Non-gaussianities for primordial black hole formation

Marco Taoso\textsuperscript{a} and Alfredo Urbano\textsuperscript{b,c}

\textsuperscript{a}I.N.F.N. sezione di Torino, via P. Giuria 1, I-10125 Torino, Italy
\textsuperscript{b}Dipartimento di Fisica, “Sapienza” Università di Roma, Piazzale Aldo Moro 5, 00185, Roma, Italy
\textsuperscript{c}I.F.P.U., Institute for Fundamental Physics of the Universe, via Beirut 2, I-34014 Trieste, Italy

E-mail: marco.taoso@to.infn.it, alfredo.urbano@uniroma1.it

Received March 8, 2021
Revised June 29, 2021
Accepted July 11, 2021
Published August 10, 2021

Abstract. We analyze primordial non-gaussianities in presence of an ultra-slow phase during the inflationary dynamics, focusing on scenarios relevant for the production of primordial black holes. We compute the three-point correlation function of comoving curvature perturbations finding that non-gaussianities are sizable, and predominantly local. In the context of threshold statistics, we analyze their impact for the abundance of primordial black holes, and their interplay with the non-gaussianities arising from the non-linear relation between density and curvature perturbations. We find that non-gaussianities significantly modify the estimate of the primordial black holes abundance obtained with the gaussian approximation. However, we show that this effect can be compensated by a small change, of a factor $2 \div 3$ at most, of the amplitude of the primordial power spectrum of curvature perturbations. This is obtained with a small tuning of the parameters of the inflationary model.

Keywords: primordial black holes, dark matter theory, inflation

ArXiv ePrint: 2102.03610
1 Motivations

The theory of inflation provides an elegant mechanism that explains the origin of structures in the Universe. In the inflationary picture, space-time fluctuates quantum mechanically around a background that is expanding exponentially fast. After the end of inflation, these fluctuations are transferred to the radiation field, creating slightly overdense and under-dense regions. These tiny ripples are the initial conditions that start the process of gravitational collapse which eventually forms the intricate architecture of galaxies that populate our observable Universe today.

If the amplitude of some of the overdense regions exceeds a specific threshold value, the pull of gravity becomes unsustainable and the gravitational collapse directly forms primordial black holes (PBHs) [1]. Interestingly, for masses in the range $10^{18} \lesssim M_{\text{PBH}[g]} \lesssim 10^{21}$, a population of such PBHs may account for the totality of dark matter observed in the Universe today, as envisaged by Hawking and Carr in their pioneering research [2, 3].

Because of their intrinsic quantum-mechanical origin, the way in which quantum fluctuations lead to a classical pattern of perturbations can be described only in a probabilistic sense, and it is often assumed that primordial perturbations follow a gaussian statistics. It is important to stress that there is no fundamental reason to believe that this assumption is true, and it should be rather considered to be an approximation. This approximation seems to work exquisitely well at the typical scales which are relevant for cosmic microwave background (CMB) anisotropy measurements, $0.005 \lesssim k [\text{Mpc}^{-1}] \lesssim 0.2$, and this is because deviations from the gaussian pattern are expected to be suppressed by powers of the slow-roll parameters [4] which during conventional slow-roll dynamics take $O(\ll 1)$ values.

As far as the formation of PBHs is concerned, however, the story might be drastically different. The process of PBH formation involves completely different scales compared to those probed by CMB observations. Typically, in order to be compatible with observational bounds on PBH abundance in the present-day Universe, one needs to consider scales $10^{12} \lesssim$
Figure 1. Time-evolution of the physical Hubble length $1/H$ (red line), written in units of its present-day value $1/H_0 \approx 4.5 \times 10^3\text{ Mpc}$ (with $H_0 \approx 10^{-42}\text{ GeV}$ in natural units), as function of the logarithm of the scale factor $a$ of the Friedmann-Lemaître-Robertson-Walker line element. We follow the evolution of $1/H$ throughout the history of our observable Universe. The inflationary phase is computed according to the model in ref. [32]. After (instantaneous) reheating, the evolution of $1/H$ during the radiation epoch, matter epoch (m.e.) and the present-day epoch ($\Lambda$) is computed according to the standard $\Lambda$CDM model. We superimpose the time-evolution of physical length scales $\lambda = a/k$, with $k$ the comoving wavenumber, that are relevant for CMB observations (the green band labelled “observational window” with $0.005 \lesssim k$ [Mpc$^{-1}] \lesssim 0.2$ and $k_* = 0.05$ [Mpc$^{-1}$]). We also show the time-evolution of physical length scales associated to perturbations that are responsible for PBH formation (the blue band labelled “PBH scales” with $10^{12} \lesssim k$ [Mpc$^{-1}] \lesssim 10^{14}$). Scales that cross the ultra slow-roll phase when they are deep in the super-Hubble regime are not sizably affected by the presence of negative friction [32]. For these scales, non-gaussianities are slow-roll suppressed. This is the case of scales relevant for CMB observations. On the contrary, scales that cross the Hubble radius close to the ultra slow-roll regime can be exponentially enhanced [32]. The interval of $e$-folds during which the ultra slow-roll phase takes place is indicated with a vertical strip. At these scales, non-gaussianities are enhanced. PBH formation takes place at time $t_f$ after the PBH scales re-enter the Hubble radius during the radiation epoch. Once formed, we consider the simple scenario in which the PBH energy density redshifts as matter.

