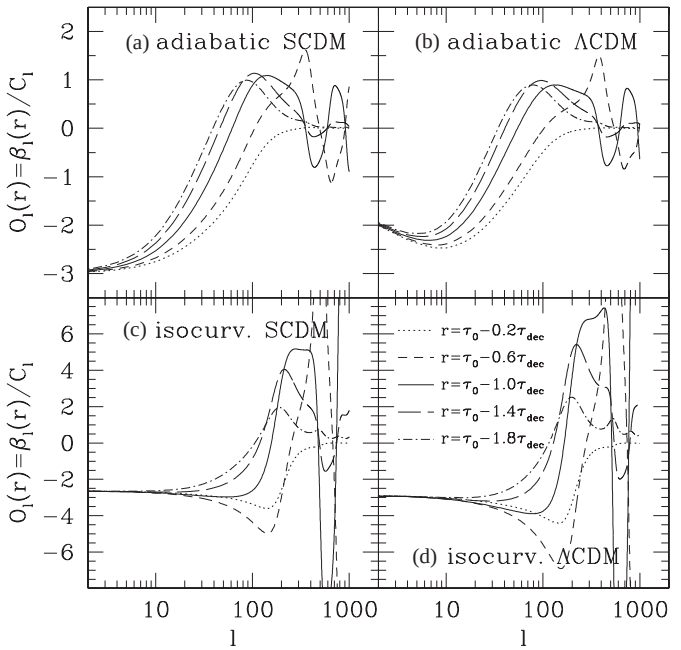

Fig. 17. Wiener filters for the primordial fluctuations [144]. Wiener filters, $\mathcal{O}_l(r) = \beta_l(r)/C_l$ (Eq. (521)). We plot (a) $\mathcal{O}_l$ for an adiabatic SCDM ($\Omega_{\rm m} = 1$, $\Omega_\Lambda = 0$, $\Omega_{\rm b} = 0.05$, $h = 0.5$), (b) for an adiabatic ΛCDM ($\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_{\rm b} = 0.04$, $h = 0.7$), (c) for an isocurvature SCDM, and (d) for an isocurvature ΛCDM. The filters are plotted at five conformal distances $r = c(\tau_0 - \tau)$ as explained in the bottom-right panel. Here τ is the conformal time (τ_0 at the present). The SCDM models have $c\tau_0 = 11.84$ Gpc and $c\tau_{\rm dec} = 0.235$ Gpc, while the ΛCDM models $c\tau_0 = 13.89$ Gpc and $c\tau_{\rm dec} = 0.277$ Gpc, where $\tau_{\rm dec}$ is the photon decoupling epoch.

Unfortunately, as g_{Tl} oscillates, our reconstruction of Φ or S from a temperature map alone is not perfect. While $\mathcal{O}_l$ reconstructs the primordial fluctuations very well on large scales via the Sachs–Wolfe effect, $\mathcal{O}_l \sim 0$ on intermediate scales ($l \sim 50$ for adiabatic and $l \sim 100$ for isocurvature), indicating loss of information on the phases of the underlying fluctuations. Then, toward smaller scales, we recover information, lose information, and so on. Exact scales at which $\mathcal{O}_l \sim 0$ depend on r and cosmology. A good news is that a high signal-to-noise map of the CMB polarization anisotropy will enable us to overcome the loss of information, as the polarization transfer function is out of phase in l compared to the temperature transfer function, filling up information at which $\mathcal{O}_l \sim 0$. In other words, the polarization anisotropy has finite information about the phases of the primordial perturbations, when the temperature anisotropy has zero information.

10.5.2. Measuring primordial non-Gaussianity in adiabatic fluctuations

Using two functions introduced in the previous section, we construct a *cubic* statistic which is optimal for the primordial non-Gaussianity. We apply filters to a_{lm}, and then transform the filtered a_{lm}'s to obtain two maps, A and B, given by

$$A(r, \hat{\mathbf{n}}) \equiv \sum_{lm} \frac{\alpha_l(r) W_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{\mathbf{n}}) \ , \tag{523}$$

$$B(r,\hat{\mathbf{n}}) \equiv \sum_{lm} \frac{\beta_l(r) W_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{\mathbf{n}}) \ . \tag{524}$$

The latter map, $B(r,\hat{\mathbf{n}})$, is exactly the $\mathcal{O}_l$-filtered map, a Wiener-filtered map of the underlying primordial fluctuations. We then form a cubic statistic given by

$$\mathcal{S}_{\rm prim} \equiv 4\pi \int r^2 \, {\rm d}r \int \frac{{\rm d}^2\hat{\mathbf{n}}}{4\pi} A(r,\hat{\mathbf{n}}) B^2(r,\hat{\mathbf{n}}) \ , \tag{525}$$

where the angular average is done on the full sky, regardless of the sky cut. We find that $\mathcal{S}_{\rm prim}$ reduces *exactly* to

$$\mathcal{S}_{\rm prim} = \sum_{l_1 \leqslant l_2 \leqslant l_3} \frac{\tilde{B}^{\rm obs}_{l_1 l_2 l_3} \tilde{B}^{\rm prim}_{l_1 l_2 l_3}}{\tilde{C}_{l_1} \tilde{C}_{l_2} \tilde{C}_{l_3}} \ , \tag{526}$$

where

$$\tilde{B}_{l_1 l_2 l_3} \equiv B_{l_1 l_2 l_3} W_{l_1} W_{l_2} W_{l_3} \ , \tag{527}$$

and $B^{\rm obs}_{l_1 l_2 l_3}$ is the observed bispectrum with the effect of W_l corrected while $B^{\rm prim}_{l_1 l_2 l_3}$ is given by Eqs. (476) and (459).

The denominator of Eq. (526) is the variance of $\tilde{B}^{\rm obs}_{l_1 l_2 l_3}$ in the limit of weak non-Gaussianity (say $|f_{\rm NL}| \lesssim 10^3$) when all l's are different: $\langle \tilde{B}^2_{l_1 l_2 l_3} \rangle = \tilde{C}_{l_1} \tilde{C}_{l_2} \tilde{C}_{l_3} \Delta_{l_1 l_2 l_3}$, where $\Delta_{l_1 l_2 l_3}$ is 6 for $l_1 = l_2 = l_3$, 2 for $l_1 = l_2 \neq l_3$, etc., and 1 otherwise. The bispectrum configurations are thus summed up nearly optimally with the approximate inverse-variance weights, provided that $\Delta_{l_1 l_2 l_3}$ is approximated with $\simeq 1$. The least-square fit of $\tilde{B}^{\rm prim}_{l_1 l_2 l_3}$ to $\tilde{B}^{\rm obs}_{l_1 l_2 l_3}$ can be performed to yield

$$\mathcal{S}_{\rm prim} \simeq f_{\rm NL} \sum_{l_1 \leqslant l_2 \leqslant l_3} \frac{(\tilde{B}^{\rm prim}_{l_1 l_2 l_3})^2}{\tilde{C}_{l_1} \tilde{C}_{l_2} \tilde{C}_{l_3}} \ . \tag{528}$$

This equation gives an estimate of $f_{\rm NL}$ directly from $\mathcal{S}_{\rm prim}$.

The most time-consuming part is the back-and-forth harmonic transform necessary for pre-filtering (see Eqs. (523) and (524)), taking $N^{3/2}$ operations times the number of sampling points of r, of order 100, for evaluating the integral (Eq. (525)). This is much faster than the full bispectrum analysis which takes $N^{5/2}$, enabling us to perform a more detailed analysis of the data in a reasonable amount of computational time. For example, measurements of all bispectrum configurations up to $l_{\rm max} = 512$ take 8 h to compute on 16 processors of an SGI Origin 300; thus, even only 100 Monte Carlo simulations take 1 month to be carried out. On the other hand, $\mathcal{S}_{\rm prim}$ takes only 30 s to compute, 1000 times faster. When we measure $f_{\rm NL}$ for $l_{\rm max} = 1024$, we speed up by a factor of 4000: 11 days for the bispectrum vs. 4 min for $\mathcal{S}_{\rm prim}$. We can do 1000 simulations for $l_{\rm max} = 1024$ in 3 days.

10.5.3. Mixed fluctuations

The $\mathcal{O}_l$-filtered map, B, is an Wiener-filtered map of primordial curvature or isocurvature perturbations; however, this is correct only when correlations between the two components are negligible. On the other hand, multi-field inflation models and curvaton models naturally predict correlations. The current CMB

data are consistent with, but do not require, a correlated mixture of these fluctuation modes [12,282,225]. In this case, the Wiener filter for the primordial fluctuations (Eq. (521)) needs to be modified such that $\mathcal{O}_l(r) = \beta_l(r) W_l/\tilde{C}_l \to \tilde{\beta}_l(r) W_l/\tilde{C}_l$, where

$$\tilde{\beta}_l^{\rm adi}(r) = \frac{2}{\pi}\int k^2\,{\rm d}k[P_\Phi(k) g_{Tl}^{\rm adi}(k) + P_C(k) g_{Tl}^{\rm iso}(k)] j_l(kr)\ ,$$
$$\tilde{\beta}_l^{\rm iso}(r) = \frac{2}{\pi}\int k^2\,{\rm d}k[P_S(k) g_{Tl}^{\rm iso}(k) + P_C(k) g_{Tl}^{\rm adi}(k)] j_l(kr)\ ,$$

for curvature (adi) and isocurvature (iso) perturbations, respectively. Here P_Φ is the primordial power-spectrum of curvature perturbations, P_S of isocurvature perturbations, and P_C of cross correlations.

For measuring non-Gaussianity from the correlated fluctuations, we use Eq. (466) as a model for Φ- and S-field non-Gaussianity to parameterize them with $f_{\rm NL}^{\rm adi}$ and $f_{\rm NL}^{\rm iso}$, respectively. We then form a cubic statistic similar to $\mathcal{S}_{\rm prim}$ (Eq. (525)), using $A(r,\hat{\mathbf{n}})$ and a new filtered map $\tilde{B}(r,\hat{\mathbf{n}})$ which uses $\tilde{\beta}_l(r)$. We have two cubic combinations: $A_{\rm adi}\tilde{B}_{\rm adi}^2$ for measuring $f_{\rm NL}^{\rm adi}$ and $A_{\rm iso}\tilde{B}_{\rm iso}^2$ for $f_{\rm NL}^{\rm iso}$, each of which comprises four terms including one P_Φ^2 (or P_S^2), one P_C^2, and two $P_\Phi P_C$'s (or $P_S P_C$'s). In other words, the correlated contribution makes the total number of terms contributing to the non-Gaussianity four times more than the uncorrelated-fluctuation models (see Ref. [30] for more generic cases).

10.5.4. Point-source non-Gaussianity

Next, we show that the filtering method is also useful for measuring foreground non-Gaussianity arising from extragalactic point-sources. The residual point-sources left unsubtracted in a map can seriously contaminate both the power-spectrum and the bispectrum. We can, on the other hand, use multi-band observations as well as external template maps of dust, free-free, and synchrotron emission, to remove diffuse Galactic foreground [39]. The radio sources with known positions can be safely masked.

The filtered map for the point-sources is

$$D(\hat{\mathbf{n}}) \equiv \sum_{lm}\frac{W_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{\mathbf{n}})\ . \tag{529}$$

This filtered map was actually used for detecting point-sources in the WMAP maps [39]. Using $D(\hat{\mathbf{n}})$, the cubic statistic is derived as

$$\mathcal{S}_{\rm src} \equiv \int \frac{{\rm d}^2\hat{\mathbf{n}}}{4\pi} D^3(\hat{\mathbf{n}}) = \frac{3}{2\pi}\sum_{l_1\leqslant l_2\leqslant l_3}\frac{\tilde{B}_{l_1l_2l_3}^{\rm obs}\tilde{B}_{l_1l_2l_3}^{\rm src}}{\tilde{C}_{l_1}\tilde{C}_{l_2}\tilde{C}_{l_3}}\ . \tag{530}$$

Here, $B_{l_1l_2l_3}^{\rm src}$ is the point-source bispectrum for unit white-noise bispectrum (i.e., $b^{\rm src}=1$ in Eq. (494)). When the covariance between $B_{l_1l_2l_3}^{\rm prim}$ and $B_{l_1l_2l_3}^{\rm src}$ is negligible as is the case for WMAP and Planck (see Table 5), we find

$$\mathcal{S}_{\rm src} \simeq \frac{3b^{\rm src}}{2\pi}\sum_{l_1\leqslant l_2\leqslant l_3}\frac{(\tilde{B}_{l_1l_2l_3}^{\rm src})^2}{\tilde{C}_{l_1}\tilde{C}_{l_2}\tilde{C}_{l_3}}\ . \tag{531}$$

We omit the covariance only for simplicity; however, including it would be simple [145].

Again, $\mathcal{S}_{\rm src}$ measures $b^{\rm src}$ much faster than the full bispectrum analysis, constraining effects of residual point-sources on CMB sky maps. Since $\mathcal{S}_{\rm src}$ does not contain the extra integral over r, it is even

100 times faster to compute than $\mathscr{S}_{\rm prim}$. This statistic is particularly useful because it is sometimes difficult to tell how much of C_l is due to point-sources. In Section 10.6 we see how $\mathscr{S}_{\rm src}$ (i.e., $b^{\rm src}$) is related to C_l due to the unsubtracted point-sources.

10.5.5. Incomplete sky coverage

Finally, we show how to incorporate incomplete sky coverage and pixel weights into our statistics. Suppose that we weight a sky map by $M(\hat{\mathbf{n}})$ to measure the harmonic coefficients,

$$a_{lm}^{\rm obs} = \frac{1}{T}\int {\rm d}^2\hat{\mathbf{n}} M(\hat{\mathbf{n}})\Delta T(\hat{\mathbf{n}}) Y_{lm}^*(\hat{\mathbf{n}}) \ . \tag{532}$$

A full-sky a_{lm} is related to $a_{lm}^{\rm obs}$ through the coupling matrix $M_{ll'mm'} \equiv \int {\rm d}^2\hat{\mathbf{n}} M(\hat{\mathbf{n}}) Y_{lm}^*(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}})$ by $a_{lm}^{\rm obs} = \sum_{l'm'} a_{l'm'} M_{ll'mm'}$. In this case the observed bispectrum is biased by a factor of $\int {\rm d}^2\hat{\mathbf{n}} M^3(\hat{\mathbf{n}})/(4\pi)$; thus, we need to divide $S_{\rm prim}$ and S_{ps} by this factor. If only the sky cut is considered, then this factor is the fraction of the sky covered by observations (see Eq. (452)).

Monte Carlo simulations of non-Gaussian sky maps computed with Eq. (518) (see Appendix A of Ref. [139]) show that $\mathscr{S}_{\rm prim}$ reproduces the input $f_{\rm NL}$'s accurately both on full sky and incomplete sky with modest Galactic cut and inhomogeneous noise on the WMAP data, i.e., the statistic is unbiased. The error on $f_{\rm NL}$ from $\mathscr{S}_{\rm prim}$ is as small as that from the full bispectrum analysis; however, one cannot make a sky cut very large, e.g., more than 50% of the sky, as for it the covariance matrix of $\tilde{B}_{l_1l_2l_3}$ is no longer diagonal. The cubic statistic does not include the off-diagonal terms of the covariance matrix [see Eq. (526)]; however, it works fine for WMAP sky maps for which one can use more than 75% of the sky. Also, Eq. (531) correctly estimates $b^{\rm src}$ using simulated realizations of point-sources (see Appendix B of Ref. [139]).

These fast methods allow to carry out extensive Monte Carlo simulations characterizing the effects of realistic noise properties of the experiments, sky cut, foreground sources, and so on. A reconstructed map of the primordial fluctuations, which plays a key role in the method, potentially gives other real-space statistics more sensitivity to primordial non-Gaussianity. As it has been shown, the method can be applied to the primordial non-Gaussianity arising from inflation, gravity, or correlated isocurvature fluctuations, as well as the foreground non-Gaussianity from radio point-sources, all of which can be important sources of non-Gaussian fluctuations on the CMB sky maps.

10.6. Applications to observational data

There are two approaches to testing Gaussianity of the CMB.

- Blind tests (null tests) which make no assumption about the form of non-Gaussianity. The simplest test would be measurements of deviation of one-point PDF from a Gaussian distribution. (Measurements of the skewness, kurtosis, etc., for example.) Being model-independent is a merit of this approach, while the statistical power is weak. If we had no models to test, this approach would be the only choice.
- Testing specific models of non-Gaussianity, constraining the model parameters. This approach is powerful in putting *quantitative* constraints on non-Gaussianity, at the cost of being model-dependent. If we had a sensible (yet fairly generic) model to test, this approach would be more powerful than the blind tests.

Both approaches have been applied to the CMB data on large angular scales ($\sim 7^\circ$) [134,135,108,255,84,218,51,19,67,208,179,214,251,25,227,145,137,148,7,59], on intermediate scales ($\sim 1^\circ$) [220,263], and on small scales ($\sim 10'$) [294,252,232]. So far, there is no compelling evidence for the cosmological non-Gaussianity, and the pre-WMAP constraint on $f_{\rm NL}$ was weak, $f_{\rm NL}\lesssim$(2000–3000) at 95% confidence level [145,59,252].

In this section, we briefly review results of Gaussianity tests on the WMAP data presented in Ref. [139]. The WMAP, Wilkinson Microwave Anisotropy Probe, has recently produced clean and precise sky maps of the CMB in 5 microwave bands [38], with the angular resolution 30 times better than that of the differential microwave radiometer (DMR) aboard the COBE satellite [36]. Detailed study of these sky maps offers a fundamental test of cosmology, as various cosmological effects change temperature and energy distribution of the CMB at all angular scales (e.g., [115]). The temperature and polarization power-spectra of the WMAP data [110,136] have determined the best-fit cosmological model with errors in the parameter determinations being quite small ($< 10\%$) [269,216]. The systematic errors in the parameter determinations are minimized by both the careful instrumental design [119,217,24,111] and data analysis techniques [286].

Apart from the CMB, there are a number of non-cosmological, "foreground" sources in the microwave sky. The emission from our Galaxy is the brightest component, which must be masked or subtracted out before any cosmological analysis of the CMB. Since the WMAP observes in five frequency bands, much of the Galactic emission can be reliably subtracted using the non-monochromatic nature of the Galaxy [39]. The power-spectra measured in different bands coincide with each other after the foreground subtraction, which is reassuring [110]. Actually, much more problematic a foreground component is the extragalactic radio sources. Although we can mask those positions of the sky which are known to have sources brighter than some threshold flux (which is determined by the sensitivity of observations), there always remain undetected sources. The undetected (unmasked) sources potentially contaminate the cosmological CMB signals. Since we cannot subtract them out individually, we must estimate the effect of the sources in a statistical manner.

The emission from the sources is highly non-Gaussian and only important on small angular scales; thus, we can use the non-Gaussian signals to directly estimate the source contribution. This example illustrates usefulness of the higher-order statistics in a real life.

10.6.1. Minkowski functionals

For the first test, one can use (but is not limited to) the Minkowski functionals [194,96,255,57], which measure morphological structures of the CMB, describing the properties of regions spatially bounded by a set of contours. The contours may be specified in terms of fixed temperature thresholds, $\nu = \Delta T/\sigma$, where σ is the standard deviation of the map, or in terms of the area. The three Minkowski functionals are: (1) the total area above threshold, $A(\nu)$, (2) the total contour length, $C(\nu)$, and (3) the genus, $G(\nu)$, which is the number of hot spots minus the number of cold spots. Parameterization of contours by threshold is computationally simpler, while parameterization by area reduces the correlations between the Minkowski functionals [263]; however, when a joint analysis of the three Minkowski functionals is performed, one has to explicitly include their covariance anyway. Therefore the simpler threshold parameterization will be used.