$k$ [Mpc$^{-1}] \lesssim 10^{14}$. At such small scales, the physics involved can be completely different compared to what expected at CMB scales. Indeed, the formation of fluctuations large enough to trigger gravitational collapse into PBHs requires a departure from conventional slow-roll dynamics, and the slow-roll parameters can easily reach $O(1)$ values. Consequently, the parametric suppression that limits the presence of non-gaussianities at CMB scales is not guaranteed anymore.

For illustration, we summarize the time-evolution of physical scales probed by CMB observations and scales relevant for PBH formation in figure 1.

The introductory discussion above makes clear that the possible presence of non-gaussianities plays a fundamental role in the precise computation of the PBH abundance. The formation of PBHs requires perturbations above some threshold, located in the tail of their probability distribution. Small deviations from the gaussian shape of the tail, therefore, may produce exponentially different results.
In this work we investigate the role of non-gaussianities in the context of inflationary models featuring an ultra slow roll (USR) phase [5–9]. This part of the inflationary dynamics has been invoked for the production of PBHs (see refs. [10, 11] for the earliest proposal in this direction\(^1\)). This is because, in these scenarios, an USR phase allows to boost the primordial fluctuations at small scales, leading to a sizable population of PBHs. Crucially, conventional slow-roll dynamics is violated during such a period, potentially enabling the generation of large non-gaussianities. In our analysis, we focus on single-field inflation models in which the USR phase arises when the inflaton, rolling down its potential, crosses an approximate stationary inflection point before the end of inflation. Specifically, we consider both a toy model, which allows an analytical exposition, and a more realistic parametrization of the inflaton potential.

There exist a number of preceding studies devoted to the understanding of non-gaussianities in the context of PBH formation, and we will properly acknowledge them in the course of our work. Before diving into the details of our analysis, let us sketch briefly the structure of this paper.

In section 2, we will frame in a more formal way the computation of the abundance of PBHs and the connection with inflationary dynamics. In section 3, we will compute the so-called bispectrum of comoving curvature perturbations which controls the first non-trivial (i.e. connected) deviation from exact gaussianity. The bispectrum vanishes identically for a gaussian distribution, and a large non-zero result will introduce sizable deviations from the gaussian statistics. In section 4, we will move to consider the connected part of the trispectrum of comoving curvature perturbations which controls deviation from exact gaussianity at the level of fourth-order statistical correlators. We will not compute the full connected trispectrum but we will argue that the non-zero contribution which is generated by cubic interactions plays a crucial role in the computation of the PBH abundance. In section 5 we discuss the computation of PBH abundance in the presence of non-gaussianities. We will show that their impact is not dramatic. We find that the size of non-gaussianities is large enough to significantly affect the PBH abundance. On the other hand, it is enough to change the amplitude of the power spectrum of comoving curvature perturbations of a factor \(2 \div 3\) to obtain the same abundance obtained with the gaussian approximation.

We present our conclusions in section 6. Appendix A complements the computations illustrated in the main text with further technical details.

2 On the abundance of primordial black holes: from theory to observations and back

In cosmology and in the theory of structure formation, we assume that fluctuations in the energy density or in the gravitational potential form a three-dimensional random field. A three-dimensional random field \(\delta(\vec{x})\) is a set of random variables, one for each point in the three-dimensional real space, defined by a probability functional \(P[\delta(\vec{x})]\) which specifies the probability for the occurrence of a particular realization of the field over the ensemble. In cosmology, the random field \(\delta(\vec{x})\) is the three-dimensional mass overdensity field which gives the fractional overdensity in a given region of space (and at a given cosmic time) with respect to the unperturbed background (a.k.a. density contrast). A random field is fully specified by the entire hierarchy of its correlation functions, \(\langle \delta(\vec{x}_1)\delta(\vec{x}_2)\ldots\delta(\vec{x}_n) \rangle\). The simplest type of

\(^1\text{Ref. [10] investigates a potential with a plateau and two breaks. The inflationary dynamics in presence of such features was studied in [12].} \)
random field is the gaussian random field for which the probability functional $P$ is a gaussian distribution. The statistical properties of a gaussian (subscript $\mathcal{G}$ hereafter) random field are completely specified by the two-point correlation function or, working in Fourier space, by the power spectrum. The $n$-point correlation functions either vanish (for odd $n$) or can be expressed in terms of the power spectrum as a consequence of the Isserlis’ theorem (for even $n$). In Fourier space, we have (assuming a vanishing mean value as well as homogeneity and isotropy of space)

$$
\langle \delta_k \rangle = 0, \\
\langle \delta_k \delta_{k'} \rangle = (2\pi)^3 \delta^{(3)}(k + k') \Delta_\delta(k), \\
\langle \delta_{k_1} \delta_{k_2} \delta_{k_3} \rangle = 0, \\
\langle \delta_{k_1} \delta_{k_2} \delta_{k_3} \delta_{k_4} \rangle = \langle \delta_{k_1} \delta_{k_2} \rangle \langle \delta_{k_3} \delta_{k_4} \rangle + \langle \delta_{k_1} \delta_{k_3} \rangle \langle \delta_{k_2} \delta_{k_4} \rangle + \langle \delta_{k_1} \delta_{k_4} \rangle \langle \delta_{k_2} \delta_{k_3} \rangle \\
= (2\pi)^6 \delta^{(3)}(k_1 + k_2) \delta^{(3)}(k_3 + k_4) \Delta_\delta(k_1) \Delta_\delta(k_3) + \text{two permutations},
$$

and so on. The variance of the gaussian distribution is

$$\sigma^2 = \lim_{\delta \to \delta_0} \langle \langle \delta(\vec{x}) \delta(\vec{y}) \rangle \rangle = \langle \langle \delta(\vec{x}) \delta(\vec{x}) \rangle \rangle = \int \frac{dk}{k} P_\delta(k), \quad \Delta_\delta(k) = \frac{2\pi^2}{k^3} P_\delta(k),$$

from which we see that the power spectrum of density fluctuations $P_\delta(k)$ is the contribution to the variance of the field per unit logarithmic interval of $k$, and saying that $P_\delta(k) \sim 1$ means that the Fourier modes in a unit logarithmic bin around wavenumber $k$ generate fluctuations $\delta \rho/\rho$ of order unity. Notice that $\sigma^2$ does not depend on the position, by virtue of the delta function eq. (2.2) which is indeed a consequence of homogeneity. We are interested in the situation in which there is a sizable probability to form in a given region of typical size $R$ density fluctuations $\delta \rho/\rho$ large enough to collapse and form black holes under the pull of their own gravitational weight. In order to study PBH formation and make contact with the power spectrum of comoving curvature perturbations, eq. (2.5) has to be modified in two ways. First, we convolve the random field $\delta(\vec{x})$ with some window function $W(|\vec{x} - \vec{y}|, R)$ to obtain $\delta_R(\vec{x}) = \int d^3 y W(|\vec{x} - \vec{y}|, R) \delta(\vec{y})$ so that, at every point $\vec{x}$, the smoothed field represents the weighted average of $\delta$ over a spherical region of characteristic dimension $R$ centred in $\vec{x}$. Second, we relate the density contrast to the comoving curvature perturbation $\mathcal{R}$ by means of the linear-order relation (in Fourier space) [13]

$$\delta_k = \frac{2(1 + \omega)}{(5 + 3\omega)^2} \left( \frac{k}{aH} \right)^2 \mathcal{R}_k,$$

where $\omega = p/\rho$ defines the equation of state, and we shall take $\omega = 1/3$ for radiation domination. It is important to remark that eq. (2.6) is not an exact relation but it is only valid at the linear order, and non-linear corrections to eq. (2.6) unavoidably source non-gaussianities, see [14–22] for related studies. We will come back to this point in section 5. For the moment, we shall work assuming the linear relation in eq. (2.6).

Eq. (2.5) becomes

$$\langle \delta^2 \rangle_R = 4 \pi \omega^2 \int \frac{dk}{k} W(kR)^2 \left( \frac{k}{aH} \right)^4 P_\mathcal{R}(k),$$

where $P_\mathcal{R}(k)$ is the power spectrum of the comoving curvature perturbation. $W(kR)$ is the Fourier transform of the window function introduced before, and we shall use the volume-normalized gaussian window function $W(kR) = \exp(-k^2 R^2/2)$.