In Ref. [139] the Minkowski functionals at five different resolutions from the pixel size of 3.7 degrees in diameter to 12 arcminutes have been measured. Fig. 18 shows one example at $28'$ pixel resolution. The gray band shows the 68% confidence region derived from 1000 Gaussian Monte Carlo simulations.

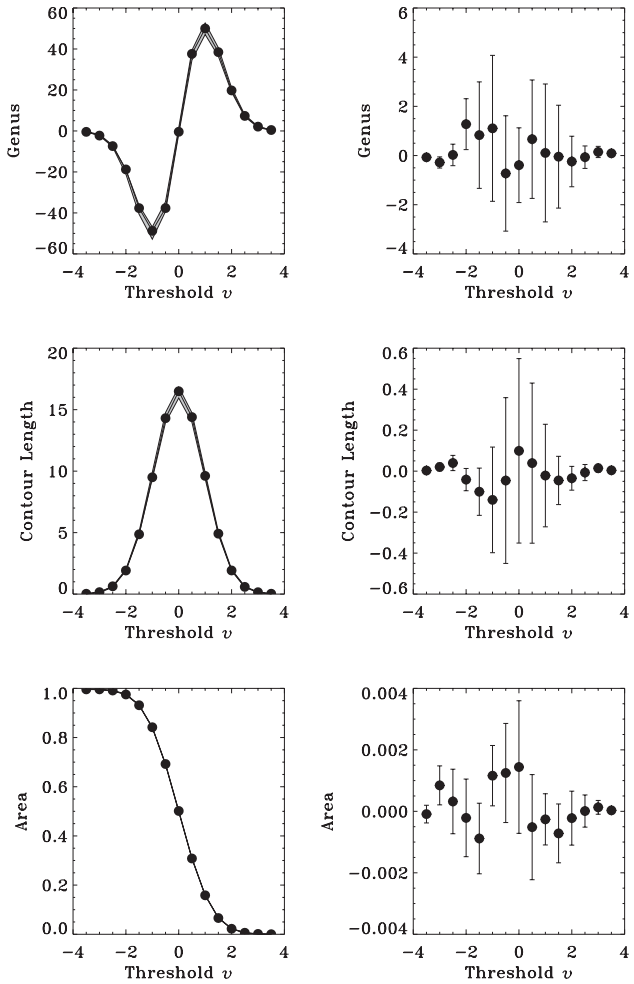

Fig. 18. The Minkowski functionals for the foreground-cleaned WMAP data [139]. The Minkowski functions at $28'$ pixel resolution (filled circles) and the residuals between the mean of the Gaussian simulations and the WMAP data. The gray band shows the 68% confidence interval for the Gaussian Monte Carlo simulations. The WMAP data are in excellent agreement with the Gaussian simulations.

(See Section 2.3 of Ref. [139] for description of the simulations.) The WMAP data are in excellent agreement with the Gaussian simulations at all resolutions. But, *how Gaussian is it?*

10.6.2. Angular bispectrum

For the second test, we use the fast cubic statistics derived in Section 10.5, which combine three-point (triangle) configurations of the angular bispectrum that are sensitive to the models under consideration.

Once again, we consider two components. The first one is the primordial non-Gaussianity from inflation parametrized by $f_{\rm NL}$ (see Section 8), which determines the amplitude of a quadratic term added to Bardeen's curvature perturbations: $\Phi(\mathbf{x}) = \Phi_{\rm L}(\mathbf{x}) + f_{\rm NL}[\Phi_{\rm L}^2(\mathbf{x}) - \langle\Phi_{\rm L}^2(\mathbf{x})\rangle]$. It is useful to estimate the

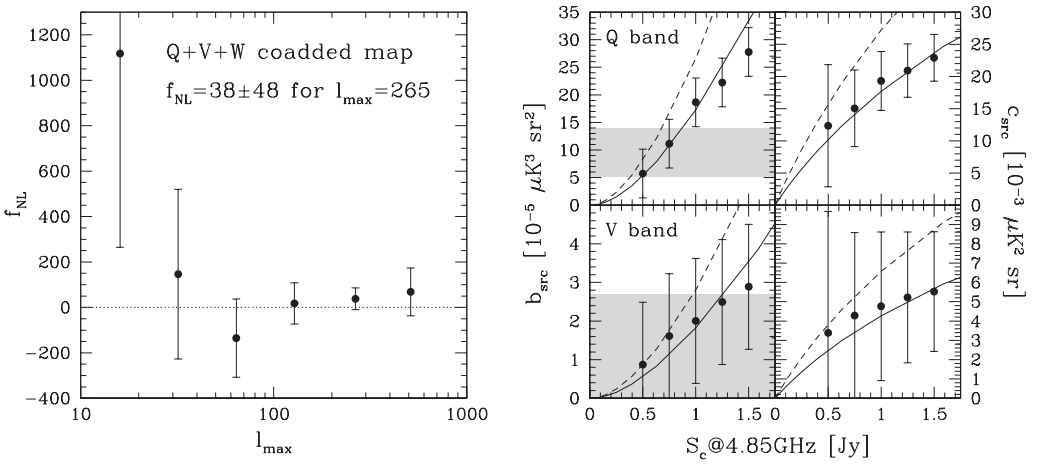

Fig. 19. Primordial non-Gaussianity and point-source contribution [139]. (*Left panel*): The non-linearity parameter, $f_{\rm NL}$, as a function of the maximum multipole $l_{\rm max}$, measured from the Q+V+W coadded map using the bispectrum estimator. The error bars at each $l_{\rm max}$ are not independent. (*Right panel*): The point-source angular bispectrum $b^{\rm src}$ and power-spectrum $c^{\rm src}$. The left panels show $b^{\rm src}$ in Q band (top panel) and V band (bottom panel). The shaded areas show measurements from the WMAP sky maps with the standard source cut, while the filled circles show those with flux thresholds $S_{\rm c}$ defined at 4.85 GHz. The dashed lines show predictions from the source count model of Ref. [279], while the solid lines are those multiplied by 0.65 to match the WMAP measurements. The right panels show $c^{\rm src}$. The filled circles are computed from the measured $b^{\rm src}$ substituted into Eq. (534). The lines are predictions. The error bars are not independent.

r.m.s. amplitude of Φ to see how important the second-order term is. One obtains $\langle\Phi^2\rangle^{1/2} \simeq \langle\Phi_{\rm L}^2\rangle^{1/2}(1 + f_{\rm NL}^2\langle\Phi_{\rm L}^2\rangle)$, where $\langle\Phi^2\rangle^{1/2} \simeq 3.3\times10^{-5}$ [36]; thus, a fractional contribution from the second term is

$$f_{\rm NL}^2\langle\Phi_{\rm L}^2\rangle \simeq 10^{-5}(f_{\rm NL}/100)^2\ . \tag{533}$$

We are talking about very small effects.

This parameterization is useful to find *quantitative* constraints on the amount of non-Gaussianity allowed by the CMB data. Also, the form is general in that $f_{\rm NL}$ parameterizes the leading-order non-linear corrections to Φ.

Fig. 19 shows $f_{\rm NL}$ measured from the foreground-cleaned Q+V+W coadded map using the cubic statistic, as a function of the maximum multipole $l_{\rm max}$ (for details of measurements, see Ref. [139]). There is no significant detection of $f_{\rm NL}$ at any angular scale. There is no significant band-to-band variation, or significant detection in any band. The best constraint is $-58 < f_{\rm NL} < 134$ (95%), which is equivalent to say that the fractional contribution to the r.m.s. value of Φ from the second-order term is smaller than 2×10^{-5}. These results support inflationary models, but still do not exclude the possibility of having a small contribution from non-linearities predicted by second-order perturbation theory.

Note that $f_{\rm NL}$ for $l_{\rm max} = 265$ has a smaller error than that for $l_{\rm max} = 512$, because the latter is dominated by the instrumental noise. Since all the pixels outside the cut region are uniformly weighted, the inhomogeneous noise in the map (pixels on the ecliptic equator are noisier than those on the north and south poles) is not accounted for. This leads to a noisier estimator than a minimum variance estimator. The constraint on $f_{\rm NL}$ for $l_{\rm max} = 512$ will likely improve with more appropriate pixel-weighting schemes [108,252]. Apparently, the fact that the constraint actually obtained from the data is worse than predicted

(cf., Table 6) should be due to sub-optimalness of the current estimator. The simple inverse noise (N^{-1}) weighting makes the constraints much worse than the uniform weighting, as it increases errors on large angular scales where the CMB signal dominates over the instrumental noise. (However, it works fine for the point-sources.) The uniform weighting is thus closer to optimal.

The Minkowski functionals shown in Fig. 18 also place constraints on $f_{\rm NL}$, comparing the data to the predictions derived from Monte Carlo simulations of the non-Gaussian CMB (for details of the simulations, see Appendix A of [139]). It has been found that $f_{\rm NL} < 139$ (95%), remarkably consistent with that from the bispectrum analysis.

10.6.3. Point-source non-Gaussianity

The second component is the foreground non-Gaussianity from radio point-sources, parameterized by the skewness, $b^{\rm src}$. The filled circles in the right panels of Fig. 19 show $b^{\rm src}$ measured in Q (top panel) and V (bottom panel) band. We have used source masks for various flux cuts, $S_{\rm c}$, defined at 4.85 GHz to make these measurements. (The masks are made from the GB6+PMN 5 GHz source catalogue.) We find that $b^{\rm src}$ increases as $S_{\rm c}$: the brighter sources being unmasked, the more non-Gaussianity is detected. On the other hand one can make predictions for $b^{\rm src}$ using the source count model. Comparing the measured values of $b^{\rm src}$ with the predicted counts by [279] (dashed lines) at 44 GHz, one finds that the measured values are smaller than the predicted values by a factor of 0.65. The solid lines show the predictions multiplied by 0.65. Our value for the correction factor matches well the one obtained from the WMAP source counts for 2–10 Jy in Q band [39].

The source bispectrum, $b^{\rm src}$, is related to the source power-spectrum, $c^{\rm src}$, by an integral relation [139],

$$c^{\rm src}(S_{\rm c}) = b^{\rm src}(S_{\rm c})[g(\nu)S_{\rm c}]^{-1} + \int_0^{S_{\rm c}} \frac{{\rm d}S}{S}\, b^{\rm src}(S)[g(\nu)S]^{-1}\ , \tag{534}$$

where $g(\nu)$ is a conversion factor from Jy sr^{-1} to μK which depends upon the observing frequency ν as $g(\nu) = (24.76\,{\rm Jy}\,\mu{\rm K}^{-1}\,{\rm sr}^{-1})^{-1}[(\sinh x/2)/x^2]^2$, $x \equiv h\nu/k_{\rm B}T_0 \simeq \nu/(56.78\,{\rm GHz})$. One can use this equation combined with the measured $b^{\rm src}$ as a function of the flux threshold $S_{\rm c}$ to directly determine $c^{\rm src}$ as a function of $S_{\rm c}$, *without relying on any extrapolations*. The right panels of Fig. 19 also show the estimated $c^{\rm src}$ as filled circles. The measurements suggest that $c^{\rm src}$ for the standard source mask (indicated by the shaded area) is $c^{\rm src} = (15 \pm 6) \times 10^{-3}\,\mu{\rm K}^2$ sr in Q band. In V band, $c^{\rm src} = (4.5 \pm 4) \times 10^{-3}\,\mu{\rm K}^2$ sr.

In addition to the bispectrum, the WMAP team has carried out other methods to estimate the source contribution: (1) extrapolation from the number counts of detected sources in the WMAP data [39], and (2) the angular power-spectrum on small angular scales [110]. These methods yield consistent results.

In summary, the WMAP 1-year data has enormously improved the sensitivity for testing the Gaussianity of the CMB. Yet, we do not have any compelling evidence for primordial non-Gaussianity. This result is consistent with what is predicted by inflation and the second-order perturbation theory. There may be some chance to find non-Gaussian signals arising from second-order perturbations. Detection can be made possible by the Planck experiment combining the temperature and polarization anisotropies. While we can detect $f_{\rm NL} \sim 5$ by using the temperature alone (see Table 6), combining the polarization measurements increases our sensitivity: we have several observables for the bispectrum such as $\langle {\rm TTT}\rangle$, $\langle {\rm TTE}\rangle$, $\langle {\rm TEE}\rangle$, and $\langle {\rm EEE}\rangle$. The future polarization-dedicated satellite experiment (e.g., CMBPol) in combination with the Planck temperature map may enable us to detect $f_{\rm NL} \sim 3$.

11. Conclusions and future prospects

Testing Gaussianity of the primordial fluctuations is and will be one of the most powerful probes of the inflationary paradigm. Gaussianity tests are complementary to conventional ones using the power-spectrum: as we have shown in this review, Gaussianity tests enable us to discriminate between different inflationary models which would be indistinguishable otherwise. We have examined various examples including the standard single-field slow-roll inflation, the curvaton model, the inhomogeneous reheating scenario, multi-field models and some unconventional scenarios, which make unique predictions for the strength of non-Gaussianity and its shape (see Table 1 for a summary). For the single-field slow-roll inflationary model, on the other hand, we have shown that inflation itself produces a negligible amount of non-Gaussianity, and the dominant contribution comes from the evolution of the ubiquitous second-order perturbations after inflation, which is potentially detectable with future observations of temperature and polarization anisotropies of the CMB. This effect *must exist* regardless of inflationary models, setting the minimum level of non-Gaussianity in the cosmological perturbations. Alternative models for the generation of perturbations might produce stronger non-Gaussianity than this minimum amount. Therefore, if we do not find any evidence for this ubiquitous non-Gaussianity, then it will challenge our understanding of the evolution of cosmological perturbations at a deeper level. (In other words, no detection of non-Gaussianity at the level of $f_{\rm NL} \sim 1$ rules out our standard cosmological model!) It is extremely important to keep improving upon our sensitivity to the primordial non-Gaussianity until we reach the critical sensitivity, $f_{\rm NL} \sim 1$.

We have reviewed in great detail the current constraints on $f_{\rm NL}$ from the angular bispectrum of the CMB. Here, let us make a remark on future prospects for observational constraints on $f_{\rm NL}$. It has been shown that the angular bispectrum of temperature anisotropy alone can detect non-Gaussianity, if $|f_{\rm NL}| > 5$ [143]. This estimate assumes that the Planck satellite is the ultimate experiment measuring temperature anisotropy in terms of primordial non-Gaussianity. Small-scale CMB experiments, such as the Atacama Cosmology Telescope [146], would detect non-Gaussianity from secondary anisotropies (see Section 10.4 and, e.g., Refs. [92,70,6]). If we add polarization information (which is assumed to be measured as accurately as temperature up to $l \sim 3000$), then one can improve it to $f_{\rm NL} > 3$. This is still a factor of 3 larger than the critical limit; however, fortunately we have many more Gaussianity tests which can, in principle, give us independent measurements of $f_{\rm NL}$. If fluctuations are Gaussian, then the power-spectrum contains all the statistical information, so that one cannot overcome cosmic variance by using other statistical tools; however, if fluctuations are non-Gaussian, then there can be many independent statistical tools measuring different aspects of the same non-Gaussianity, giving independent constraints on the strength and shape of non-Gaussianity. [34] If those statistical tools are orthogonal to the bispectrum, then one can improve the limits on $f_{\rm NL}$ by the square-root of the number of independent statistical tools (i.e., we need at least 9 completely independent methods to measure $f_{\rm NL} = 1$).

Although numerous statistical estimators have been applied to the CMB data for Gaussianity tests, only a few of these (Minkowski functionals [139], Mexican-hat wavelets [209], local curvature [58]) have been used to find limits on $f_{\rm NL}$. Also, the extent to which these statistical tools are independent

[34] Let us mention here the analysis of the 3-point function of CMB anisotropies in the WMAP data of Ref. [91]. Also interesting is a statistical method based on the multivariate empirical distribution function of the spherical harmonics, proposed in Refs. [105,106].

remains unknown (see Ref. [57] for the first attempt to address this issue); thus, studying statistical power and complementarity of the statistical tools measuring $f_{\rm NL}$ will be one of the most important goals. To achieve this goal, it is crucial to have accurate numerical simulations of the non-Gaussian CMB sky maps (both temperature and polarization), as well as analytical calculations of the effects of $f_{\rm NL}$ on the statistics can be very complicated. At the time of writing, analytical predictions exist only for the bispectrum [143] and the trispectrum [215]. Simulations of non-Gaussian temperature fluctuations with the $f_{\rm NL}$-model already exist [139,160] and can be readily extended to include polarization, as well as arbitrary non-Gaussian initial conditions in the primordial curvature and entropy perturbations. Moreover, direct simulations of the non-linear dynamics of cosmological perturbations, which have been evaluated analytically in this review, may be feasible. We clearly need a systematic study of the combined statistical power of various methods constraining $f_{\rm NL}$, using these simulations.

In addition to the CMB, we have other methods to constrain the primordial non-Gaussianity. Galaxy correlations at large distances, where non-linear clustering is modest, still preserve statistical properties of the primordial fluctuations; thus, one can use them to find limits on the primordial non-Gaussianity (see, e.g., Ref. [47] for a review). Using the three-dimensional bispectrum, the authors of Ref. [256] have found a limit $-2000 < f_{\rm NL} < 1600$ (95%), from the PSCz survey. They conclude that the bispectrum analysis of the Sloan Digital Sky Survey can reach $|f_{\rm NL}| \sim (150\text{–}200)$, and that of an idealistic all-sky redshift survey up to $z \sim 1$ can reach $|f_{\rm NL}| \sim 1$. In principle, therefore, the LSS data might become as competitive as the CMB data. One big advantage of the LSS data is three-dimensional information. Since the CMB data give us only two-dimensional information, the number of modes of the bispectrum that one can measure is fairly limited even on the full sky, and this is the fundamental limitation of the CMB bispectrum. On the other hand, the number of modes available in the three-dimensional bispectrum is enormous, and helps to obtain tight limits on $f_{\rm NL}$. Of course, there are disadvantages of the LSS data: non-linear clustering and bias producing spurious non-Gaussian signals [285]. Combination of CMB and LSS data will thus offer a systematic error check and potentially an improved signal-to-noise ratio for detection of $f_{\rm NL} \sim 1$.

Yet another tool is the number of massive halos (e.g., clusters of galaxies) at high z [169,147,193,244,293,61]. The number of massive clusters is very sensitive to statistical properties of the primordial fluctuations. For example, one can calculate the number of clusters corresponding to density peaks of, say, 3-σ, for a Gaussian distribution. Since these objects are rare, the number is very sensitive to the exact shape of the tail of the probability distribution function of density fluctuations. Even a slight amount of non-Gaussianity can change it rather dramatically. This method is powerful when density fields are positively skewed, giving more objects for a given mass and redshift. (A positive skewness in density fluctuations corresponds $f_{\rm NL} < 0$.) Although the current limits on $f_{\rm NL}$ from the WMAP constrain deviation of the number of massive clusters from the Gaussian prediction to within $\sim$50% for $z < 1$ and $M < 10^{15} M_{\odot}$, the constraints rapidly improve as one goes to higher z [139]. Therefore, one needs to go to high z to look for signatures of non-Gaussianity in cluster abundance. One major problem of this method is, however, as it has been correctly pointed out by several authors [284,256], that one needs very accurate (of order a few percent) determinations of the mass of clusters at $z > 1$, in order to find $f_{\rm NL} \sim -100$. It seems rather difficult to achieve this accuracy for many clusters. Yet, *if* one finds *one* exceptionally massive cluster ($\sim 10^{15} M_{\odot}$) at a very high z (~ 3), then it should tantalizingly indicate the presence of non-Gaussianity. A preliminary, lower z version of this methodology was attempted in Ref. [293] for a massive cluster, MS 1054-03, at $z = 0.83$, where evidence for non-Gaussianity was claimed; however, unfortunately uncertainties in mass determinations are still too large to claim a robust detection of non-Gaussianity

(the WMAP limit is inconsistent with this detection), but in principle one can extend it to higher z with a better determination of the mass.