\footnote{See [23] for a study of the role of the window function for the PBHs formation.}
are formed when the density perturbation exceeds the threshold value \( \delta_{th} \), the fraction of the Universe ending up in PBHs is given by the tail of the density contrast distribution \( \beta(M_{PBH}) = \gamma \int_{\delta_{th}}^{\infty} P(\delta) d\delta \), and in the gaussian approximation we have

\[
\beta_G(M_{PBH}) = \gamma \int_{\delta_{th}}^{\infty} \frac{d\delta}{\sqrt{2\pi}\sigma_R(M_{PBH})} \exp \left[ -\frac{\delta^2}{2\sigma_R(M_{PBH})^2} \right],
\]

where we convert the smoothed variance \( \sigma_R \) to a function of the PBH mass taking into account that the size \( R \) is related to the mass \( M_{PBH} \) as \( R \approx 2G_N M_{PBH}/\gamma a_{form} \). The parameter \( \gamma \) depends on the details of the gravitational collapse and \( a_{form} \) is the scale factor at the time at which the PBH is formed.

The distribution of PBHs as function of their mass is strongly peaked at the value that maximizes \( \sigma(M_{PBH}) \), in view of the exponential factor in eq. (2.8). In terms of the power spectrum \( P_R \), this means that the PBH abundance roughly scales as \( e^{-1/P_R} \). Finally, the present-day fractional abundance of dark matter in the form of PBH is given by

\[
f_{PBH} = \left( \frac{\gamma}{0.2} \right)^{1/2} \left[ \frac{\beta(M_{PBH})}{1.6 \times 10^{-16}} \right] \left[ \frac{g_*(T_{form})}{106.75} \right]^{-1/4} \left[ \frac{M_{PBH}}{10^{18} \text{ g}} \right]^{-1/2}.
\]

The computation of \( \beta \) by means of eq. (2.8) is known as threshold statistics. The value of \( \delta_{th} \) is subject to some uncertainty (varying in the range \( 0.4 \lesssim \delta_{th} \lesssim 2/3 \)), and also depends on the specific shape of the power spectrum. In our analysis, for simplicity, we adopt the fixed values \( \delta_{th} = 0.45 \) and \( \gamma = 0.2 \) [24].

An alternative way of computing the PBH abundance relies on the use of peak theory [25], in which one computes the number density of peaks of the overdensity field \( \delta \) (which in general contains many peaks, i.e. local maxima, and valleys, i.e. local minima) and integrate above some threshold \( \delta_c \) [26]. The latter in general differs from \( \delta_{th} \) and depends on the shape of the power spectrum [27] (see also ref. [14] in which the full non-linear relation between density and curvature perturbation is explored). Using again the gaussian approximation, one finds [15, 25, 26]

\[
\beta_G^{pk}(M_{PBH}) \approx \frac{\gamma}{3\pi} \frac{(k^2)^{3/2}}{3} r_m^3 \frac{\sigma^2_c}{\langle \delta^2 \rangle} e^{-\frac{\delta^2_c}{2\sigma^2}} \quad \text{with} \quad \langle k^2 \rangle = \frac{1}{\sigma^2} \int \frac{d^3k}{(2\pi)^3} k^2 \Delta_s(k),
\]

where \( \delta_c = \delta_{th}/3\psi(r_m) \) with \( \psi \) the average density profile and \( r_m \) is the comoving distance where the so-called compaction function is maximized (we refer to refs. [27, 28] for a comprehensive discussion and more detailed definitions).\(^3\)

\(^3\)Eq. (2.10) is obtained as follows. Introducing \( \nu \equiv \delta/\sigma \), the comoving number density of peaks is [26]

\[
n_{G,com}^{pk} = \left( \frac{2\pi}{3} \right)^{3/2} \nu^2 e^{-\nu^2}, \tag{2.11}
\]

and in physical space one has \( n_{G,phys}^{pk} = n_{G,com}^{pk}/a_f^4 \), with \( a_f \) is the scale factor at the formation time \( t_f \). The PBH mass reads \( M_{PBH}(\nu) = \mathcal{K} M_{PBH}(t_m) \sigma^5 (\nu - \nu_c)^\gamma \) where \( \gamma \) is a critical exponent that depends on the equation of state in the formation era (with \( \gamma \approx 0.36 \) for radiation [29]) and \( \mathcal{K} \sim O(1) \). \( M_{PBH}(t_m) \) is the horizon mass at horizon-crossing time \( t_m \) defined implicitly by the relation \( a(t_m) H(t_m) r_m = 1 \). One can write