Non-cosmological non-Gaussianities from Galactic emission and extragalactic point-sources are serious contaminants. Fortunately the shape of the angular bispectrum from point sources is very different from that of the primordial bispectrum, and they can be separated very well [143,139]. A problem occurs when we find a non-Gaussian signal, but we do not know what the origin is. Although many authors claim detection of non-Gaussianity in the WMAP 1-yr data [62,219,65,287,71,81,104,209,153], none of its can be accounted for by the $f_{\rm NL}$ model, and their origin is unclear (also, one should keep in mind that the statistical significance of these detections is less than around 3-σ). Clearly, understanding possible foreground contamination and other possible systematics in the data are critical issues for measurements of primordial non-Gaussianity.

Therefore, testing the Gaussianity of primordial perturbations represents a challenge for the present and future CMB experiments, as well as for LSS observations, which might reveal the ultimate origin of the structures we see in the Universe today.

Acknowledgements

We thank all our colleagues with whom we have discussed many issues concerning non-Gaussianity from inflation. In particular, we are indebted to Viviana Acquaviva, Paolo Creminelli, Domenico Marinucci, Andrew Liddle, Michele Liguori, David Lyth, Silvia Mollerach, Jim Peebles, Licia Verde, Nicola Vittorio and David Wands. N.B. acknowledges PPARC for financial support. E.K. acknowledges that Section 10.6 is based upon work in collaboration with the WMAP science team: C. Barnes, C. Bennett (PI), M. Halpern, R. Hill, G. Hinshaw, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, S. Meyer, N. Odegard, L. Page, H. Peiris, D. Spergel, G. Tucker, L. Verde, J. Weiland, E. Wollack, and E. Wright. The WMAP mission is made possible by the support of the Office of Space Sciences at NASA Headquarters and by the hard and capable work of scores of scientists, engineers, technicians, machinists, data analysts, budget analysts, managers, administrative staff, and reviewers. S.M. and A.R. thank the INFN Gruppo IV for financial support within the PD51 Project. S.M. also thanks partial financial support from INAF (progetto di ricerca *Non-Gaussian primordial perturbations: constraints from CMB and redshift surveys*).

Appendix A. Second-order gravitational perturbations

A.1. Basic notation

The number of spatial dimensions is $n = 3$. Greek indices ($\alpha, \beta, \ldots, \mu, \nu, \ldots$) run from 0 to 3, while latin indices ($a, b, \ldots, i, j, k, \ldots m, n, \ldots$) run from 1 to 3. The total space–time metric $g_{\mu\nu}$ has signature $(-,+,+,+)$. The connection coefficients are defined as

$$\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2}\, g^{\alpha\rho}\left(\frac{\partial g_{\rho\gamma}}{\partial x^{\beta}} + \frac{\partial g_{\beta\rho}}{\partial x^{\gamma}} - \frac{\partial g_{\beta\gamma}}{\partial x^{\rho}}\right) . \tag{A.1}$$

The Riemann tensor is defined as

$$R^{\alpha}{}_{\beta\mu\nu} = \Gamma^{\alpha}_{\beta\nu,\mu} - \Gamma^{\alpha}_{\beta\mu,\nu} + \Gamma^{\alpha}_{\lambda\mu}\Gamma^{\lambda}_{\beta\nu} - \Gamma^{\alpha}_{\lambda\nu}\Gamma^{\lambda}_{\beta\mu} . \tag{A.2}$$

The Ricci tensor is a contraction of the Riemann tensor

$$R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu} \tag{A.3}$$

and in terms of the connection coefficient it reads

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\mu \Gamma^\alpha_{\nu\alpha} + \Gamma^\alpha_{\sigma\alpha}\Gamma^\sigma_{\mu\nu} - \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\mu\alpha} \ . \tag{A.4}$$

The Ricci scalar is given by contracting the Ricci tensor

$$R = R^\mu{}_\mu \ . \tag{A.5}$$

The Einstein tensor is defined as

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}\, g_{\mu\nu} R \ . \tag{A.6}$$

The Einstein equations are written as $G_{\mu\nu} = \kappa^2 T_{\mu\nu}$, so that $\kappa^2 \equiv 8\pi G_{\rm N}$, where $G_{\rm N}$ is Newton's constant.

In the following expressions we have chosen a specific ordering of the terms. In the expressions in which two spatial indices appear, such as Eq. (A.18), we have assembled together the terms proportional to δ_{ij}. The intrinsically second-order terms precede the source terms which are quadratic in the first-order perturbations. The second-order fluctuations have been listed in the following order as $\phi^{(2)}$, $\psi^{(2)}$, $\omega^{(2)}$, $\omega_i^{(2)}$, $\chi^{(2)}$, $\chi_i^{(2)}$ and $\chi_{ij}^{(2)}$, respectively. This ordering simplifies the analogy between the first-order and the second-order equations and allows to obtain immediately the expressions in a given gauge.

A.2. The connection coefficients

In a spatially flat FRW background the connection coefficients are

$$\Gamma^0_{00} = \frac{a'}{a}; \quad \Gamma^i_{0j} = \frac{a'}{a}\,\delta^i_j; \quad \Gamma^0_{ij} = \frac{a'}{a}\,\delta_{ij} \ ; \tag{A.7}$$

$$\Gamma^i_{00} = \Gamma^0_{0i} = \Gamma^i_{jk} = 0 \ . \tag{A.8}$$

The first-order perturbed connection coefficients corresponding to first-order metric perturbations in Eq. (103) are

$$\delta^{(1)}\Gamma^0_{00} = \phi^{(1)\prime} \ , \tag{A.9}$$

$$\delta^{(1)}\Gamma^0_{0i} = \partial_i \phi^{(1)} + \frac{a'}{a}\,\partial_i \omega^{(1)} \ , \tag{A.10}$$

$$\delta^{(1)}\Gamma^i_{00} = \frac{a'}{a}\,\partial^i \omega^{(1)} + \partial^i \omega^{(1)\prime} + \partial^i \phi^{(1)} \ , \tag{A.11}$$

$$\delta^{(1)}\Gamma^0_{ij} = -2\frac{a'}{a}\phi^{(1)}\delta_{ij} - \partial_i\partial_j\omega^{(1)} - 2\frac{a'}{a}\psi^{(1)}\delta_{ij} - \psi^{(1)'}\delta_{ij} - \frac{a'}{a}D_{ij}\chi^{(1)} + \frac{1}{2}D_{ij}\chi^{(1)'}\,, \tag{A.12}$$

$$\delta^{(1)}\Gamma^i_{0j} = -\psi^{(1)'}\delta^i_j + \tfrac{1}{2}D^i_j\chi^{(1)'}\,, \tag{A.13}$$

$$\delta^{(1)}\Gamma^i_{jk} = -\partial_j\psi^{(1)}\delta^i_k - \partial_k\psi^{(1)}\delta^i_j + \partial^i\psi^{(1)}\delta_{jk} - \frac{a'}{a}\partial^i\omega^{(1)}\delta_{jk} + \tfrac{1}{2}\partial_j D^i_k\chi^{(1)} + \tfrac{1}{2}\partial_k D^i_j\chi^{(1)} - \tfrac{1}{2}\partial^i D_{jk}\chi^{(1)}\,. \tag{A.14}$$

At second order we get

$$\delta^{(2)}\Gamma^0_{00} = \frac{1}{2}\phi^{(2)'} - 2\phi^{(1)}\phi^{(1)'} + \partial^k\phi^{(1)}\partial_k\omega^{(1)} + \frac{a'}{a}\partial^k\omega^{(1)}\partial_k\omega^{(1)} + \partial^k\omega^{(1)}\partial_k\omega^{(1)'}\,, \tag{A.15}$$

$$\delta^{(2)}\Gamma^0_{0i} = \frac{1}{2}\partial_i\phi^{(2)} + \frac{1}{2}\frac{a'}{a}(\partial_i\omega^{(2)} + \omega^{(2)}_i) - 2\phi^{(1)}\partial_i\phi^{(1)} - 2\frac{a'}{a}\phi^{(1)}\partial_i\omega^{(1)} - \psi^{(1)'}\partial_i\omega^{(1)} + \frac{1}{2}\partial^k\omega^{(1)}D_{ik}\chi^{(1)'}\,, \tag{A.16}$$

$$\delta^{(2)}\Gamma^i_{00} = \frac{1}{2}\partial^i\phi^{(2)} + \frac{1}{2}\frac{a'}{a}(\partial^i\omega^{(2)} + \omega^{i(2)}) + \frac{1}{2}(\partial^i\omega^{(2)'} + (\omega^{i(2)})') + 2\psi^{(1)}\partial^i\phi^{(1)} - \phi^{(1)'}\partial^i\omega^{(1)} + 2\frac{a'}{a}\psi^{(1)}\partial^i\omega^{(1)} + 2\psi^{(1)}\partial^i\omega^{(1)'} - \partial_k\phi^{(1)}D^{ik}\chi^{(1)} - \frac{a'}{a}\partial_k\omega^{(1)}D^{ik}\chi^{(1)} - \partial_k\omega^{(1)'}D^{ik}\chi^{(1)}\,, \tag{A.17}$$

$$\delta^{(2)}\Gamma^0_{ij} = \left(-\frac{a'}{a}\phi^{(2)} - \frac{1}{2}\psi^{(2)'} - \frac{a'}{a}\psi^{(2)} + 4\frac{a'}{a}(\phi^{(1)})^2 + 2\phi^{(1)}\psi^{(1)'} + 4\frac{a'}{a}\phi^{(1)}\psi^{(1)} + \partial^k\omega^{(1)}\partial_k\psi^{(1)} - \frac{a'}{a}\partial^k\omega^{(1)}\partial_k\omega^{(1)}\right)\delta_{ij} - \frac{1}{2}\partial_i\partial_j\omega^{(2)} + \tfrac{1}{4}(D_{ij}\chi^{(2)'} + \partial_j\chi^{(2)'}_i + \partial_i\chi^{(2)'}_j + (\chi^{(2)}_{ij})') + \frac{1}{2}\frac{a'}{a}(D_{ij}\chi^{(2)} + \partial_i\chi^{(2)}_j + \partial_j\chi^{(2)}_i + \chi^{(2)}_{ij}) - \frac{1}{4}(\partial_i\omega^{(2)}_j + \partial_j\omega^{(2)}_i) + 2\phi^{(1)}\partial_i\partial_j\omega^{(1)} - \partial_i\psi^{(1)}\partial_j\omega^{(1)} - \partial_j\psi^{(1)}\partial_i\omega^{(1)} - \phi^{(1)}D_{ij}\chi^{(1)'} + \tfrac{1}{2}\partial^k\omega^{(1)}\partial_i D_{kj}\chi^{(1)} + \tfrac{1}{2}\partial^k\omega^{(1)}\partial_j D_{ik}\chi^{(1)} - \tfrac{1}{2}\partial^k\omega^{(1)}\partial_k D_{ij}\chi^{(1)}\,, \tag{A.18}$$

$$\begin{aligned}\delta^{(2)}\Gamma^i_{0j} = & -\tfrac{1}{2}\psi^{(2)\prime}\delta^i_j + \tfrac{1}{4}\left(D^i_j\chi^{(2)\prime} + \partial_j(\chi^{(2)i})' + \partial^i(\chi^{(2)}_j)'\right.\\ & \left. + (\chi^{(2)i}_j)'\right) + \tfrac{1}{4}\left(\partial_j\omega^{i(2)} - \partial^i\omega^{(2)}_j\right)\\ & - 2\psi^{(1)}\psi^{(1)\prime}\delta^i_j - \partial^i\omega^{(1)}\partial_j\phi^{(1)} - \frac{a'}{a}\,\partial^i\omega^{(1)}\partial_j\omega^{(1)}\\ & + \psi^{(1)}D^i_j\chi^{(1)\prime} + \psi^{(1)\prime}D^i_j\chi^{(1)} - \tfrac{1}{2}\,D^{ik}\chi^{(1)}D_{kj}\chi^{(1)\prime}\ , \end{aligned} \tag{A.19}$$

$$\begin{aligned}\delta^{(2)}\Gamma^i_{jk} = & \tfrac{1}{2}\left(-\partial_j\psi^{(2)}\delta^i{}_k - \partial_k\psi^{(2)}\delta^i_j + \partial^i\psi^{(2)}\delta_{jk}\right)\\ & + \tfrac{1}{4}\left(\partial_j D^i{}_k\chi^{(2)} + \partial_k D^i{}_k\chi^{(2)} - \partial^i D_{jk}\chi^{(2)}\right)\\ & + \tfrac{1}{2}\,\partial_j\partial_k\chi^{i(2)} + \tfrac{1}{4}(\partial_j\chi^{i(2)}{}_k + \partial_k\chi^{i(2)}_j - \partial^i\chi^{(2)}_{jk})\\ & - \frac{1}{2}\,\frac{a'}{a}\,(\partial^i\omega^{(2)} + \omega^{i(2)})\delta_{jk} + 2\psi^{(1)}(-\partial_j\psi^{(1)}\delta^i{}_k - \partial_k\psi^{(1)}\delta^i{}_j + \partial^i\psi^{(1)}\delta_{jk})\\ & + 2\,\frac{a'}{a}\,\phi^{(1)}\partial^i\omega^{(1)}\delta_{jk} + \partial^i\omega^{(1)}\partial_j\partial_k\omega^{(1)} + \psi^{(1)\prime}\partial^i\omega^{(1)}\delta_{jk}\\ & + \psi^{(1)}(\partial_j D^i{}_k\chi^{(1)} + \partial_k D^i_j\chi^{(1)} - \partial^i D_{jk}\chi^{(1)}) + \partial_j\psi^{(1)}D^i{}_k\chi^{(1)}\\ & + \partial_k\psi^{(1)}D^i_j\chi^{(1)} - \partial_m\psi^{(1)}D^{im}\chi^{(1)}\delta_{jk} - \frac{a'}{a}\,\partial^i\omega^{(1)}D_{jk}\chi^{(1)}\\ & + \frac{a'}{a}\,\partial^m\omega^{(1)}D^i{}_m\chi^{(1)}\delta_{jk} - \tfrac{1}{2}\,\partial^i\omega^{(1)}D_{jk}\chi^{(1)\prime}\\ & - \tfrac{1}{2}\,D^{im}\chi^{(1)}\partial_j D_{mk}\chi^{(1)} - \tfrac{1}{2}\,D^{im}\chi^{(1)}\partial_k D_{mj}\chi^{(1)} + \tfrac{1}{2}\,D^{im}\chi^{(1)}\partial_m D_{jk}\chi^{(1)}\ . \end{aligned} \tag{A.20}$$

A.3. The Ricci tensor components

In a spatially flat FRW background the components of the Ricci tensor $R_{\mu\nu}$ are given by

$$R_{00} = -3\,\frac{a''}{a} + 3\left(\frac{a'}{a}\right)^2; \quad R_{0i} = 0\ ; \tag{A.21}$$

$$R_{ij} = \left[\frac{a''}{a} + \left(\frac{a'}{a}\right)^2\right]\delta_{ij}\ . \tag{A.22}$$

The first-order perturbed Ricci tensor components read

$$\delta^{(1)}R_{00} = \frac{a'}{a}\,\partial_i\partial^i\omega^{(1)} + \partial_i\partial^i\omega^{(1)\prime} + \partial_i\partial^i\phi^{(1)} + 3\psi^{(1)\prime\prime} + 3\,\frac{a'}{a}\,\psi^{(1)\prime} + 3\,\frac{a'}{a}\,\phi^{(1)\prime}\ , \tag{A.23}$$

$$\delta^{(1)}R_{0i} = \frac{a''}{a}\,\partial_i\omega^{(1)} + \left(\frac{a'}{a}\right)^2\partial_i\omega^{(1)} + 2\partial_i\psi^{(1)\prime} + 2\,\frac{a'}{a}\,\partial_i\phi^{(1)} + \frac{1}{2}\,\partial_k D^k{}_i\chi^{(1)\prime}\,, \tag{A.24}$$

$$\begin{aligned}\delta^{(1)}R_{ij} = {} & \left[-\frac{a'}{a}\,\phi^{(1)\prime} - 5\,\frac{a'}{a}\,\psi^{(1)\prime} - 2\,\frac{a''}{a}\,\phi^{(1)} - 2\left(\frac{a'}{a}\right)^2\phi^{(1)} - 2\,\frac{a''}{a}\,\psi^{(1)}\right.\\ & \left. -2\left(\frac{a'}{a}\right)^2\psi^{(1)} - \psi^{(1)\prime\prime} + \partial_k\partial^k\psi^{(1)} - \frac{a'}{a}\,\partial_k\partial^k\omega^{(1)}\right]\delta_{ij} - \partial_i\partial_j\omega^{(1)\prime}\\ & + \frac{a'}{a}\,D_{ij}\chi^{(1)\prime} + \frac{a''}{a}\,D_{ij}\chi^{(1)} + \left(\frac{a'}{a}\right)^2 D_{ij}\chi^{(1)} + \frac{1}{2}\,D_{ij}\chi^{(1)\prime\prime} + \partial_i\partial_j\psi^{(1)}\\ & - \partial_i\partial_j\phi^{(1)} - 2\,\frac{a'}{a}\,\partial_i\partial_j\omega^{(1)} + \frac{1}{2}\,\partial_k\partial_i D^k{}_j\chi^{(1)} + \frac{1}{2}\,\partial_k\partial_j D^k{}_i\chi^{(1)} - \frac{1}{2}\,\partial_k\partial^k D_{ij}\chi^{(1)}\,.\end{aligned} \tag{A.25}$$

At second order we obtain

$$\begin{aligned}\delta^{(2)}R_{00} = {} & \frac{3}{2}\,\frac{a'}{a}\,\phi^{(2)\prime} + \frac{1}{2}\,\nabla^2\phi^{(2)} + \frac{3}{2}\,\frac{a'}{a}\,\psi^{(2)\prime} + \frac{3}{2}\,\psi^{(2)\prime\prime} + \frac{1}{2}\,\frac{a'}{a}\,\partial_k\partial^k\omega^{(2)}\\ & + \frac{1}{2}\,\partial_k\partial^k\omega^{(2)\prime} - 6\,\frac{a'}{a}\,\phi^{(1)}\phi^{(1)\prime} - \partial^k\phi^{(1)}\partial_k\phi^{(1)} - 3\phi^{(1)\prime}\psi^{(1)\prime}\\ & + 2\psi^{(1)}\nabla^2\,\phi^{(1)} - \partial_k\psi^{(1)}\partial^k\phi^{(1)} + 6\,\frac{a'}{a}\,\psi^{(1)}\psi^{(1)\prime} + 6\psi^{(1)}\psi^{(1)\prime\prime}\\ & + 3(\psi^{(1)\prime})^2 - \phi^{(1)\prime}\nabla^2\omega^{(1)} + \frac{a'}{a}\,\partial^k\omega^{(1)}\partial_k\phi^{(1)} + \frac{a''}{a}\,\partial^k\omega^{(1)}\partial_k\omega^{(1)}\\ & + \left(\frac{a'}{a}\right)^2\partial^k\omega^{(1)}\partial_k\omega^{(1)} - \frac{a'}{a}\,\partial_k\psi^{(1)}\partial^k\omega^{(1)} + 2\,\frac{a'}{a}\,\psi^{(1)}\nabla^2\omega^{(1)} - \partial_k\psi^{(1)}\partial^k\omega^{(1)\prime}\\ & + 2\psi^{(1)}\nabla^2\omega^{(1)\prime} + 3\,\frac{a'}{a}\,\partial^k\omega^{(1)}\partial_k\omega^{(1)\prime} - \partial_k\phi^{(1)}\partial_i D^{ik}\chi^{(1)} - \partial_i\partial_k\phi^{(1)}D^{ik}\chi^{(1)}\\ & - \frac{a'}{a}\,\partial_i\partial_k\omega^{(1)}D^{ik}\chi^{(1)} - \frac{a'}{a}\,\partial_k\omega^{(1)}\partial_i D^{ik}\chi^{(1)} - \partial_k\omega^{(1)\prime}\partial_i D^{ik}\chi^{(1)}\\ & - \partial_i\partial_k\omega^{(1)\prime}D^{ik}\chi^{(1)} + \tfrac{1}{2}\,D^{ik}\chi^{(1)}D_{ki}\chi^{(1)\prime\prime} + \tfrac{1}{4}\,D^{ik}\chi^{(1)\prime}D_{ki}\chi^{(1)\prime}\\ & + \frac{1}{2}\,\frac{a'}{a}\,D^{ik}\chi^{(1)}D_{ki}\chi^{(1)\prime}\,.\end{aligned} \tag{A.26}$$

$$
\begin{aligned}
\delta^{(2)}R_{0i} = {} & \frac{a'}{a}\,\partial_i\phi^{(2)} + \partial_i\psi^{(2)\prime} + \frac{1}{4}\,\partial_k D_i^k\chi^{(2)\prime} + \frac{1}{4}\,\nabla^2\chi_i^{(2)\prime} - \frac{1}{4}\,\nabla^2\omega_i^{(2)} \\
& + \frac{1}{2}\,\frac{a''}{a}\,(\partial_i\omega^{(2)} + \omega_i^{(2)}) + \frac{1}{2}\left(\frac{a'}{a}\right)^2(\partial_i\omega^{(2)} + \omega_i^{(2)}) - 4\,\frac{a'}{a}\,\phi^{(1)}\partial_i\phi^{(1)} \\
& - 2\psi^{(1)\prime}\partial_i\phi^{(1)} + 4\psi^{(1)\prime}\partial_i\psi^{(1)} + 4\psi^{(1)}\partial_i\psi^{(1)\prime} - 2\,\frac{a''}{a}\,\phi^{(1)}\partial_i\omega^{(1)} \\
& - 2\left(\frac{a'}{a}\right)^2\phi^{(1)}\partial_i\omega^{(1)} - \frac{a'}{a}\,\phi^{(1)\prime}\partial_i\omega^{(1)} - \nabla^2\omega^{(1)}\partial_i\phi^{(1)} \\
& - \partial^k\omega^{(1)}\partial_i\partial_k\phi^{(1)} + \partial^k\phi^{(1)}\partial_i\partial_k\omega^{(1)} - \partial^k\omega^{(1)}\partial_i\partial_k\omega^{(1)\prime} - \frac{a'}{a}\,\nabla^2\omega^{(1)}\partial_i\omega^{(1)} \\
& - \psi^{(1)\prime\prime}\partial_i\omega^{(1)} - 5\,\frac{a'}{a}\,\psi^{(1)\prime}\partial_i\omega^{(1)} - \frac{1}{2}\,\partial^k\phi^{(1)}D^{ik}\chi^{(1)\prime} + \psi^{(1)}\partial_k D^k{}_i\chi^{(1)\prime} \\
& + \psi^{(1)\prime}\partial_k D^k{}_i\chi^{(1)} - \frac{1}{2}\,\partial_k\psi^{(1)}D_i^k\chi^{(1)\prime} + \partial_k\psi^{(1)\prime}D^k{}_i\chi^{(1)} + \frac{a'}{a}\,\partial^k\omega^{(1)}D_{ik}\chi^{(1)\prime} \\
& + \tfrac{1}{2}\,\partial^k\omega^{(1)}D_{ik}\chi^{(1)\prime\prime} - \tfrac{1}{2}\,\partial_k D^{km}\chi^{(1)}D_{mi}\chi^{(1)\prime} - \tfrac{1}{2}\,D^{km}\chi^{(1)}\partial_k D_{mi}\chi^{(1)\prime} \\
& + \tfrac{1}{2}\,D^{km}\chi^{(1)\prime}\partial_i D_{mk}\chi^{(1)} + \tfrac{1}{4}\,D^{km}\chi^{(1)}\partial_i D_{mk}\chi^{(1)\prime}\,. \qquad \text{(A.27)}
\end{aligned}
$$

The expression for the purely spatial part of $\delta^{(2)}\mathscr{R}_{\mu\nu}$ is very long, thus for simplicity we will divide it into two parts: the diagonal part $\delta^{(2)}R_{ij}^d$, proportional to δ_{ij}, and the non-diagonal part R_{ij}^{nd}.

$$
\begin{aligned}
\delta^{(2)}R_{ij}^d = {} & \Bigg[-\left(\frac{a'}{a}\right)^2\phi^{(2)} - \frac{1}{2}\,\frac{a'}{a}\,\phi^{(2)\prime} - \frac{a''}{a}\,\phi^{(2)} - \frac{5}{2}\,\frac{a'}{a}\,\psi^{(2)\prime} - \left(\frac{a'}{a}\right)^2\psi^{(2)} - \frac{1}{2}\,\psi^{(2)\prime\prime} \\
& - \frac{a''}{a}\,\psi^{(2)} + \frac{1}{2}\,\nabla^2\psi^{(2)} - \frac{1}{2}\,\frac{a'}{a}\,\nabla^2\omega^{(2)} + 4\left(\left(\frac{a'}{a}\right)^2 + \frac{a''}{a}\right)(\phi^{(1)})^2 \\
& + 4\,\frac{a'}{a}\,\phi^{(1)}\phi^{(1)\prime} + 10\,\frac{a'}{a}\,\phi^{(1)}\psi^{(1)\prime} + 2\,\frac{a'}{a}\,\phi^{(1)\prime}\psi^{(1)} + \phi^{(1)\prime}\psi^{(1)\prime} + 2\phi^{(1)}\psi^{(1)\prime\prime} \\
& + 4\left(\left(\frac{a'}{a}\right)^2 + \frac{a''}{a}\right)\phi^{(1)}\psi^{(1)} + \partial_k\psi^{(1)}\partial^k\phi^{(1)} + (\psi^{(1)\prime})^2 + \partial_k\psi^{(1)}\partial^k\psi^{(1)} \\
& + 2\psi^{(1)}\nabla^2\psi^{(1)} + \frac{a'}{a}\,\partial_k\phi^{(1)}\partial^k\omega^{(1)} + 2\,\frac{a'}{a}\,\phi^{(1)}\nabla^2\omega^{(1)} - \left(\frac{a'}{a}\right)^2\partial_k\omega^{(1)}\partial^k\omega^{(1)} \\
& - \frac{a''}{a}\,\partial_k\omega^{(1)}\partial^k\omega^{(1)} - \frac{a'}{a}\,\partial_k\omega^{(1)}\partial^k\omega^{(1)\prime} + 3\,\frac{a'}{a}\,\partial_k\omega^{(1)}\partial^k\psi^{(1)} \\
& + 2\partial_k\psi^{(1)\prime}\partial^k\omega^{(1)} + \psi^{(1)\prime}\nabla^2\omega^{(1)} + \partial_k\psi^{(1)}\partial^k\omega^{(1)\prime} - \partial_m\psi^{(1)}\partial_k D^{km}\chi^{(1)} \\
& - \partial_k\partial_m\psi^{(1)}D^{km}\chi^{(1)} + \frac{a'}{a}\,\partial_m\partial^k\omega^{(1)}D_k^m\chi^{(1)} + \frac{a'}{a}\,\partial^k\omega^{(1)}\partial_m D_k^m\chi^{(1)} \\
& - \frac{1}{2}\,\frac{a'}{a}\,D^{mk}\chi^{(1)}D_{km}\chi^{(1)\prime}\Bigg]\delta_{ij}\,, \qquad \text{(A.28)}
\end{aligned}
$$

$$
\begin{aligned}
\delta^{(2)}R^{nd}_{ij} = & -\frac{1}{2}\,\partial_i\partial_j\phi^{(2)} + \frac{1}{2}\,\partial_i\partial_j\psi^{(2)} - \frac{a'}{a}\,\partial_i\partial_j\omega^{(2)} - \frac{1}{2}\,\partial_i\partial_j\omega^{(2)\prime} - \frac{1}{2}\,\frac{a'}{a}\,(\partial_i\omega_j^{(2)} + \partial_j\omega_i^{(2)}) \\
& - \frac{1}{4}\,(\partial_i\omega_j^{(2)\prime} + \partial_j\omega_i^{(2)\prime}) + \frac{1}{2}\left(\left(\frac{a'}{a}\right)^2 + \frac{a''}{a}\right)(D_{ij}\chi^{(2)} + \partial_i\chi_j^{(2)} + \partial_j\chi_i^{(2)} + \chi_{ij}^{(2)}) \\
& + \frac{1}{2}\,\frac{a'}{a}\,(D_{ij}\chi^{(2)\prime} + \partial_i\chi_j^{(2)\prime} + \partial_j\chi_i^{(2)\prime} + (\chi_{ij}^{(2)})') + \frac{1}{2}\,\partial_k\partial_i D^k{}_j\chi^{(2)} - \frac{1}{4}\,\nabla^2 D_{ij}\chi^{(2)} \\
& - \tfrac{1}{4}\,\nabla^2\chi_{ij}^{(2)} + \tfrac{1}{4}(D_{ij}\chi^{(2)\prime\prime} + \partial_i\chi_j^{(2)\prime\prime} + \partial_j\chi_i^{(2)\prime\prime} + (\chi_{ij}^{(2)})'') + \partial_i\phi^{(1)}\partial_j\phi^{(1)} \\
& + 2\phi^{(1)}\partial_i\partial_j\phi^{(1)} - \partial_j\phi^{(1)}\partial_i\psi^{(1)} - \partial_i\phi^{(1)}\partial_j\psi^{(1)} \\
& + 3\partial_i\psi^{(1)}\partial_j\psi^{(1)} + 2\psi^{(1)}\partial_i\partial_j\psi^{(1)} + 4\,\frac{a'}{a}\,\phi^{(1)}\partial_i\partial_j\omega^{(1)} + \phi^{(1)\prime}\partial_i\partial_j\omega^{(1)} \\
& + 2\phi^{(1)}\partial_i\partial_j\omega^{(1)\prime} + \nabla^2\omega^{(1)}\partial_i\partial_j\omega^{(1)} - \partial_j\partial^k\omega^{(1)}\partial_i\partial_k\omega^{(1)} - 2\,\frac{a'}{a}\,\partial_i\psi^{(1)}\partial_j\omega^{(1)} \\
& - 2\,\frac{a'}{a}\,\partial_i\omega^{(1)}\partial_j\psi^{(1)} - \partial_i\psi^{(1)\prime}\partial_j\omega^{(1)} - \partial_j\psi^{(1)\prime}\partial_i\omega^{(1)} - \partial_i\psi^{(1)}\partial_j\omega^{(1)\prime} \\
& - \partial_j\psi^{(1)}\partial_i\omega^{(1)\prime} + \psi^{(1)\prime}\partial_i\partial_j\omega^{(1)} - 2\left(\frac{a'}{a}\right)^2\phi^{(1)}D_{ij}\chi^{(1)} \\
& - 2\,\frac{a''}{a}\,\phi^{(1)}D_{ij}\chi^{(1)} - 2\,\frac{a'}{a}\,\phi^{(1)}D_{ij}\chi^{(1)\prime} - \frac{a'}{a}\,\phi^{(1)\prime}D_{ij}\chi^{(1)} - \frac{1}{2}\phi^{(1)\prime}D_{ij}\chi^{(1)\prime} \\
& - \phi^{(1)}D_{ij}\chi^{(1)\prime\prime} + \tfrac{1}{2}\,\partial_k\phi^{(1)}\partial_i D^k{}_j\chi^{(1)} + \tfrac{1}{2}\partial_k\phi^{(1)}\partial_j D^k{}_i\chi^{(1)} \\
& - \frac{1}{2}\,\partial_k\phi^{(1)}\partial^k D_{ij}\chi^{(1)} - 3\,\frac{a'}{a}\,\psi^{(1)\prime}D_{ij}\chi^{(1)} + \frac{1}{2}\,\psi^{(1)\prime}D_{ij}\chi^{(1)\prime} + \frac{1}{2}\partial_k\psi^{(1)}\partial_i D^k{}_j\chi^{(1)} \\
& + \tfrac{1}{2}\,\partial_k\psi^{(1)}\partial_j D^k{}_i\chi^{(1)} - \tfrac{3}{2}\,\partial_k\psi^{(1)}\partial^k D_{ij}\chi^{(1)} + \psi^{(1)}\partial_k\partial_i D^k{}_j\chi^{(1)} + \psi^{(1)}\partial_k\partial_j D^k{}_i\chi^{(1)} \\
& - \psi^{(1)}\partial_k\partial^k D_{ij}\chi^{(1)} + \partial_i\psi^{(1)}\partial_k D^k{}_j\chi^{(1)} + \partial_j\psi^{(1)}\partial_k D^k{}_i\chi^{(1)} + \partial_k\partial_i\psi^{(1)}D^k{}_j\chi^{(1)} \\
& + \partial_k\partial_j\psi^{(1)}D^k{}_i\chi^{(1)} + \tfrac{1}{2}\,\partial_k\partial_i\omega^{(1)}D^k{}_j\chi^{(1)\prime} + \tfrac{1}{2}\,\partial_k\partial_j\omega^{(1)}D^k{}_i\chi^{(1)\prime} - \tfrac{1}{2}\,\partial_k\partial^k\omega^{(1)}D_{ij}\chi^{(1)\prime} \\
& + \tfrac{1}{2}\,\partial^k\omega^{(1)}\partial_i D_{kj}\chi^{(1)\prime} + \tfrac{1}{2}\,\partial^k\omega^{(1)}\partial_j D_{ki}\chi^{(1)\prime} - \partial^k\omega^{(1)}\partial_k D_{ij}\chi^{(1)\prime} + \tfrac{1}{2}\,\partial^k\omega^{(1)\prime}\partial_i D_{kj}\chi^{(1)} \\
& + \tfrac{1}{2}\,\partial^k\omega^{(1)\prime}\partial_j D_{ki}\chi^{(1)} - \tfrac{1}{2}\,\partial^k\omega^{(1)\prime}\partial_k D_{ij}\chi^{(1)} + \frac{a'}{a}\,\partial^k\omega^{(1)}\partial_i D_{kj}\chi^{(1)} \\
& + \frac{a'}{a}\,\partial^k\omega^{(1)}\partial_j D_{ki}\chi^{(1)} - \frac{a'}{a}\,\partial^k\omega^{(1)}\partial_k D_{ij}\chi^{(1)} - \frac{a'}{a}\,\partial_k\partial^k\omega^{(1)}D_{ij}\chi^{(1)} \\
& - \tfrac{1}{2}\,D_i^k\chi^{(1)\prime}D_{kj}\chi^{(1)\prime} - \tfrac{1}{2}\,\partial_i D_{mj}\chi^{(1)}\partial_k D^{km}\chi^{(1)} - \tfrac{1}{2}\,\partial_j D_{mi}\chi^{(1)}\partial_k D^{km}\chi^{(1)} \\
& + \tfrac{1}{2}\,\partial_m D_{ij}\chi^{(1)}\partial_k D^{km}\chi^{(1)} - \tfrac{1}{2}\,\partial_k\partial_i D_{mj}\chi^{(1)}D^{km}\chi^{(1)} - \tfrac{1}{2}\,\partial_k\partial_j D_{mi}\chi^{(1)}D^{km}\chi^{(1)} \\
& + \tfrac{1}{2}\,\partial_k\partial_m D_{ij}\chi^{(1)}D^{km}\chi^{(1)} + \tfrac{1}{2}\,D^{km}\chi^{(1)}\partial_i\partial_j D_{km}\chi^{(1)} + \tfrac{1}{4}\,\partial_i D^{mk}\chi^{(1)}\partial_j D_{mk}\chi^{(1)}\ . \qquad \text{(A.29)}
\end{aligned}
$$

A.4. The Ricci scalar

At zeroth order the Ricci scalar R reads

$$R = \frac{6}{a^2}\,\frac{a''}{a}\ . \qquad \text{(A.30)}$$

The first-order perturbation of R is

$$\delta^{(1)}R = \frac{1}{a^2}\left(-6\frac{a'}{a}\partial_i\partial^i\omega^{(1)} - 2\partial_i\partial^i\omega^{(1)\prime} - 2\partial_i\partial^i\phi^{(1)} - 6\psi^{(1)\prime\prime} \right.$$
$$\left. -6\frac{a'}{a}\phi^{(1)\prime} - 18\frac{a'}{a}\psi^{(1)\prime} - 12\frac{a''}{a}\phi^{(1)} + 4\partial_i\partial^i\psi^{(1)} + \partial_k\partial^i D^k{}_i\chi^{(1)}\right) . \tag{A.31}$$

At second order we find

$$\begin{aligned}
\delta^{(2)}R = & -\nabla^2\phi^{(2)} - 3\frac{a'}{a}\phi^{(2)\prime} - 6\frac{a''}{a}\phi^{(2)} + 2\nabla^2\psi^{(2)} - 9\frac{a'}{a}\psi^{(2)\prime} - 3\psi^{(2)\prime\prime} - \nabla^2\omega^{(2)\prime} \\
& - 3\frac{a'}{a}\nabla^2\omega^{(2)} + \frac{1}{2}\partial_k\partial_i D^{ki}\chi^{(2)} + 24\frac{a''}{a}(\phi^{(1)})^2 + 2\partial_k\phi^{(1)}\partial^k\phi^{(1)} + 4\phi^{(1)}\nabla^2\phi^{(1)} \\
& + 24\frac{a'}{a}\phi^{(1)}\phi^{(1)\prime} + 6\phi^{(1)\prime}\psi^{(1)\prime} + 36\frac{a'}{a}\phi^{(1)}\psi^{(1)\prime} + 2\partial_k\psi^{(1)}\partial^k\phi^{(1)} - 4\psi^{(1)}\nabla^2\phi^{(1)} \\
& + 12\phi^{(1)}\psi^{(1)\prime\prime} - 12\psi^{(1)}\psi^{(1)\prime\prime} - 36\frac{a'}{a}\psi^{(1)\prime}\psi^{(1)} + 6\partial_k\psi^{(1)}\partial^k\psi^{(1)} + 16\psi^{(1)}\nabla^2\psi^{(1)} \\
& + 6\frac{a'}{a}\partial^k\omega^{(1)}\partial_k\phi^{(1)} + 12\frac{a'}{a}\phi^{(1)}\nabla^2\omega^{(1)} + 4\phi^{(1)}\nabla^2\omega^{(1)\prime} + 2\phi^{(1)\prime}\nabla^2\omega^{(1)} \\
& - 5\frac{a''}{a}\partial_k\omega^{(1)}\partial^k\omega^{(1)} - 6\frac{a'}{a}\partial_k\omega^{(1)}\partial^k\omega^{(1)\prime} + \nabla^2\omega^{(1)}\nabla^2\omega^{(1)} - \partial^i\partial^k\omega^{(1)}\partial_i\partial_k\omega^{(1)} \\
& + 8\partial_k\omega^{(1)}\partial^k\psi^{(1)\prime} + 2\partial_k\omega^{(1)\prime}\partial^k\psi^{(1)} - 4\psi^{(1)}\nabla^2\omega^{(1)\prime} - 12\frac{a'}{a}\psi^{(1)}\nabla^2\omega^{(1)} \\
& + 4\psi^{(1)\prime}\nabla^2\omega^{(1)} + 2\partial_k\phi^{(1)}\partial_i D^{ik}\chi^{(1)} + 2\partial_i\partial_k\phi^{(1)}D^{ik}\chi^{(1)} + 4\psi^{(1)}\partial_k\partial_i D^{ki}\chi^{(1)} \\
& - 2\partial_k\partial_i\psi^{(1)}D^{ik}\chi^{(1)} + 3\partial_k\omega^{(1)}\partial^i D^k{}_i\chi^{(1)\prime} + 6\frac{a'}{a}\partial^k\omega^{(1)}\partial_i D^i{}_k\chi^{(1)} + 2\partial_i\omega^{(1)\prime}\partial_k D^{ik}\chi^{(1)} \\
& + 2\partial_k\partial_i\omega^{(1)\prime}D^{ik}\chi^{(1)} + 6\frac{a'}{a}\partial_k\partial_i\omega^{(1)}D^{ki}\chi^{(1)} - D^{ik}\chi^{(1)}D_{ik}\chi^{(1)\prime\prime} - \tfrac{3}{4}D^{ik}\chi^{(1)\prime}D_{ki}\chi^{(1)\prime} \\
& - 3\frac{a'}{a}D^{ik}\chi^{(1)}D_{ik}\chi^{(1)\prime} - 2\partial_k\partial^i D_{mi}\chi^{(1)}D^{km}\chi^{(1)} + \nabla^2 D_{im}\chi^{(1)}D^{mi}\chi^{(1)} \\
& - \partial_k D^{km}\chi^{(1)}\partial^i D_{mi}\chi^{(1)} + \tfrac{1}{4}\partial^i D^{km}\chi^{(1)}\partial_i D_{mk}\chi^{(1)} .
\end{aligned} \tag{A.32}$$