\[
\beta_G^{pk} = \int_{\nu_c}^{\infty} d\nu \frac{M_{PBH}(\nu) n_G^{pk,phys}(\nu) / \rho_f}{\rho_f} \quad \text{where} \quad \rho_f \text{ is the background density at the formation time. Using the result from numerical simulation} \ a_f / a_m \sim 3 [28, 30] \text{ and in the limit} \ \nu_c \equiv \delta_c / \sigma \gg 1 \text{ one obtains eq. (2.10)} \quad \text{with} \ \gamma = 3\mathcal{K}(1 + \gamma) \nu_c^{-5} \sigma^7.\]
Notice also that, following refs. [15, 27, 28], we do not use the smoothing in eq. (2.10) when dealing with peak theory. The reason is that when applying peak theory, and in particular when computing the threshold for collapse, the scale $r_m$ in position space associated to the peak of the power spectrum arises naturally.

On general ground, the abundance obtained by means of peak theory is larger than the one provided by threshold statistics [25, 27]. In this paper, however, our interest is not a comparison between the two approaches. Instead, the issue that we want to explore is whether the gaussian approximation in eqs. (2.8), (2.10) is justified. It is clear that the formation of PBH is a tail effect and, as such, strongly depends on possible deviations from the gaussian approximation. Investigating the possible presence of non-gaussianities in the context of realistic inflationary models in which one generates a sizable abundance of PBH is precisely the main goal of this work. After clarifying whether non-gaussianities are present and what are their main properties, we will quantify their impact on the PBH abundance focusing on threshold statistics. We will consider the case of peak theory and present a comparison with threshold statistics in a forthcoming work [31].

Let us, therefore, illustrate in more details what are the conditions that the inflationary dynamics should possess in order to generate the large density fluctuations needed for PBH formation.

In the gaussian approximation, a sizable PBH abundance is obtained for $P_R \simeq 10^{-1} - 10^{-2}$ at the PBH scales, requiring therefore a huge enhancement of $P_R$ with respect to its value at the CMB scales, $P_R \simeq 10^{-9}$. This can happen in presence of an USR phase during inflation. In fact, during such a transitory part of the inflationary dynamics, the amplitude of the comoving curvature perturbations can be exponentially enhanced. Let us introduce the Hubble parameters $\epsilon$ and $\eta$:

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \quad \eta \equiv -\frac{\ddot{H}}{2HH},$$

(2.12)

where $H$ is the Hubble function and dot is the derivative with respect to time. During standard slow roll inflation $|\epsilon| \ll 1$, $|\eta| \ll 1$ and, in the super-horizon limit, the evolution of $R_k$ is a linear combination of a constant mode and an exponentially decaying one. The latter becomes quickly negligible and the amplitude of $R_k$ freezes to a constant value until the end of inflation. Instead, if $3 + \epsilon - 2\eta \simeq 3 - 2\eta < 1$, a condition which we take as a definition of the USR phase, the decaying mode becomes a growing one. In practice $R_k$ evolves as a dumped harmonic oscillator, with the sign of the friction term controlled by the combination $3 + \epsilon - 2\eta$. During the USR phase this sign flips and the friction term becomes a driving force. This is responsible for the exponential growth of $R_k$ and the enhancement of $P_R$ at small scales. Notice however that for certain comoving wavelengths the amplitude of the growing mode (once it freezes to a constant value after the USR phase) is opposite to the one of the constant mode. This cancellation produces a dip in the power spectrum. The outcome of this intricate dynamics can produce a power spectrum of comoving curvature perturbations as the one in figure 3.

To make a quantitative analysis we shall consider an explicit inflationary model where the USR phase is realized. We focus on the single field model discussed in [32], described by the following action in the Jordan frame

$$S = \int d^4x \sqrt{-g} \left[ -\frac{1}{2} \left( M_P^2 + \xi \phi^2 \right) R + \frac{1}{2} g_{\mu\nu} \partial^\mu \phi \partial^\nu \phi - V(\phi) \right].$$

(2.13)
Figure 2. The blue dashed line shows the evolution of $\eta$ as a function of the number of e-folds in the model of ref. [32]. Specifically, within this solution (depicted with a cyan star in [32]) all the dark matter is in form of PBHs. The black line shows the behaviour of $\eta$ in our toy-model.

where $\phi$ is the inflaton field, $M_P$ is the reduced Planck mass and $\xi$ the non-minimal coupling to gravity. The potential is a polynomial

$$V(\phi) = a_2\phi^2 + a_3\phi^3 + a_4\phi^4 + \sum_{n=5}^{N} a_n\phi^n.$$  \hspace{1cm} (2.14)

and in full generality contains also higher order operators ($n > 4$). These are not mandatory for the production of PBH but, as noted in [32], they improve the compatibility of the model with the CMB observables. After a judicious tuning of the coefficients of the quadratic and cubic terms, the potential presents an approximate stationary inflection point. This key features induces an USR phase in the inflationary dynamics. The dashed blue line in figure 2 shows the evolution of $\eta$ as a function of the e-fold time for a specific choice of the parameters of the model. As it is evident, the condition $\eta > 3/2$, and therefore the USR phase, is realized for a while. This solution, depicted with a cyan star in [32], is in agreement with CMB observables and present 100% of DM in the form of PBHs (the abundance is computed with eq. (2.8)). For concreteness, in the rest of the paper, we focus on this benchmark point for our numerical analysis. We refer to it as numerical model in the following. However, let us stress that our results will not be limited to this specific solution. On the contrary, as we will see, our conclusion will be of general validity for all the single-field inflationary models in which PBH production is induced by the presence of an approximate stationary inflection point before the end of inflation.

As shown by the black line in figure 2, the evolution of $\eta$ can be approximated, at least at the qualitative level, by a piecewise function. This observation has prompted a simple analytical model to describe the inflationary dynamics [32, 33]. The inflationary phase is splited in the three periods, each of them with constant values of $\eta$ (and $\epsilon \ll 1$). In this way one can obtain an analytical solution for the evolution of $R_k$ and consequently its power spectrum. The advantage is that one can gain more insight into the features of the USR dynamics, as the enhancement or decrease of $R_k$ for certain comoving wavelengths (see [32] for an in-depth discussion). Moreover, the analytical model does not rely on a specific
inflationary model, rather, changing its parameters, it can describe different USR phases. The drawback is that it is not realistic enough for a detailed comparison with observables. In the rest of the paper we will compare the results based on the analytical model with the outcome of a numerical analysis performed in the context of the inflationary model described above.

Let us briefly summarize the main features of the analytical model. The three regions discussed above are as follows.

- **Region I.** It corresponds to a standard slow roll phase. We assume $\eta_I = 0$ and therefore we have a constant $\epsilon_I \ll 1$.

- **Region II.** It corresponds to the USR phase, which extends from the $e$-fold time $N_{\text{in}}$ to $N_{\text{end}}$. We define $\Delta N \equiv N_{\text{end}} - N_{\text{in}}$ and take $\eta_{\text{II}} \gg 3/2$. Considering $\epsilon \ll \eta$, its evolution is $\epsilon_{\text{II}}(N) = \epsilon_I e^{-2\eta_I(N-N_{\text{in}})}$.

- **Region III.** It corresponds to the last part of the inflationary dynamics. We assume a negative value for $\eta_{\text{III}}$. $\epsilon$ evolves according to $\epsilon_{\text{III}}(N) = \epsilon_I e^{-2\eta_I \Delta N e^{-2\eta_{\text{II}}(N-N_{\text{end}})}}$.

In each region the evolution of the $R_k$ can be obtained analytically by solving the Mukhanov-Sasaki equation. Switching to conformal time $\tau$ (we remind that $dN/c = a H$), for each region (i) we have

$$ R_k^{(i)}(\tau) = \frac{H\sqrt{\pi}}{2}\sqrt{2\epsilon^{(i)}(\tau)}(-\tau)^{3/2} \left[ \alpha_k^{(i)} H_\nu^{(1)}(-k\tau) e^{i(\nu+1/2)\pi/2} + \beta_k^{(i)} H_\nu^{(2)}(-k\tau) e^{-i(\nu+1/2)\pi/2} \right], $$

where $H$ is taken constant since we always have $\epsilon \ll 1$. $H_\nu^{(1)}$ and $H_\nu^{(2)}$ are the Hankel function of first and second kind respectively and $\nu = \sqrt{9/4 - \eta_I (3 - \eta_I)}$. Imposing Bunch-Davies initial conditions, one has $\alpha_k^{(1)} = 1$ and $\beta_k^{(1)} = 0$ in the first region. The complex coefficients $\alpha_k^{(i)}$ and $\beta_k^{(i)}$ in the regions II and III are obtained requiring continuity of $R_k$ and its first derivative across the three regions. Finally, the power spectrum $P_R(k)$ is computed at super-Hubble scales as $P_R(k) \equiv \lim_{\tau \rightarrow -0} \frac{k^2}{2\pi^2} |R_k(\tau)|^2$.