A.5. The Einstein tensor components

The Einstein tensor in a spatially flat FRW background is given by

$$G^i{}_j = -\frac{3}{a^2}\left(\frac{a'}{a}\right)^2 , \tag{A.33}$$

$$G^i_j = -\frac{1}{a^2}\left(2\frac{a''}{a} - \left(\frac{a'}{a}\right)^2\right)\delta^i_j , \tag{A.34}$$

$$G^0{}_i = G^i{}_0 = 0 . \tag{A.35}$$

The first-order perturbations of the Einstein tensor components are

$$\delta^{(1)}G^0{}_0 = \frac{1}{a^2}\left[6\left(\frac{a'}{a}\right)^2\phi^{(1)} + 6\,\frac{a'}{a}\,\psi^{(1)\prime} + 2\,\frac{a'}{a}\,\nabla^2\omega^{(1)} - 2\nabla^2\psi^{(1)} - \frac{1}{2}\,\partial_k\partial^i D_i^k\chi^{(1)}\right], \tag{A.36}$$

$$\delta^{(1)}G^0{}_i = \frac{1}{a^2}\left(-2\,\frac{a'}{a}\,\partial_i\phi^{(1)} - 2\partial_i\psi^{(1)\prime} - \frac{1}{2}\,\partial_k D^k{}_i\chi^{(1)\prime}\right), \tag{A.37}$$

$$\begin{aligned}\delta^{(1)}G^i{}_j = \frac{1}{a^2}\Bigg[&\left(2\,\frac{a'}{a}\,\phi^{(1)\prime} + 4\,\frac{a''}{a}\,\phi^{(1)} - 2\left(\frac{a'}{a}\right)^2\phi^{(1)} + \nabla^2\phi^{(1)} + 4\,\frac{a'}{a}\,\psi^{(1)\prime}\right.\\ &\left.+2\psi^{(1)\prime\prime} - \nabla^2\psi^{(1)} + 2\,\frac{a'}{a}\,\nabla^2\omega^{(1)} + \nabla^2\omega^{(1)\prime} + \frac{1}{2}\,\partial_k\partial^m D^k{}_m\chi^{(1)}\right)\delta^i{}_j\\ &- \partial^i\partial_j\phi^{(1)} + \partial^i\partial_j\psi^{(1)} - 2\,\frac{a'}{a}\,\partial^i\partial_j\omega^{(1)} - \partial^i\partial_j\omega^{(1)\prime}\\ &+\frac{a'}{a}\,D^i{}_j\chi^{(1)\prime} + \frac{1}{2}\,D^i{}_j\chi^{(1)\prime\prime} + \frac{1}{2}\,\partial_k\partial^i D^k{}_j\chi^{(1)} + \frac{1}{2}\,\partial_k\partial_j D^{ik}\chi^{(1)} - \frac{1}{2}\,\partial_k\partial^k D^i_j\chi^{(1)}\Bigg].\end{aligned} \tag{A.38}$$

The second-order perturbed Einstein tensor components are given by

$$\begin{aligned}\delta^{(2)}G^0{}_0 = \frac{1}{a^2}\Bigg(&3\left(\frac{a'}{a}\right)^2\phi^{(2)} + 3\,\frac{a'}{a}\,\psi^{(2)\prime} - \nabla^2\psi^{(2)} + \frac{a'}{a}\,\nabla^2\omega^{(2)} - \frac{1}{4}\,\partial_k\partial_i D^{ki}\chi^{(2)}\\ &- 12\left(\frac{a'}{a}\right)^2(\phi^{(1)})^2 - 12\,\frac{a'}{a}\,\phi^{(1)}\psi^{(1)\prime} - 3\partial_i\psi^{(1)}\partial^i\psi^{(1)}\\ &- 8\psi^{(1)}\nabla^2\psi^{(1)} + 12\,\frac{a'}{a}\,\psi^{(1)}\psi^{(1)\prime} - 3(\psi^{(1)\prime})^2 + 4\,\frac{a'}{a}\,\phi^{(1)}\nabla^2\omega^{(1)}\\ &- 2\,\frac{a'}{a}\,\partial_k\omega^{(1)}\partial^k\phi^{(1)} - \frac{1}{2}\,\frac{a''}{a}\,\partial_k\omega^{(1)}\partial^k\omega^{(1)}\\ &+ \tfrac{1}{2}\,\partial_i\partial_k\omega^{(1)}\partial^i\partial^k\omega^{(1)} - \tfrac{1}{2}\,\partial_k\partial^k\omega^{(1)}\partial_k\partial^k\omega^{(1)}\\ &- 2\,\frac{a'}{a}\,\partial_k\psi^{(1)}\partial^k\omega^{(1)} + 4\,\frac{a'}{a}\,\psi^{(1)}\nabla^2\omega^{(1)}\\ &- 2\partial_k\omega^{(1)}\partial^k\psi^{(1)\prime} - 2\psi^{(1)\prime}\nabla^2\omega^{(1)} - \phi^{(1)}\partial_i\partial^k D^i{}_k\chi^{(1)} - 2\psi^{(1)}\partial_k\partial^i D^k{}_i\chi^{(1)}\\ &+ \partial_k\partial_i\psi^{(1)}D^{ki}\chi^{(1)} - 2\,\frac{a'}{a}\,\partial_i\partial_k\omega^{(1)}D^{ik}\chi^{(1)}\\ &- 2\,\frac{a'}{a}\,\partial_k\omega^{(1)}\partial_i D^{ik}\chi^{(1)} - \partial_k\omega^{(1)}\partial^i D^k{}_i\chi^{(1)\prime}\\ &- \tfrac{1}{2}\,\nabla^2 D_{mk}\chi^{(1)}D^{km}\chi^{(1)} + \partial_m\partial^k D_{ik}\chi^{(1)}D^{im}\chi^{(1)} + \tfrac{1}{2}\,\partial_k D^{km}\chi^{(1)}\partial^i D_{mi}\chi^{(1)}\\ &-\frac{1}{8}\,\partial^i D^{km}\chi^{(1)}\partial_i D_{km}\chi^{(1)} + \frac{1}{8}\,D^{ik}\chi^{(1)\prime}D_{ki}\chi^{(1)\prime} + \frac{a'}{a}\,D^{ki}\chi^{(1)}D_{ik}\chi^{(1)\prime}\Bigg),\end{aligned} \tag{A.39}$$

$$
\begin{aligned}
\delta^{(2)}G^{i}{}_{0} = \frac{1}{a^2}\Bigg(&\frac{a'}{a}\,\partial^i\phi^{(2)} + \partial^i\psi^{(2)\prime} + \frac{1}{4}\,\partial_k D^{ki}\chi^{(2)\prime} + \frac{1}{4}\,\nabla^2\chi^{i\,(2)\prime} - \frac{1}{4}\,\nabla^2\omega^{i(2)}\\
&- \frac{a''}{a}\,\partial^i\omega^{(2)} - \frac{a''}{a}\,\omega^{i(2)} + 2\left(\frac{a'}{a}\right)^2\partial^i\omega^{(2)} + 2\left(\frac{a'}{a}\right)^2\omega^{i(2)}\\
&- 4\,\frac{a'}{a}\,\phi^{(1)}\partial^i\phi^{(1)} + 4\,\frac{a'}{a}\,\psi^{(1)}\partial^i\phi^{(1)} - 2\psi^{(1)\prime}\partial^i\phi^{(1)}\\
&+ 4\psi^{(1)\prime}\partial^i\psi^{(1)} + 8\psi^{(1)}\partial^i\psi^{(1)\prime} - \partial^i\phi^{(1)}\nabla^2\omega^{(1)} - \partial^k\omega^{(1)}\partial^i\partial_k\phi^{(1)}\\
&+ \nabla^2\phi^{(1)}\partial^i\omega^{(1)} + \partial^i\partial_k\omega^{(1)}\partial^k\phi^{(1)} + 4\,\frac{a''}{a}\,\phi^{(1)}\partial^i\omega^{(1)} - 8\left(\frac{a'}{a}\right)^2\phi^{(1)}\partial^i\omega^{(1)}\\
&+ 2\,\frac{a'}{a}\,\phi^{(1)\prime}\partial^i\omega^{(1)} + \nabla^2\omega^{(1)\prime}\partial^i\omega^{(1)} - \partial^k\omega^{(1)}\partial^i\partial_k\omega^{(1)\prime} + 2\psi^{(1)\prime\prime}\partial^i\omega^{(1)}\\
&+ 8\left(\frac{a'}{a}\right)^2\psi^{(1)}\partial^i\omega^{(1)} - 4\,\frac{a''}{a}\,\psi^{(1)}\partial^i\omega^{(1)} - 2\,\frac{a'}{a}\,\psi^{(1)\prime}\partial^i\omega^{(1)} - \frac{1}{2}\,\partial^k\phi^{(1)}\,D^i{}_k\chi^{(1)\prime}\\
&- 2\,\frac{a'}{a}\,\partial_k\phi^{(1)}D^{ki}\chi^{(1)} - \tfrac{1}{2}\,\partial_k\psi^{(1)}D^{ki}\chi^{(1)\prime} + 2\psi^{(1)}\partial_k D^{ki}\chi^{(1)\prime} + \psi^{(1)\prime}\partial_k D^{ki}\chi^{(1)}\\
&- \partial_k\psi^{(1)\prime}D^{ki}\chi^{(1)} + \tfrac{1}{2}\,\partial^k\omega^{(1)}\,D^i{}_k\chi^{(1)\prime\prime}\\
&+ \frac{a'}{a}\,\partial^k\omega^{(1)}\,D^i{}_k\chi^{(1)\prime} - 4\left(\frac{a'}{a}\right)^2\partial_k\omega^{(1)}D^{ik}\chi^{(1)}\\
&+ 2\,\frac{a''}{a}\,\partial_k\omega^{(1)}D^{ik}\chi^{(1)} - \frac{1}{2}\,\partial_k D^{km}\chi^{(1)}D^i_m\chi^{(1)\prime} - \frac{1}{2}\,\partial_k D^i{}_m\chi^{(1)\prime}D^{km}\chi^{(1)}\\
&+ \tfrac{1}{4}\,\partial^i D_{mk}\chi^{(1)}D^{km}\chi^{(1)\prime} + \tfrac{1}{2}\,\partial^i D_{mk}\chi^{(1)\prime}D^{km}\chi^{(1)} - \tfrac{1}{2}\,D^{ik}\chi^{(1)}\partial_m D^m{}_k\chi^{(1)\prime}\Bigg)\,, \qquad \text{(A.40)}
\end{aligned}
$$

$$
\begin{aligned}
\delta^{(2)}G^{0}{}_{i} = \frac{1}{a^2}\Bigg(&-\frac{a'}{a}\,\partial_i\phi^{(2)} - \partial_i\psi^{(2)\prime} - \frac{1}{4}\partial_k D^k{}_i\chi^{(2)\prime} - \frac{1}{4}\,\nabla^2\chi_i^{(2)\prime} + \frac{1}{4}\,\nabla^2\omega_i^{(2)}\\
&+ 8\,\frac{a'}{a}\,\phi^{(1)}\partial_i\phi^{(1)} + 4\phi^{(1)}\partial_i\psi^{(1)\prime} + 2\psi^{(1)\prime}\partial_i\phi^{(1)} - 4\psi^{(1)\prime}\partial_i\psi^{(1)} - 4\psi^{(1)}\partial_i\psi^{(1)\prime}\\
&+ \partial_i\phi^{(1)}\nabla^2\omega^{(1)} - \partial_i\partial_k\omega^{(1)}\partial^k\phi^{(1)} + 8\,\frac{a''}{a}\,\phi^{(1)}\partial_i\omega^{(1)} - 4\left(\frac{a'}{a}\right)^2\phi^{(1)}\partial_i\omega^{(1)}\\
&- 2\,\frac{a'}{a}\,\partial^k\omega^{(1)}\partial_i\partial_k\omega^{(1)} + \nabla^2\psi^{(1)}\partial_i\omega^{(1)} + \partial^k\omega^{(1)}\partial_i\partial_k\psi^{(1)} - \partial_k\phi^{(1)}D^k{}_i\chi^{(1)\prime}\\
&+ \tfrac{1}{2}\,\partial^k\phi^{(1)}D_{ik}\chi^{(1)\prime} - \psi^{(1)}\partial_k D^k{}_i\chi^{(1)\prime} + \tfrac{1}{2}\,\partial_k\psi^{(1)}\,D^k{}_i\chi^{(1)\prime} - \psi^{(1)\prime}\partial_k D^k{}_i\chi^{(1)}\\
&- \partial_k\psi^{(1)\prime}D^k{}_i\chi^{(1)} + \partial_i\omega^{(1)}\partial_k\partial^m D^k{}_m\chi^{(1)}\\
&- 2\,\frac{a''}{a}\,\partial^k\omega^{(1)}D_{ik}\chi^{(1)} + \left(\frac{a'}{a}\right)^2\partial^k\omega^{(1)}D_{ik}\chi^{(1)} + \partial^k\omega^{(1)}\partial_m\partial_i D^m_k\chi^{(1)}\\
&- \tfrac{1}{2}\,\partial^m\omega^{(1)}\partial_k\partial^k D_{im}\chi^{(1)} + \tfrac{1}{2}\,\partial_k D^{km}\chi^{(1)}D_{im}\chi^{(1)\prime} + \tfrac{1}{2}\,\partial_k D_{im}\chi^{(1)\prime}D^{km}\chi^{(1)}\\
&- \tfrac{1}{4}\,\partial_i D_{mk}\chi^{(1)}D^{km}\chi^{(1)\prime} - \tfrac{1}{2}\,\partial_i D_{mk}\chi^{(1)\prime}D^{km}\chi^{(1)}\Bigg)\,, \qquad \text{(A.41)}
\end{aligned}
$$

$$
\begin{aligned}
\delta^{(2)}G^{d\,i}{}_{j} = \frac{1}{a^2}\Bigg(&\frac{1}{2}\nabla^2\phi^{(2)} + \frac{a'}{a}\phi^{(2)\prime} + 2\frac{a''}{a}\phi^{(2)} - \left(\frac{a'}{a}\right)^2\phi^{(2)} - \frac{1}{2}\nabla^2\psi^{(2)} + \psi^{(2)\prime\prime}\\
&+ 2\frac{a'}{a}\psi^{(2)\prime} + \frac{a'}{a}\nabla^2\omega^{(2)} + \frac{1}{2}\nabla^2\omega^{(2)\prime} - \frac{1}{4}\partial_k\partial_i D^{ki}\chi^{(2)} + 4\left(\frac{a'}{a}\right)^2(\phi^{(1)})^2\\
&- 8\frac{a''}{a}(\phi^{(1)})^2 - 8\frac{a'}{a}\phi^{(1)}\phi^{(1)\prime} - \partial_k\phi^{(1)}\partial^k\phi^{(1)} - 2\phi^{(1)}\nabla^2\phi^{(1)} - 4\phi^{(1)}\psi^{(1)\prime\prime}\\
&- 2\phi^{(1)\prime}\psi^{(1)\prime} - 8\frac{a'}{a}\phi^{(1)}\psi^{(1)\prime} - 2\partial_k\psi^{(1)}\partial^k\psi^{(1)} - 4\psi^{(1)}\nabla^2\psi^{(1)} + (\psi^{(1)\prime})^2\\
&+ 8\frac{a'}{a}\psi^{(1)}\psi^{(1)\prime} + 4\psi^{(1)}\psi^{(1)\prime\prime} + 2\psi^{(1)}\nabla^2\phi^{(1)} - \phi^{(1)\prime}\nabla^2\omega^{(1)}\\
&- 2\phi^{(1)}\nabla^2\omega^{(1)\prime} - 2\frac{a'}{a}\partial_k\omega^{(1)}\partial^k\phi^{(1)} - 4\frac{a'}{a}\phi^{(1)}\nabla^2\omega^{(1)} + \frac{3}{2}\frac{a''}{a}\partial_k\omega^{(1)}\partial^k\omega^{(1)}\\
&- \left(\frac{a'}{a}\right)^2\partial_k\omega^{(1)}\partial^k\omega^{(1)} + 2\frac{a'}{a}\partial_k\omega^{(1)}\partial^k\omega^{(1)\prime} - \frac{1}{2}\nabla^2\omega^{(1)}\nabla^2\omega^{(1)}\\
&+ \frac{1}{2}\partial^m\partial^k\omega^{(1)}\partial_m\partial_k\omega^{(1)} + 4\frac{a'}{a}\psi^{(1)}\nabla^2\omega^{(1)} + 2\psi^{(1)}\nabla^2\omega^{(1)\prime}\\
&- 2\partial_k\omega^{(1)}\partial^k\psi^{(1)\prime} - \psi^{(1)\prime}\nabla^2\omega^{(1)} - \partial_k\partial_m\phi^{(1)}D^{km}\chi^{(1)} - \partial_k\phi^{(1)}\partial_m D^{mk}\chi^{(1)}\\
&- \partial_k\psi^{(1)}\partial_m D^{mk}\chi^{(1)} - \tfrac{3}{2}\partial_k\omega^{(1)}\partial^i D^k{}_i\chi^{(1)\prime} - \partial_k\omega^{(1)\prime}\partial_m D^{mk}\chi^{(1)}\\
&- \partial_k\partial_m\omega^{(1)\prime}D^{km}\chi^{(1)} - 2\frac{a'}{a}\partial^k\omega^{(1)}\partial_m D^m_k\chi^{(1)} - 2\frac{a'}{a}\partial_m\partial^k\omega^{(1)}D^m{}_k\chi^{(1)}\\
&+ \tfrac{3}{4}\partial_k\partial^l D_{ml}\chi^{(1)}D^{km}\chi^{(1)} - \tfrac{1}{2}\nabla^2 D_{ml}\chi^{(1)}D^{ml}\chi^{(1)} + \tfrac{1}{4}\partial_m\partial^k D_{lk}\chi^{(1)}D^{lm}\chi^{(1)}\\
&+ \tfrac{1}{2}\partial_k D_{km}\chi^{(1)}\partial^l D^{ml}\chi^{(1)} - \tfrac{1}{8}\partial^l D_{km}\chi^{(1)}\partial_l D^{km}\chi^{(1)}\\
&+ \frac{1}{2}D^{mk}\chi^{(1)}D_{mk}\chi^{(1)\prime\prime} + \frac{3}{8}D^{mk}\chi^{(1)\prime}D_{mk}\chi^{(1)\prime} + \frac{a'}{a}D^{mk}\chi^{(1)}D_{km}\chi^{(1)\prime}\Bigg)\delta^i_j. \qquad \text{(A.42)}
\end{aligned}
$$

$$
\begin{aligned}
\delta^{(2)}G^{nd^1}{}_{j} = \frac{1}{a^2}\Bigg[&-\frac{1}{2}\partial^i\partial_j\phi^{(2)} + \frac{1}{2}\partial^i\partial_j\psi^{(2)} - \frac{a'}{a}\partial^i\partial_j\omega^{(2)} - \frac{1}{2}\partial^i\partial_j\omega^{(2)\prime}\\
&- \frac{1}{2}\frac{a'}{a}(\partial^i\omega^{(2)}_j + \partial_j\omega^{i(2)}) - \frac{1}{4}(\partial^i\omega^{(2)\prime}_j + \partial_j\omega^{i(2)\prime})\\
&+ \frac{1}{2}\frac{a'}{a}(D^i_j\chi^{(2)\prime} + \partial^i\chi^{(2)\prime}_j + \partial_j\chi^{i(2)\prime} + \chi^{i(2)\prime}_j) + \frac{1}{2}\partial_k\partial^i D^k{}_j\chi^{(2)}\\
&- \tfrac{1}{4}\nabla^2 D^i{}_j\chi^{(2)} - \tfrac{1}{4}\nabla^2\chi^{i(2)}{}_j + \tfrac{1}{4}(D^i{}_j\chi^{(2)\prime\prime} + \partial^i\chi^{(2)\prime\prime}_j + \partial_j\chi^{i(2)\prime\prime} + \chi^{i(2)\prime\prime}_j)
\end{aligned}
$$

$$
\begin{aligned}
&+\partial^i\phi^{(1)}\partial_j\phi^{(1)}+2\phi^{(1)}\partial^i\partial_j\phi^{(1)}-2\psi^{(1)}\partial^i\partial_j\phi^{(1)}\\
&-\partial_j\phi^{(1)}\partial^i\psi^{(1)}-\partial^i\phi^{(1)}\partial_j\psi^{(1)}\\
&+3\partial^i\psi^{(1)}\partial_j\psi^{(1)}+4\psi^{(1)}\partial^i\partial_j\psi^{(1)}+2\frac{a'}{a}\partial^i\omega^{(1)}\partial_j\phi^{(1)}+4\frac{a'}{a}\phi^{(1)}\partial^i\partial_j\omega^{(1)}\\
&+\phi^{(1)'}\partial^i\partial_j\omega^{(1)}+2\phi^{(1)}\partial^i\partial_j\omega^{(1)'}+\nabla^2\omega^{(1)}\partial^i\partial_j\omega^{(1)}-\partial_j\partial^k\omega^{(1)}\partial^i\partial_k\omega^{(1)}\\
&-2\frac{a'}{a}\partial^i\psi^{(1)}\partial_j\omega^{(1)}-2\frac{a'}{a}\partial^i\omega^{(1)}\partial_j\psi^{(1)}-\partial^i\psi^{(1)'}\partial_j\omega^{(1)}+\partial_j\psi^{(1)'}\partial^i\omega^{(1)}\\
&-\partial^i\psi^{(1)}\partial_j\omega^{(1)'}-\partial_j\psi^{(1)}\partial^i\omega^{(1)'}-2\psi^{(1)}\partial^i\partial_j\omega^{(1)'}+\psi^{(1)'}\partial^i\partial_j\omega^{(1)}\\
&-4\frac{a'}{a}\psi^{(1)}\partial^i\partial_j\omega^{(1)}-2\frac{a'}{a}\phi^{(1)}D^i_j\chi^{(1)'}-\frac{1}{2}\phi^{(1)'}D^i_j\chi^{(1)'}-\phi^{(1)}D^i_j\omega^{(1)''}\\
&+\tfrac{1}{2}\partial_k\phi^{(1)}\partial^i D^k{}_j\chi^{(1)}+\tfrac{1}{2}\partial_k\phi^{(1)}\partial_j D^{ki}\chi^{(1)}-\tfrac{1}{2}\partial_k\phi^{(1)}\partial^k D^i_j\chi^{(1)}+\partial_j\partial_k\phi^{(1)}D^{ki}\chi^{(1)}\\
&+\frac{1}{2}\psi^{(1)'}D^i_j\chi^{(1)'}+\psi^{(1)''}D^i_j\chi^{(1)}+2\frac{a'}{a}\psi^{(1)'}D^i_j\chi^{(1)}+\frac{1}{2}\partial_k\psi^{(1)}\partial^i D^k{}_j\chi^{(1)}\\
&+2\frac{a'}{a}\psi^{(1)}D^i_j\chi^{(1)'}+\psi^{(1)}D^i_j\omega^{(1)''}+\frac{1}{2}\partial_k\psi^{(1)}\partial_j D^{ki}\chi^{(1)}-\frac{3}{2}\partial_k\psi^{(1)}\partial^k D^i_j\chi^{(1)}\\
&+2\psi^{(1)}\partial_k\partial^i D^k{}_j\chi^{(1)}+2\psi^{(1)}\partial_k\partial_j D^{ki}\chi^{(1)}-2\psi^{(1)}\partial_k\partial^k D^i_j\chi^{(1)}-\nabla^2\psi^{(1)}D^i_j\chi^{(1)}\\
&+\partial^i\psi^{(1)}\partial_k D^k{}_j\chi^{(1)}+\partial_j\psi^{(1)}\partial_k D^{ki}\chi^{(1)}+\partial_k\partial^i\psi^{(1)}D^k{}_j\chi^{(1)}+\tfrac{1}{2}\partial^i\omega^{(1)}\partial_k D^k{}_j\chi^{(1)'}\\
&+\tfrac{1}{2}\partial_k\partial^i\omega^{(1)}D^k{}_j\chi^{(1)'}+\tfrac{1}{2}\partial_k\partial_j\omega^{(1)}D^{ki}\chi^{(1)'}-\tfrac{1}{2}\partial_k\partial^k\omega^{(1)}D^i_j\chi^{(1)'}+\tfrac{1}{2}\partial^k\omega^{(1)}\partial^i D_{kj}\chi^{(1)'}\\
&+\tfrac{1}{2}\partial^k\omega^{(1)}\partial_j D^i_k\chi^{(1)'}-\partial^k\omega^{(1)}\partial_k D^i_j\chi^{(1)'}+\tfrac{1}{2}\partial^k\omega^{(1)'}\partial^i D_{kj}\chi^{(1)}+\tfrac{1}{2}\partial^k\omega^{(1)'}\partial_j D^i_k\chi^{(1)}\\
&-\tfrac{1}{2}\partial^k\omega^{(1)'}\partial_k D^i_j\chi^{(1)}+\partial_k\partial_j\omega^{(1)'}D^{ik}\chi^{(1)}+\frac{a'}{a}\partial^k\omega^{(1)}\partial^i D_{kj}\chi^{(1)}+\frac{a'}{a}\partial^k\omega^{(1)}\partial_j D^i_k\chi^{(1)}\\
&-\frac{a'}{a}\partial^k\omega^{(1)}\partial_k D^i_j\chi^{(1)}+2\frac{a'}{a}\partial_k\partial_j\omega^{(1)}D^{ik}\chi^{(1)}-\frac{1}{2}D^{ki}\chi^{(1)'}D_{kj}\chi^{(1)'}\\
&-\tfrac{1}{2}\partial^i D_{mj}\chi^{(1)}\partial_k D^{km}\chi^{(1)}-\tfrac{1}{2}\partial_j D^i_m\chi^{(1)}\partial_k D^{km}\chi^{(1)}+\tfrac{1}{2}\partial_m D^i_j\chi^{(1)}\partial_k D^{km}\chi^{(1)}\\
&-\tfrac{1}{2}\partial_k\partial^i D_{mj}\chi^{(1)}D^{km}\chi^{(1)}-\tfrac{1}{2}\partial_k\partial_j D^i_m\chi^{(1)}D^{km}\chi^{(1)}+\tfrac{1}{2}\partial_k\partial_m D^i_j\chi^{(1)}D^{km}\chi^{(1)}\\
&+\tfrac{1}{2}D^{km}\chi^{(1)}\partial^i\partial_j D_{km}\chi^{(1)}+\tfrac{1}{4}\partial^i D^{mk}\chi^{(1)}\partial_j D_{mk}\chi^{(1)}-\partial_k\partial^m D^k_m\chi^{(1)}D^i_j\chi^{(1)}\\
&-\frac{a'}{a}D_{kj}\chi^{(1)'}D^{ik}\chi^{(1)}-\frac{1}{2}D_{kj}\omega^{(1)''}D^{ki}\chi^{(1)}-\partial_m\partial_k D^m_j\chi^{(1)}D^{ki}\chi^{(1)}\\
&+\tfrac{1}{2}\partial_m\partial^m D_{kj}\chi^{(1)}D^{ki}\chi^{(1)}\Big]\,,
\end{aligned}
\tag{A.43}
$$

where $\delta^{(2)}G^{d^i}_{\ j}$ stands for the diagonal part of $\delta^{(2)}G^i{}_j$, which is proportional to $\delta^i{}_j$, and $\delta^{(2)}G^{nd^i}_{\ j}$ is the non-diagonal contribution.

Appendix B. Perturbing the Klein–Gordon equation

In the homogeneous background the Klein–Gordon equation for the scalar field φ is

$$\varphi_0'' + 2\,\frac{a'}{a}\varphi_0' = -\frac{\partial V}{\partial\varphi}\,a^2\ . \tag{B.1}$$

The perturbed Klein–Gordon equation at first order is

$$\begin{aligned}&\delta^{(1)}\varphi'' + 2\,\frac{a'}{a}\,\delta^{(1)}\varphi' - \nabla^2\delta^{(1)}\varphi - \phi^{(1)'}\varphi_0' - 3\psi^{(1)'}\varphi_0' - \nabla^2\omega^{(1)}\varphi_0'\\ &\quad = -\delta^{(1)}\varphi\,\frac{\partial^2 V}{\partial\varphi^2}\,a^2 - 2\phi^{(1)}\,\frac{\partial V}{\partial\varphi}\ .\end{aligned} \tag{B.2}$$

At second order we get

$$\begin{aligned}&-\frac{1}{2}\,\delta^{(2)}\varphi'' - \frac{a'}{a}\,\delta^{(2)}\varphi' + \frac{1}{2}\,\nabla^2\delta^{(2)}\varphi + \phi^{(2)}\varphi_0'' + 2\,\frac{a'}{a}\,\phi^{(2)}\varphi_0' + \frac{1}{2}\,\phi^{(2)'}\varphi_0'\\ &+\frac{3}{2}\,\psi^{(2)'}\varphi_0' + \frac{1}{2}\,\nabla^2\omega^{(2)}\varphi_0' - 4(\phi^{(1)})^2\varphi_0'' - 8\,\frac{a'}{a}\,(\phi^{(1)})^2\varphi_0' - 4\phi^{(1)}\phi^{(1)'}\varphi_0'\\ &+2\phi^{(1)}\delta^{(1)}\varphi'' + \phi^{(1)'}\delta^{(1)}\varphi' + 4\,\frac{a'}{a}\,\phi^{(1)}\delta^{(1)}\varphi' + \partial^k\phi^{(1)}\partial_k\delta^{(1)}\varphi - 6\phi^{(1)}\psi^{(1)'}\varphi_0'\\ &+6\psi^{(1)}\psi^{(1)'}\varphi_0' + 3\psi^{(1)'}\delta^{(1)}\varphi' - \partial^k\psi^{(1)}\partial_k\delta^{(1)}\varphi + 2\psi^{(1)}\nabla^2\delta^{(1)}\varphi\\ &-2\phi^{(1)}\nabla^2\omega^{(1)}\varphi_0' - \partial^k\omega^{(1)}\partial_k\phi^{(1)}\varphi_0' - \partial^k\omega^{(1)}\partial_k\psi^{(1)}\varphi_0' + 2\psi^{(1)}\nabla^2\omega^{(1)}\varphi_0'\\ &+\partial^k\omega^{(1)}\partial_k\omega^{(1)}\varphi_0'' + 2\,\frac{a'}{a}\,\partial^k\omega^{(1)}\partial_k\omega^{(1)}\varphi_0' + \partial_k\omega^{(1)}\partial^k\omega^{(1)'}\varphi_0' + 2\partial^k\omega^{(1)}\partial_k\delta^{(1)}\varphi'\\ &+2\,\frac{a'}{a}\,\partial^k\omega^{(1)}\partial_k\delta^{(1)}\varphi + \nabla^2\omega^{(1)}\delta^{(1)}\varphi' + \partial^k\omega^{(1)'}\partial_k\delta^{(1)}\varphi - \partial^k\omega^{(1)}\partial_i D^i{}_k\chi^{(1)}\varphi_0'\\ &-\partial_i\partial_k\omega^{(1)}D^{ik}\chi^{(1)}\varphi_0' - \partial_i\partial_k\delta^{(1)}\varphi D^{ik}\chi^{(1)} - \partial_k\delta^{(1)}\varphi\partial^i D^k{}_i\chi^{(1)}\\ &+\frac{1}{2}\,D^{ik}\chi^{(1)}D_{ki}\chi^{(1)'}\varphi_0' = \frac{1}{2}\,\frac{\partial^2 V}{\partial\varphi^2}\,\delta^{(2)}\varphi a^2 + \frac{1}{2}\,\frac{\partial^3 V}{\partial\varphi^3}(\delta^{(1)}\varphi)^2 a^2\ .\end{aligned} \tag{B.3}$$

To obtain the Klein–Gordon equation in the Poisson gauge one can simply set $\omega^{(1)}=\omega^{(2)}=0$, $\chi^{(1)}=\chi^{(2)}=0$, and $\phi^{(1)}=\psi^{(1)}$. Thus at first order we find

$$\delta^{(1)}\varphi'' + 2\,\frac{a'}{a}\,\delta^{(1)}\varphi' - \nabla^2\delta^{(1)}\varphi - 4\phi^{(1)'}\varphi_0' = -\delta^{(1)}\varphi\frac{\partial^2 V}{\partial\varphi^2}a^2 - 2\phi^{(1)}\,\frac{\partial V}{\partial\varphi}\ , \tag{B.4}$$

while at second order the equation is

$$\begin{aligned}&\frac{1}{2}\,\delta^{(2)}\varphi'' + \frac{a'}{a}\,\delta^{(2)}\varphi' - \frac{1}{2}\,\nabla^2\delta^{(2)}\varphi - \phi^{(2)}\left(\varphi_0'' + 2\,\frac{a'}{a}\,\varphi_0'\right)\\ &-\tfrac{1}{2}\,\phi^{(2)'}\varphi_0' - \tfrac{3}{2}\,\psi^{(2)'}\varphi_0' - 4\phi^{(1)}\phi^{(1)'}\varphi_0' - 4\phi^{(1)'}\delta^{(1)}\varphi'\\ &-4\phi^{(1)}\nabla^2\delta^{(1)}\varphi = -2\phi^{(1)}\delta^{(1)}\varphi\,\frac{\partial^2 V}{\partial\varphi^2}\,a^2 - \frac{1}{2}\,\delta^{(2)}\varphi\,\frac{\partial^2 V}{\partial\varphi^2}\,a^2 - \frac{1}{2}\,(\delta^{(1)}\varphi)^2\,\frac{\partial^3 V}{\partial\varphi^3}\,a^2\ ,\end{aligned} \tag{B.5}$$

where we have used the background equation (B.1) and the first-order perturbed equation (B.4) to simplify some terms.

Appendix C. Wigner 3-*j* symbol

In this appendix, we summarize basic properties of the Wigner 3-*j* symbol, following Ref. [245]. The Wigner 3-*j* symbol characterizes geometric properties of the angular bispectrum.

C.1. Triangle conditions

The Wigner 3-*j* symbol,

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} , \tag{C.1}$$

is related to the Clebsh–Gordan coefficients which describe coupling of two angular momenta in the quantum mechanics. In the quantum mechanics, l is the eigenvalue of the angular momentum operator, $\mathbf{L}=\mathbf{r}\times\mathbf{p}$: $\mathbf{L}^2 Y_{lm}=l(l+1)Y_{lm}$. m is the eigenvalue of the z-direction component of the angular momentum, $L_z Y_{lm} = mY_{lm}$.

The symbol such as

$$(-1)^{m_3}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \tag{C.2}$$

describes coupling of two angular-momentum states, $\mathbf{L}_1$ and $\mathbf{L}_2$, forming a coupled state, $\mathbf{L}_3=\mathbf{L}_1+\mathbf{L}_2$. It follows from $\mathbf{L}_1+\mathbf{L}_2-\mathbf{L}_3=0$ that $m_1+m_2-m_3=0$; thus, the Wigner 3-*j* symbol (C.1) describes three angular momenta forming a triangle, $\mathbf{L}_1+\mathbf{L}_2+\mathbf{L}_3=0$, and satisfies $m_1+m_2+m_3=0$.

Since $\mathbf{L}_1$, $\mathbf{L}_2$, and $\mathbf{L}_3$ form a triangle, they have to satisfy the triangle conditions, $|L_i-L_j|\leqslant L_k\leqslant L_i+L_j$, where $L_i\equiv|\mathbf{L}_i|$. Hence, l_1, l_2, and l_3 also satisfy the triangle conditions,

$$|l_i-l_j|\leqslant l_k\leqslant l_i+l_j; \tag{C.3}$$

otherwise, the Wigner 3-*j* symbol vanishes. The triangle conditions also include $m_1+m_2+m_3=0$. These properties may regard (l, m) as vectors, $\mathbf{l}$, which satisfy $\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3=0$. Note that, however, $\mathbf{L}\neq\mathbf{l}$.

For $l_1=l_2$ and $l_3=m_3=0$, the Wigner 3-*j* symbol reduces to

$$(-1)^m\begin{pmatrix} l & l & 0 \\ m & -m & 0 \end{pmatrix} = \frac{(-1)^l}{\sqrt{2l+1}} . \tag{C.4}$$

In Section 10.1, we have used this relation to reduce the covariance matrix of the angular bispectrum and trispectrum. We have also used this relation to reduce the angular trispectrum for $L=0$ (see Eq. (438)).