We can now try to relate the parameters of the model to observable quantities (see table 1). For this purpose in the left panel of figure 3 we show the power spectrum of comoving curvature perturbations as a function of the comoving scale $k$ for a choice of the parameters of the analytical model. The large enhancement at small scales allows for a sizable PBH production. As anticipated, the analytical model reproduces, at least qualitatively, the main features obtained with a more accurate numerical calculation, namely the presence of a peak and a dip in the power spectrum. The correspondence between what we compute and what we observe can be explained with the following points.

- **i)** We fix the position of the peak of $P_R(k)$ at $k_{\text{peak}} = 10^{14} \text{Mpc}^{-1}$. This choice, in turn, gives a PBH distribution peaked around $M_{\text{PBH}} = 10^{18} \text{g}$. These values are motivated by the fact that such population of PBHs can successfully account for the totality of dark matter in the present-day Universe. Furthermore, from the value $M_{\text{PBH}} = 10^{18} \text{g}$ we know that the comoving wavenumber $k_{\text{peak}}$ crosses the horizon approximately 36 $e$-folds after the perturbation associated to the pivot wavenumber associated with the CMB scale $k_*=0.05 \text{Mpc}^{-1}$ does it. From that, we infer the value of $N_{\text{peak}}$ (and $N_{\text{in}}$).
ii) For a given value of $\Delta N$ and $\eta_{II}$, we can compute the position of the dip and, consequently, the number of e-folds between horizon crossing of the pivot scale $k_*$ and $k_{\text{dip}}$. We can, therefore, compute the normalization of the power spectrum $H^2/8\pi^2\epsilon_1$ by rescaling the observed value at the pivot scale, namely $A_s \approx 2 \times 10^{-9}$, up to $k_{\text{dip}}$ by means of the slow-roll evolution $P_R(k) = A_s(k/aH)^{n_s-1}$ where we take $n_s = 0.965$.

iii) Finally, the values of $\Delta N$, $\eta_{II}$ and $\eta_{III}$ are readjusted — recomputing for each choice the normalization described at point ii) — by requiring $f_{\text{PBH}} = 1$. The value of $\eta_{III}$ mostly controls the dynamics after the USR phase (although it also enters in the computation of the power spectrum), and its value can be gauged in order to guarantee that inflation lasts approximately 60 e-folds.

Typical values of $\Delta N$, $\eta_{II}$ and $\eta_{III}$ which generate the correct PBH abundance of dark matter are shown in figure 3. Despite the simplicity of the model, it is nice to see that these numbers are in full agreement with those obtained, by means of a careful numerical analysis, in the context of more complete models [32].

However, this discussion makes also clear that the picture risks being grossly incomplete without including the effect of non-gaussianities. This is mostly because the integral in eq. (2.8) is dominated by the tail of the distribution where deviation from the gaussian approximation may give a sizable effect.

On general ground, deviations from the gaussian approximation are expected in realistic models of inflation. At CMB scales, non-gaussianities are typically negligible since suppressed by the small values of the Hubble parameters $\epsilon$ and $\eta$ (see ref. [44] for the latest experimental constraints on primordial non-gaussianities at the scales probed by the Planck experiment). However, we are interested in the opposite situation in which slow-roll is violated, and in particular $\eta$ takes sizable values. It is, therefore, plausible to expect sizable non-gaussianities at scales much smaller that those probed by CMB measurements where the USR dynamics might take place.

Deviation from normality can be explored by considering skewness and kurtosis. Skewness ($S$) and kurtosis ($K$) measure deviations with respect to a gaussian normal distribution. Gaussian distributions are symmetric around their mean while, for a generic distribution, skewness measures the degree of departure from the symmetric shape. Kurtosis, on the contrary, is a measure of how differently shaped are the tails of a distribution with respect to the tails of the gaussian distribution. In the presence of non-gaussianities, eq. (2.8) is modified

| Model parameter | $H^2/8\pi^2\epsilon_1$ | $N_{in}$ | $\Delta N$ | $\eta_{II}$ | $\eta_{III}$ |
|-----------------|------------------------|---------|-----------|-------------|-------------|
| Description     | $P_R$ in region I       | beginning of USR | duration of USR | value of $\eta$ during USR | value of $\eta$ after USR |
| Features in the | normalization of $P_R$  | $k_{\text{peak}}$ | $k_{\text{dip}}, P_R(k_{\text{peak}})$ | $k_{\text{dip}}, P_R(k_{\text{peak}})$ | $P_R(k)$ |
| power spectrum  | at large scales         |          |           |             |             |
| Observable      | $A_s$                   | $M_{\text{PBH}}$ | $n_s$     | $f_{\text{PBH}}$ | total number of e-folds |