C.2. Symmetry

The Wigner 3-*j* symbol is invariant under even permutations,

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} l_3 & l_1 & l_2 \\ m_3 & m_1 & m_2 \end{pmatrix} = \begin{pmatrix} l_2 & l_3 & l_1 \\ m_2 & m_3 & m_1 \end{pmatrix} , \tag{C.5}$$

while it changes the phase for odd permutations if $l_1 + l_2 + l_3 = \text{odd}$,

$$(-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \tag{C.6}$$

$$= \begin{pmatrix} l_2 & l_1 & l_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = \begin{pmatrix} l_1 & l_3 & l_2 \\ m_1 & m_3 & m_2 \end{pmatrix} = \begin{pmatrix} l_3 & l_2 & l_1 \\ m_3 & m_2 & m_1 \end{pmatrix} . \tag{C.7}$$

The phase also changes under the transformation of $m_1 + m_2 + m_3 \to -(m_1 + m_2 + m_3)$, if $l_1 + l_2 + l_3 = \text{odd}$,

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} . \tag{C.8}$$

If there is no z-direction component of the angular momenta in the system, i.e., $m_i = 0$, then the Wigner 3-j symbol of the system,

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} , \tag{C.9}$$

is non-zero only if $l_1 + l_2 + l_3 = \text{even}$. This symbol is invariant under any permutations of l_i.

In Section 10.2, we have frequently used the Gaunt integral, $\mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3}$, defined by

$$\mathscr{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} \equiv \int \mathrm{d}^2\hat{\mathbf{n}} Y_{l_1 m_1}(\hat{\mathbf{n}}) Y_{l_2 m_2}(\hat{\mathbf{n}}) Y_{l_3 m_3}(\hat{\mathbf{n}})$$

$$= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} , \tag{C.10}$$

to calculate the angular bispectrum. By definition, the Gaunt integral is invariant under both the odd and the even permutations, and non-zero only if $l_1 + l_2 + l_3 = \text{even}$, $m_1 + m_2 + m_3 = 0$, and $|l_i - l_j| \leqslant l_k \leqslant l_i + l_j$. In other words, the Gaunt integral describes fundamental geometric properties of the angular bispectrum such as the triangle conditions.

The Gaunt integral for $m_i = 0$ gives the identity for the Legendre polynomials,

$$\int_{-1}^{1} \frac{\mathrm{d}x}{2} P_{l_1}(x) P_{l_2}(x) P_{l_3}(x) = \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 . \tag{C.11}$$

In Section 10.1, we have used this identity to derive the bias for the angular bispectrum on the incomplete sky (Eq. (452)). Here, we have used

$$Y_{l0}(\hat{\mathbf{n}}) = \sqrt{\frac{4\pi}{2l+1}} P_l(\cos\theta) . \tag{C.12}$$

C.3. Orthogonality

The Wigner 3-j symbol has the following orthogonality properties:

$$\sum_{l_3 m_3} (2l_3 + 1) \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1' & m_2' & m_3 \end{pmatrix} = \delta_{m_1 m_1'} \delta_{m_2 m_2'} , \tag{C.13}$$

and

$$\sum_{m_1 m_2} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l'_3 \\ m_1 & m_2 & m'_3 \end{pmatrix} = \frac{\delta_{l_3 l'_3} \delta_{m_3 m'_3}}{2l_3 + 1} \ , \tag{C.14}$$

or

$$\sum_{\text{all } m} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 = 1 \ . \tag{C.15}$$

The orthogonality properties are essential for any basic calculations involving the Wigner 3-*j* symbols. Note that these orthogonality properties are consistent with orthonormality of the angular-momentum eigenstate vectors, and unitarity of the Clebsh–Gordan coefficients, by definition.

C.4. Rotation matrix

A finite rotation operator for the Euler angles α, β, and γ, $D(\alpha, \beta, \gamma)$, comprises angular momentum operators,

$$D(\alpha, \beta, \gamma) = \mathrm{e}^{-\mathrm{i}\alpha L_z} \mathrm{e}^{-\mathrm{i}\beta L_y} \mathrm{e}^{-\mathrm{i}\gamma L_z} \ . \tag{C.16}$$

Since the Wigner 3-*j* symbol describes coupling of two angular momenta, it also describes coupling of two rotation operators. Using the rotation matrix element, $D^{(l)}_{m'm} = \langle l, m' | D | l, m \rangle$, we have

$$D^{(l_1)}_{m'_1 m_1} D^{(l_2)}_{m'_2 m_2} = \sum_{l_3} (2l_3 + 1) \sum_{m_3 m'_3} D^{(l_3)*}_{m'_3 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} \ . \tag{C.17}$$

In Section 10.1, we have used this relation to evaluate rotationally invariant harmonic spectra. Note that the rotation matrix is orthonormal,

$$\sum_m D^{(l)*}_{m'm} D^{(l)}_{m''m} = \delta_{m'm''} \ . \tag{C.18}$$

C.5. Wigner 6-j symbol

The Wigner 6-*j* symbol,

$$\begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \end{Bmatrix} \ , \tag{C.19}$$

describes coupling of three angular momenta. We often encounter the Wigner 6-*j* symbol, when we calculate the angular bispectrum which has more complicated geometric structures (see, e.g., Ref. [92], Appendix C of Ref. [137]). The angular trispectrum also often includes the Wigner 6-*j* symbol [114].

The Wigner 6-*j* symbol is related to the Wigner 3-*j* symbols through

$$\begin{aligned} (-1)^{l'_1 + l'_2 + l'_3} & \begin{Bmatrix} l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \end{Bmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \\ & = \sum_{\text{all } m'} (-1)^{m'_1 + m'_2 + m'_3} \begin{pmatrix} l_1 & l'_2 & l'_3 \\ m_1 & m'_2 & -m'_3 \end{pmatrix} \begin{pmatrix} l'_1 & l_2 & l'_3 \\ -m'_1 & m_2 & m'_3 \end{pmatrix} \begin{pmatrix} l'_1 & l'_2 & l_3 \\ m'_1 & -m'_2 & m_3 \end{pmatrix} \ . \end{aligned} \tag{C.20}$$

By using Eq. (C.15), we also obtain

$$(-1)^{l_1'+l_2'+l_3'}\begin{Bmatrix} l_1 & l_2 & l_3 \\ l_1' & l_2' & l_3' \end{Bmatrix} = \sum_{\text{all } mm'} (-1)^{m_1'+m_2'+m_3'} \times \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} l_1 & l_2' & l_3' \\ m_1 & m_2' & -m_3' \end{pmatrix} \times \begin{pmatrix} l_1' & l_2 & l_3' \\ -m_1' & m_2 & m_3' \end{pmatrix}\begin{pmatrix} l_1' & l_2' & l_3 \\ m_1' & -m_2' & m_3 \end{pmatrix} . \tag{C.21}$$

References

[1] L.F. Abbott, E. Farhi, M.B. Wise, Phys. Lett. B 117 (1982) 29.
[2] L.F. Abbott, M.B. Wise, Nucl. Phys. B 244 (1984) 541.
[3] V. Acquaviva, N. Bartolo, S. Matarrese, A. Riotto, Nucl. Phys. B 667 (2003) 119.
[4] N. Afshordi, R. Brandenberger, Phys. Rev. D 63 (2001) 123505.
[5] N. Aghanim, P.G. Castro, O. Forni, M. Kunz, Astron. Astrophys. 406 (2003) 797.
[6] N. Aghanim, O. Forni, Astron. Astrophys. 347 (1999) 409.
[7] N. Aghanim, O. Forni, F.R. Bouchet, Astron. Astrophys. 365 (2001) 341.
[8] A. Albrecht, P.J. Steinhardt, M.S. Turner, F. Wilczek, Phys. Rev. Lett. 48 (1982) 1437.
[9] M. Alishahiha, E. Silverstein, D. Tong, arXiv:hep-th/0404084.
[10] R. Allahverdi, Phys. Rev. D 70 (2004) 043507.
[11] T.J. Allen, B. Grinstein, M.B. Wise, Phys. Lett. B 197 (1987) 66.
[12] L. Amendola, C. Gordon, D. Wands, M. Sasaki, Phys. Rev. Lett. 88 (2002) 211302.
[13] F. Argüeso, J. González-Nuevo, L. Toffolatti, Astrophys. J. 598 (2003) 86.
[14] N. Arkani-Hamed, H.C. Cheng, M.A. Luty, S. Mukohyama, JHEP 0405 (2004) 074.
[15] N. Arkani-Hamed, P. Creminelli, S. Mukohyama, M. Zaldarriaga, JCAP 0404 (2004) 001.
[16] F. Atrio-Barandela, J.P. Mücket, Astrophys. J. 515 (1999) 465.
[17] V. Avila-Reese, P. Colin, G. Piccinelli, C. Firmani, Astrophys. J. 598 (2003) 36.
[18] D. Babich, P. Creminelli, M. Zaldarriaga, JCAP 0408 (2004) 009.
[19] A.J. Banday, S. Zaroubi, K.M. Górski, Astrophys. J. 533 (2000) 575.
[20] J.M. Bardeen, Phys. Rev. D 22 (1980) 1882.
[21] J.M. Bardeen, in: A. Zee (Ed.), Particle Physics and Cosmology, Gordon and Breach, New York, 1989.
[22] J.M. Bardeen, P.J. Steinhardt, M.S. Turner, Phys. Rev. D 28 (1983) 679.
[23] V. Barger, H.S. Lee, D. Marfatia, Phys. Lett. B 565 (2003) 33.
[24] C. Barnes, et al., Astrophys. J. 148 (Suppl.) (2003) 51.
[25] R.B. Barreiro, et al., Mon. Not. Roy. Astron. Soc. 322 (2000) 411.
[26] J.D. Barrow, P. Coles, Mon. Not. Roy. Astron. Soc. 244 (1990) 188.
[27] N. Bartolo, P.S. Corasaniti, A.R. Liddle, M. Malquarti, Phys. Rev. D 70 (2004) 043532.
[28] N. Bartolo, S. Matarrese, A. Riotto, Phys. Rev. D 64 (2001) 083514.
[29] N. Bartolo, S. Matarrese, A. Riotto, Phys. Rev. D 64 (2001) 123504.
[30] N. Bartolo, S. Matarrese, A. Riotto, Phys. Rev. D 65 (2002) 103505.
[31] N. Bartolo, S. Matarrese, A. Riotto, JHEP 0404 (2004) 006.
[32] N. Bartolo, S. Matarrese, A. Riotto, Phys. Rev. D 69 (2004) 043503.
[33] N. Bartolo, S. Matarrese, A. Riotto, JCAP 0401 (2004) 003.
[34] N. Bartolo, S. Matarrese, A. Riotto, arXiv:astro-ph/0407505.
[35] B.A. Bassett, F. Tamburini, D.I. Kaiser, R. Marteens, Nucl. Phys. B 561 (1999) 188.
[36] C.L. Bennett, et al., Astrophys. J. 464 (1996) L1.
[38] C.L. Bennett, et al., Astrophys. J. 148 (Suppl.) (2003) 1.
[39] C.L. Bennett, et al., Astrophys. J. 148 (Suppl.) (2003) 97.

[40] A. Berera, Phys. Rev. Lett. 75 (1995) 3218.
[41] A. Berera, Phys. Rev. D 54 (1996) 2519.
[42] K. Benabed, F. Bernardeau, Phys. Rev. D 61 (2000) 123510.
[43] F. Bernardeau, Astron. Astrophys. 324 (1997) 15.
[44] F. Bernardeau, J.-P. Uzan, Phys. Rev. D 66 (2002) 103506.
[45] F. Bernardeau, J.-P. Uzan, Phys. Rev. D 67 (2003) 121301.
[46] F. Bernardeau, T. Brunier, J.-P. Uzan, Phys. Rev. D 69 (2004) 063520.
[47] F. Bernardeau, S. Colombi, E. Gaztanaga, R. Scoccimarro, Phys. Rep. 367 (2002) 1.
[48] E. Bertschinger, Cosmology and large-scale structure, in: R. Shaeffer, J. Silk, M. Spiro, V. Zinn-Justin (Eds.), Proceedings of the Les Houches School, vol. LX, Elsevier, Netherland, 1996.
[49] N. Bevis, M. Hindmarsh, M. Kunz, Phys. Rev. D 70 (2004) 043508.
[50] S.M. Bilenky, Lectures given at the 1999 European School of High Energy Physics, Casta Papiernicka, Slovakia, August 22–September 4, 1999, arXiv:hep-ph/0001311.
[51] B.C. Bromley, M. Tegmark, Astrophys. J. 524 (1999) L79.
[52] M. Bruni, L. Gualtieri, C. Sopuerta, Class. Quant. Grav. 20 (2003) 535.
[53] M. Bruni, S. Matarrese, S. Mollerach, S. Sonego, Class. Quant. Grav. 14 (1997) 2585.
[54] M. Bucher, Y. Zhu, Phys. Rev. D 55 (1997) 7415.
[55] T.S. Bunch, P.C.W. Davies, Proc. Roy. Soc. A 360 (1978) 117.
[56] E.F. Bunn, M. White, Astrophys. J. 480 (1997) 6.
[57] P. Cabella, F.K. Hansen, D. Marinucci, D. Pagano, N. Vittorio, Phys. Rev. D 69 (2004) 063007.
[58] P. Cabella, M. Liguori, F.K. Hansen, D. Marinucci, S. Matarrese, L. Moscardini, N. Vittorio, arXiv:astro-ph/0406026.
[59] L. Cayón, et al., Mon. Not. Roy. Astron. Soc. 339 (2003) 1189.
[60] R. Cen, J.P. Ostriker, Astrophys. J. 514 (1999) 1.
[61] X. Chen, A. Cooray, N. Yoshida, N. Sugiyama, Mon. Not. Roy. Astron. Soc. 346 (2003) L31.
[62] L.Y. Chiang, P.D. Naselsky, O.V. Verkhodanov, M.J. Way, Astrophys. J. 590 (2003) L65.
[63] S. Cole, N. Kaiser, Mon. Not. Roy. Astron. Soc. 233 (1988) 637.
[64] P. Coles, J.D. Barrow, Mon. Not. Roy. Astron. Soc. 228 (1987) 407.
[65] P. Coles, P. Dineen, J. Earl, D. Wright, Mon. Not. Roy. Astron. Soc. 350 (2004) 983.
[66] P. Coles, F. Lucchin, Cosmology, The Origin and Evolution of Cosmic Structure, Wiley, Chichester, 1995.
[67] C.R. Contaldi, P.G. Ferreira, J. Magueijo, K.M. Górski, Astrophys. J. 534 (2000) 25.
[68] A. Cooray, Phys. Rev. D 62 (2000) 103506.
[69] A. Cooray, Phys. Rev. D 65 (2002) 083518.
[70] A. Cooray, W. Hu, Astrophys. J. 534 (2000) 533.
[71] C.J. Copi, D. Huterer, G.D. Starkman, arXiv:astro-ph/0310511.
[72] P. Creminelli, JCAP 0310 (2003) 003.
[73] P. Creminelli, M. Zaldarriaga, arXiv:astro-ph/0405428.
[74] S. Dodelson, W.H. Kinney, E.W. Kolb, Phys. Rev. D 56 (1997) 3207.
[75] A.D. Dolgov, A.D. Linde, Phys. Lett. B 116 (1982) 329.
[76] G. Dvali, A. Gruzinov, M. Zaldarriaga, Phys. Rev. D 69 (2004) 023505.
[77] G. Dvali, A. Gruzinov, M. Zaldarriaga, Phys. Rev. D 69 (2004) 083505.
[78] G.F.R. Ellis, M. Bruni, Phys. Rev. D 40 (1989) 1804.
[79] K. Enqvist, M.S. Sloth, Nucl. Phys. B 626 (2002) 395.
[80] K. Enqvist, A. Vaihkonen, arXiv:hep-ph/0405103
[81] H.K. Eriksen, F.K. Hansen, A.J. Banday, K.M. Górski, P.B. Lilje, Astrophys. J. 605 (2004) 14.
[82] T. Falk, R. Rangarajan, M. Srednicki, Astrophys. J. 403 (1993) L1.
[83] Z. Fan, J.M. Bardeen, preprint UW-PT-92-11, 1992, unpublished.
[84] P.G. Ferreira, J. Magueijo, K.M. Górski, Astrophys. J. 503 (1998) L1.
[85] A. Gangui, Phys. Rev. D 50 (1994) 3684.
[86] A. Gangui, F. Lucchin, S. Matarrese, S. Mollerach, Astrophys. J. 430 (1994) 447.
[87] A. Gangui, J. Martin, Mon. Not. Roy. Astron. Soc. 313 (2000) 323.
[88] A. Gangui, J. Martin, Phys. Rev. D 62 (2000) 103004.
[89] A. Gangui, J. Martin, M. Sakellariadou, Phys. Rev. D 66 (2002) 083502.

[90] J. Garcia-Bellido, D. Wands, Phys. Rev. D 53 (1996) 5437.
[91] E. Gaztanaga, J. Wagg, Phys. Rev. D 68 (2003) 021302.
[92] D.M. Goldberg, D.N. Spergel, Phys. Rev. D 59 (1999) 103002.
[93] C. Gordon, D. Wands, B.A. Bassett, R. Maartens, Phys. Rev. D 63 (2001) 023506.
[94] K.M. Górski, et al., Astrophys. J. 464 (1996) L11.
[95] J.R. III Gott, Astrophys. J. 288 (1985) 422.
[96] J. R. III Gott, et al., Astrophys. J. 352 (199) 1
[97] S. Groot Nibbelink, B.J. van Tent, Class. Quant. Grav. 19 (2002) 613.
[98] S. Gupta, arXiv:astro-ph/0310460.
[99] S. Gupta, A. Berera, A.F. Heavens, S. Matarrese, Phys. Rev. D 66 (2002) 043510.
[100] S. Gupta, K.A. Malik, D. Wands, Phys. Rev. D 69 (2004) 063513.
[101] A. Guth, Phys. Rev. D 23 (1981) 347.
[102] A. Guth, S.Y. Pi, Phys. Rev. Lett. 49 (1982) 1110.
[103] T. Hamazaki, H. Kodama, Prog. Theor. Phys. 96 (1996) 1123.
[104] F.K. Hansen, P. Cabella, D. Marinucci, N. Vittorio, Astrophys. J. 607 (2004) L67.
[105] F.K. Hansen, D. Marinucci, P. Natoli, N. Vittorio, Phys. Rev. D 66 (2002) 063006.
[106] F.K. Hansen, D. Marinucci, N. Vittorio, Phys. Rev. D 67 (2003) 123004.
[107] S.W. Hawking, Phys. Lett. B 115 (1982) 295.
[108] A.F. Heavens, Mon. Not. Roy. Astron. Soc. 299 (1998) 805.
[109] A.F. Heavens, S. Gupta, Mon. Not. Roy. Astron. Soc. 324 (2001) 960.
[110] G. Hinshaw, et al., Astrophys. J. 148 (Suppl.) (2003) 135.
[111] G. Hinshaw, et al., Astrophys. J. 148 (Suppl.) (2003) 63.
[112] H.M. Hodges, G.R. Blumenthal, L.A. Kofman, J.R. Primack, Nucl. Phys. B 335 (1990) 197.
[113] W. Hu, Phys. Rev. D 62 (2000) 043007.
[114] W. Hu, Phys. Rev. D 64 (2001) 083005.
[115] W. Hu, N. Sugiyama, Astrophys. J. 444 (1995) 489.
[116] J.c. Hwang, H. Noh, Phys. Lett. B 495 (2000) 277.
[117] J.c. Hwang, H. Noh, Class. Quant. Grav. 19 (2002) 527.
[118] T.K. Inoue, Exploring topology of the universe in the cosmic microwave background, Ph.D. Thesis, Kyoto University, 2001.
[119] N. Jarosik, et al., Astrophys. J. 148 (Suppl.) (2003) 29.
[120] N. Kaiser, A. Stebbins, Nature 310 (1984) 391.
[121] M. Kamionkowski, A. Kosowsky, A. Stebbins, Phys. Rev. Lett. 78 (1997) 2058.
[122] M. Kesden, A. Cooray, M. Kamionkowski, Phys. Rev. Lett. 89 (2002) 011304.
[123] W.H. Kinney, E.W. Kolb, A. Melchiorri, A. Riotto, Phys. Rev. D 69 (2004) 103516.
[124] T. Kitayama, Y. Suto, Astrophys. J. 490 (1997) 557.
[125] L. Knox, Phys. Rev. D 48 (1995) 3502.
[126] L. Knox, Y. -Song, Phys. Rev. Lett. 89 (2002) 011303.
[127] H. Kodama, T. Hamazaki, Prog. Theor. Phys. 96 (1996) 949.
[128] H. Kodama, M. Sasaki, Prog. Theor. Phys. 78 (Suppl.) (1984) 1.
[129] L.A. Kofman, Phys. Lett. B 173 (1986) 400.
[130] L.A. Kofman, arXiv:astro-ph/0303614.
[131] L.A. Kofman, G.R. Blumenthal, H. Hodges, J.R. Primack, Large-scale structures and peculiar motions in the universe, in: D.W. Latham, L.N. deCosta (Eds.), ASP Conference Series, vol. 15, 1991, p. 339.
[132] L.A. Kofman, A. Linde, Nucl. Phys. B 282 (1987) 555.
[133] L. Kofman, A.D. Linde, A.A. Starobinsky, Phys. Rev. Lett. 73 (1994) 3195.
[134] A. Kogut, et al., Astrophys. J. 464 (1996) L5.
[135] A. Kogut, et al., Astrophys. J. 464 (1996) L29.
[136] A. Kogut, et al., Astrophys. J. 148 (Suppl.) (2003) 161.
[137] E. Komatsu, The pursuit of non-Gaussian fluctuations in the cosmic microwave background, Ph.D. Thesis, Tohoku University, arXiv:astro-ph/0206039, 2001.
[138] E. Komatsu, et al., Publ. Astron. Soc. Japan 53 (2001) 57.

[139] E. Komatsu, et al., Astrophys. J. 148 (Suppl.) (2003) 119.
[140] E. Komatsu, T. Kitayama, Astrophys. J. 526 (1999) L1.
[141] E. Komatsu, U. Seljak, Mon. Not. Roy. Astron. Soc. 327 (2001) 1353.
[142] E. Komatsu, U. Seljak, Mon. Not. Roy. Astron. Soc. 336 (2002) 1256.
[143] E. Komatsu, D.N. Spergel, Phys. Rev. D 63 (2001) 063002.
[144] E. Komatsu, D.N. Spergel, B.D. Wandelt, arXiv:astro-ph/0305189.
[145] E. Komatsu, B.D. Wandelt, D.N. Spergel, A.J. Banday, K.M. Górski, Astrophys. J. 566 (2002) 19.
[146] A. Kosowsky, New Astronomy Reviews 47 (2004) 939.
[147] K. Koyama, J. Soda, A. Taruya, Mon. Not. Roy. Astron. Soc. 310 (1999) 1111.
[148] M. Kunz, et al., Astrophys. J. 563 (2001) L99.
[149] M. Landriau, E.P.S. Shellard, Phys. Rev. D 67 (2003) 103512.
[150] M. Landriau, E.P.S. Shellard, Phys. Rev. D 69 (2004) 023003.
[151] M. Landriau, E.P.S. Shellard, arXiv:astro-ph/0310229
[152] D. Langlois, Phys. Rev. D 59 (1999) 123512.
[153] D.L. Larson, B.D. Wandelt, arXiv:astro-ph/0404037.
[154] S.M. Leach, A.R. Liddle, Phys. Rev. D 68 (2003) 123508.
[155] J. Lesgourgues, D. Polarski, A.A. Starobinsky, Nucl. Phys. B 497 (1997) 479.
[156] A.R. Liddle, D.H. Lyth, Phys. Rep. 231 (1993) 1.
[157] A.R. Liddle, D.H. Lyth, Cosmological Inflation and the Large-Scale Structure, Cambridge University Press, Cambridge, 2000.
[158] A.R. Liddle, P. Parsons, J.D. Barrow, Phys. Rev. D 50 (1994) 7222.
[159] J.E. Lidsey, A.R. Liddle, E.W. Kolb, E.J. Copeland, T. Barreiro, M. Abney, Rev. Mod. Phys. 69 (1997) 373.
[160] M. Liguori, S. Matarrese, L. Moscardini, Astrophys. J. 597 (2003) 57.
[161] E. Komatsu, M. Liguori, S. Matarrese, A. Riotto, 2004, in preparation.
[162] A.D. Linde, Phys. Lett. B 108 (1982) 389.
[163] A.D. Linde, Phys. Lett. B 116 (1982) 335.
[164] A.D. Linde, JETP Lett. 40 (1984) 1333.
[165] A.D. Linde, Rep. Prog. Phys. 47 (1984) 925.
[166] A.D. Linde, Phys. Lett. B 158 (1985) 375.
[167] A.D. Linde, Particle Physics and Inflationary Cosmology, Contemporary Concepts in Physics, vol. 5, Harwood Academic, Switzerland, 1990.
[168] A.D. Linde, V. Mukhanov, Phys. Rev. D 56 (1997) 535.
[169] F. Lucchin, S. Matarrese, Astrophys. J. 330 (1988) 535.
[170] V. Lukash, Zh. Eksp. Teor. Fiz. 79 (1980) 1601 [Sov. Phys. JETP 52 (1980) 807].
[171] X.c. Luo, Astrophys. J. 427 (1994) L71.
[172] X.c. Luo, D.N. Schramm, Phys. Rev. Lett. 71 (1993) 1124.
[173] D.H. Lyth, Phys. Rev. D 31 (1985) 1792.
[174] D.H. Lyth, A. Riotto, Phys. Rep. 314 (1999) 1.
[175] D.H. Lyth, C. Ungarelli, D. Wands, Phys. Rev. D 67 (2003) 023503.
[176] D.H. Lyth, D. Wands, Phys. Lett. B 524 (2002) 5.
[177] D.H. Lyth, D. Wands, Phys. Rev. D 68 (2003) 103515.
[178] C.P. Ma, E. Bertschinger, Astrophys. J. 455 (1995) 1.
[179] J. Magueijo, Astrophys. J. 528 (2000) L57.
[180] N. Makino, Y. Suto, Astrophys. J. 405 (1993) 1.
[181] J.M. Maldacena, Adv. Theor. Math. Phys. 2 (1998) 231 [Int. J. Theor. Phys. 38 (1999) 1113].
[182] J. Maldacena, JHEP 0305 (2003) 013.
[183] K.A. Malik, Ph.D. Thesis, arXiv:astro-ph/0101563.
[184] K.A. Malik, D. Wands, gr-qc/9804046.
[185] K.A. Malik, D. Wands, Class. Quant. Grav. 21 (2004) L65.
[186] K.A. Malik, D. Wands, C. Ungarelli, Phys. Rev. D 67 (2003) 063516.
[187] J. Martin, A. Riazuelo, M. Sakellariadou, Phys. Rev. D 61 (2000) 083518.
[188] S. Matarrese, S. Mollerach, M. Bruni, Phys. Rev. D 58 (1998) 043504.

[189] S. Matarrese, M.A. Musso, A. Riotto, JCAP 0405 (2004) 008.
[190] S. Matarrese, O. Pantano, D. Sáez, Phys. Rev. Lett. 72 (1994) 320.
[191] S. Matarrese, O. Pantano, D. Sáez, Mon. Not. Roy. Astron. Soc. 271 (1994) 513.
[192] S. Matarrese, A. Riotto, JCAP 0308 (2003) 007.
[193] S. Matarrese, L. Verde, R. Jimenez, Astrophys. J. 541 (2000) 10.
[194] H. Minkowski, Mathematische Annalen 57 (1903) 447.
[195] S. Mollerach, Phys. Rev. D 42 (1990) 313.
[196] S. Mollerach, A. Gangui, F. Lucchin, S. Matarrese, Astrophys. J. 453 (1995) 1.
[197] S. Mollerach, D. Harari, S. Matarrese, Phys. Rev. D 69 (2004) 063002.
[198] S. Mollerach, S. Matarrese, Phys. Rev. D 45 (1992) 1961.
[199] S. Mollerach, S. Matarrese, Phys. Rev. D 56 (1997) 4494.
[200] S. Mollerach, S. Matarrese, A. Ortolan, F. Lucchin, Phys. Rev. D 44 (1991) 1670.
[201] T. Moroi, T. Takahashi, Phys. Lett. B 522 (2001) 215 [Erratum-ibid. B 539 (2002) 303].
[202] L. Moscardini, M. Bartelmann, S. Matarrese, P. Andreani, Mon. Not. Roy. Astron. Soc. 335 (2002) 984.
[203] L. Moscardini, S. Borgani, P. Coles, F. Lucchin, S. Matarrese, A. Messina, M. Plionis, Astrophys. J. 413 (1993) L55.
[204] L. Moscardini, S. Matarrese, F. Lucchin, A. Messina, Mon. Not. Roy. Astron. Soc. 248 (1991) 424.
[205] V.F. Mukhanov, Zh. Éksp. Teor. Fiz. 94 (1988) 1 [Sov. Phys. JETP 68 (1988) 1297].
[206] V.F. Mukhanov, G.V. Chibisov, JETP Lett. 33 (1981) 532.
[207] V.F. Mukhanov, H.A. Feldman, R.H. Brandenberger, Phys. Rep. 215 (1992) 203.
[208] P. Mukherjee, M.P. Hobson, A.N. Lasenby, Mon. Not. Roy. Astron. Soc. 318 (2000) 1157.
[209] P. Mukherjee, Y. Wang, arXiv:astro-ph/0402602.
[210] H.F. Muller, C. Schmid, arXiv:gr-qc/9401020 (unpublished).
[211] D. Munshi, T. Souradeep, A.A. Starobinsky, Astrophys. J. 454 (1995) 552.
[212] K. Nakamura, Prog. Theor. Phys. 110 (2003) 723.
[213] H. Noh, J.c. Hwang, Phys. Rev. D 69 (2004) 104011.
[214] D. Novikov, J. Schmalzing, V.F. Mukhanov, Astron. Astrophys. 364 (2000) 17.
[215] T. Okamoto, W. Hu, Phys. Rev. D 66 (2002) 063008.
[216] L. Page, et al., Astrophys. J. 148 (Suppl.) (2003) 233.
[217] L. Page, et al., Astrophys. J. 148 (Suppl.) (2003) 39.
[218] J. Pando, D. Valls-Gabaud, L.-Z. Fang, Phys. Rev. Lett. 81 (1988) 4568.
[219] C.G. Park, Mon. Not. Roy. Astron. Soc. 349 (2004) 313.
[220] C.-G. Park, C. Park, B. Ratra, M. Tegmark, Astrophys. J. 556 (2001) 582.
[221] P.J.E. Peebles, The Large-Scale Structure of the Universe, Princeton University Press, Princeton, 1980.
[222] P.J.E. Peebles, Astrophys. J. 483 (1997) L1.
[223] P.J.E. Peebles, Astrophys. J. 510 (1999) 523.
[224] P.J.E. Peebles, Astrophys. J. 510 (1999) 531.
[225] H.V. Peiris, et al., Astrophys. J. 148 (Suppl.) (2003) 213.
[226] F.M. Persi, D.N. Spergel, R. Cen, J.P. Ostriker, Astrophys. J. 442 (1995) 1.
[227] N.G. Phillips, A. Kogut, Astrophys. J. 548 (2001) 540.
[228] E. Pierpaoli, J. Garcia-Bellido, S. Borgani, JHEP 9910 (1999) 015.
[229] L. Pilo, A. Riotto, A. Zaffaroni, arXiv:astro-ph/0401302.
[230] L. Pogosian, S.H.H. Tye, I. Wasserman, M. Wyman, Phys. Rev. D 68 (2003) 023506.
[231] D. Polarsky, A.A. Starobinsky, Phys. Rev. D 50 (1994) 6123.
[232] G. Polenta, et al., Astrophys. J. 572 (2002) L27.
[233] T. Pyne, M. Birkinshaw, Astrophys. J. 415 (1993) 459.
[234] T. Pyne, S.M. Carroll, Phys. Rev. D 53 (1996) 2920.
[235] L. Randall, R. Sundrum, Phys. Rev. Lett. 83 (1999) 4690.
[236] M. Rees, D.W. Sciama, Nature 217 (1968) 511.
[237] A. Refregier, E. Komatsu, D.N. Spergel, U.-L. Pen, Phys. Rev. D 60 (2000) 123001.
[238] A. Refregier, D.N. Spergel, T. Herbig, Astrophys. J. 531 (2000) 31.
[239] A. Refregier, R. Teyssier, Phys. Rev. D 66 (2002) 043002.
[240] G.I. Rigopoulos, Class. Quant. Grav. 21 (2004) 1737.

[241] G.I. Rigopoulos, E.P. Shellard, Phys. Rev. D 68 (2003) 123518.
[242] G.I. Rigopoulos, E.P.S. Shellard, arXiv:astro-ph/0405185.
[243] A. Riotto, arXiv:hep-ph/0210162, Lectures delivered at the "ICTP Summer School on Astroparticle Physics and Cosmology", Trieste, 17 June–5 July 2002.
[244] J. Robinson, E. Gawiser, J. Silk, Astrophys. J. 532 (2000) 1.
[245] M. Rotenberg, R. Bivins, N. Metropolis, J.K. Wooten Jr., The 3-*j* and 6-*j* Symbols, The Technology Press, Massachusetts Institute of Technology, MA, 1959.
[246] R.K. Sachs, in: C. DeWitt, B. DeWitt (Eds.), Relativity, Groups, and Topology, Gordon and Breach, New York, 1964.
[247] R.K. Sachs, A.M. Wolfe, Astrophys. J. 147 (1967) 73.
[248] D.S. Salopek, Phys. Rev. D 45 (1992) 1139.
[249] D.S. Salopek, J.R. Bond, Phys. Rev. D 42 (1990) 3936.
[250] D.S. Salopek, J.R. Bond, Phys. Rev. D 43 (1991) 10005.
[251] H.B. Sandvik, J. Magueijo, Mon. Roy. Astron. Soc. 325 (2001) 463.
[252] M.G. Santos, et al., Mon. Roy. Astron. Soc. 341 (2003) 623.
[253] M. Sasaki, Prog. Theor. Phys. 76 (1986) 1036.
[254] R. Scherrer, Astrophys. J. 390 (1992) 330.
[255] J. Schmalzing, K.M. Górski, Mon. Not. Roy. Astron. Soc. 297 (1998) 355.
[256] R. Scoccimarro, E. Sefusatti, M. Zaldarriaga, Phys. Rev. D 69 (2004) 103513.
[257] U. Seljak, Astrophys. J. 463 (1996) 1.
[258] U. Seljak, J. Burwell, U.-L. Pen, Phys. Rev. D 63 (2001) 063001.
[259] U. Seljak, C.M. Hirata, Phys. Rev. D 69 (2004) 043005.
[260] U. Seljak, M. Zaldarriaga, Astrophys. J. 469 (1996) 437.
[261] U. Seljak, M. Zaldarriaga, Phys. Rev. Lett. 78 (1997) 2054.
[262] L. Senatore, arXiv:astro-ph/0406187
[263] S.F. Shandarin, H.A. Feldman, Y. Xu, M. Tegmark, Astrophys. J. 141 (Suppl.) (2002) 1.
[264] E. Silverstein, D. Tong, arXiv:hep-th/0310221.
[265] S. Smith, et al., Mon. Not. Roy. Astron. Soc., arXiv:astro-ph/0401618, in press.
[266] G.F. Smoot, et al., Astrophys. J. 396 (1992) L1.
[267] S. Sonego, M. Bruni, Comm. Math. Phys. 193 (1998) 209.
[268] D.N. Spergel, D.M. Goldberg, Phys. Rev. D 59 (1999) 103001.
[269] D.N. Spergel, et al., Astrophys. J. 148 (Suppl.) (2003) 175.
[270] V. Springel, M. White, L. Hernquist, Astrophys. J. 549 (2001) 681.
[271] M. Srednicki, Astrophys. J. 416 (1993) L1.
[272] A.A. Starobinsky, JETP Lett. 130 (1980) 682.
[273] A.A. Starobinsky, Phys. Lett. B 117 (1982) 175.
[274] A.A. Starobinsky, JETP Lett. 42 (1985) 152.
[275] A.A. Starobinsky, in: H.J. De Vega, N. Sanchez (Eds.), Field Theory Quantum Gravity and Strings, vol. 107, 1986.
[276] J.M. Stewart, M. Walker, Proc. R. Soc. London A 341 (1974) 49.
[277] J.M. Stewart, Class. Quant. Grav. 7 (1990) 1169.
[278] A. Taruya, Y. Nambu, Phys. Lett. B 428 (1998) 37.
[279] L. Toffolatti, et al., Mon. Not. Roy. Astron. Soc. 297 (1998) 117.
[280] K. Tomita, Prog. Theor. Phys. 45 (1971) 1747.
[281] K. Tomita, Prog. Theor. Phys. 47 (1972) 416.
[282] R. Trotta, A. Riazuelo, R. Durrer, Phys. Rev. Lett. 87 (2001) 231301.
[283] M.S. Turner, Phys. Rev. D 28 (1983) 1243.
[284] L. Verde, R. Jimenez, M. Kamionkowski, S. Matarrese, Mon. Not. Roy. Astron. Soc. 325 (2001) 412.
[285] L. Verde, L. Wang, A.F. Heavens, M. Kamionkowski, Mon. Not. Roy. Astron. Soc. 313 (2000) 141.
[286] L. Verde, et al., Astrophys. J. 148 (Suppl.) (2003) 195.
[287] P. Vielva, E. Martinez-Gonzalez, R.B. Barreiro, J.L. Sanz, L. Cayon, arXiv:astro-ph/0310273.
[288] B.D. Wandelt, E. Hivon, K.M. Górski, Phys. Rev. D 64 (2000) 083003.
[289] D. Wands, N. Bartolo, S. Matarrese, A. Riotto, Phys. Rev. D 66 (2002) 043520.
[290] D. Wands, K.A. Malik, D.H. Lyth, A.R. Liddle, Phys. Rev. D 62 (2000) 043527.

[291] L. Wang, M. Kamionkowski, Phys. Rev. D 61 (2000) 063504.
[292] D.H. Weinberg, S. Cole, Mon. Not. Roy. Astron. Soc. 259 (1992) 652.
[293] J.A. Willick, Astrophys. J. 530 (2000) 80.
[294] J.H.P. Wu, et al., Phys. Rev. Lett. 87 (2001) 1303.
[295] K. Yamamoto, M. Nagasawa, M. Sasaki, H. Suzuki, Phys. Rev. D 46 (1992) 4206.
[296] I. Yi, E.T. Vishniac, Phys. Rev. D 45 (1992) 3441.
[297] I. Yi, E.T. Vishniac, Astrophys. J. 86 (1993) 333.
[298] I. Yi, E.T. Vishniac, S. Mineshige, Phys. Rev. D 43 (1991) 362.
[299] M. Zaldarriaga, Phys. Rev. D 62 (2000) 063510.
[300] M. Zaldarriaga, Phys. Rev. D 69 (2004) 043508.
[301] M. Zaldarriaga, U. Seljak, Phys. Rev. D 59 (1999) 123507.
[302] Ya.B. Zel'dovich, R.A. Sunyaev, Astrophys. Space. Sci. 4 (1969) 301.
[303] P. Zhang, U.-L. Pen, Astrophys. J. 549 (2001) 18.