Table 1. Correspondence between the parameters of the model introduced in section 2 and the cosmological observables relevant for PBH formation.
as \cite{45, 46}

$$\beta_{NG}(M_{PBH}) = \beta_{G}(M_{PBH}) + \frac{e^{-\delta_{th}/2\sigma_{R}^{2}}}{\sqrt{\pi}} \left[ \frac{S_{R}}{2^{5/2}3!} H_{2} \left( \frac{\delta_{th}}{2\sigma_{R}} \right) + \frac{K_{R}}{2^{7/2}4!} H_{3} \left( \frac{\delta_{th}}{2\sigma_{R}} \right) + \ldots \right],$$

(2.16)

which can be considered as an application of the Gram-Charlier series, that is the expansion of the probability density function (PDF) in terms of its cumulant (sometimes also called Edgeworth expansion, the latter being the same series but with a different truncation and ordering of terms). In eq. (2.16), \(S_{R}\) and \(K_{R}\) are the skewness and the kurtosis smoothed on the scale \(R\) (in analogy with eq. (2.7)) and their explicit definitions will be discussed in the next sections. In short, skewness is related to the three-point correlation function which vanishes in the gaussian limit (see, e.g., eq. (2.3)) while kurtosis is related to the connected part of the four-point correlation function (that is the part of the four-point correlator that remains after subtracting the gaussian contribution, see, e.g., eq. (2.4)). The dots in eq. (2.8) indicate contributions from higher order correlators \(n \geq 5\) and \(H_{n}\) are the Hermite polynomials.\textsuperscript{4} In the following, we shall proceed with the discussion of non-gaussianities assuming eq. (2.16).

In the next section we will compute the skewness and in section 4 we will focus on the kurtosis. With these results, in section 5 we will investigate the impact of non-gaussianities on the abundance of PBHs.

\textsuperscript{4}Notice that we use the so-called “physicists’ Hermite polynomials” which are defined by \(H_{n}(x) = (-1)^{n}e^{x^{2}}d^{n}e^{-x^{2}}/dx^{n}\) instead of the so-called “probabilists’ Hermite polynomials” \(He_{n}(x)\) often preferred in statistics theory. The relation among the two is \(He_{n}(x) = 2^{-n/2}H_{n}(x)\).
3 The bispectrum and the skewness

The three-point correlation function in momentum space is

\[ \langle R_{k_1} R_{k_2} R_{k_3} \rangle = (2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) B_R(k_1, k_2, k_3), \] (3.1)

where \( B_R(k_1, k_2, k_3) \) defines the bispectrum. In the gaussian approximation, the bispectrum vanishes. While the power spectrum is a function of \( B \) only, the bispectrum is a three-dimensional function defined on a tetrahedral region for which \( \vec{k}_1, \vec{k}_2 \) and \( \vec{k}_3 \) satisfy the triangle condition enforced–as illustrated in the schematic picture above–by the delta function in eq. (3.1).\(^5\) It is convenient to define the reduced bispectrum \( B_R(k_1, k_2, k_3) \) by means of

\[ B_R(k_1, k_2, k_3) = B_R(k_1, k_2, k_3) [\Delta_R(k_1) \Delta_R(k_2) + \Delta_R(k_1) \Delta_R(k_3) + \Delta_R(k_2) \Delta_R(k_3)] . \] (3.2)

This is a useful definition since the reduced bispectrum \( B_R \) is dimensionless (roughly speaking, the scaling \( B_R \sim 1/k^6 \) is absorbed by \( \Delta_R^2 \sim 1/k^6 \)). In position space, the skewness is

\[ S = \frac{1}{(\sigma^2)^{3/2}} \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_3}{(2\pi)^3} \langle R_{k_1} R_{k_2} R_{k_3} \rangle \]
\[ = \frac{1}{(\sigma^2)^{3/2}} \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_3}{(2\pi)^3} B_R(k_1, k_2, k_3) \bigg|_{k_1+k_2+k_3=0} . \] (3.3)

We take this opportunity to give a more general definition. We define the \( n \)-th cumulant

\[ C^{(n)} = \frac{\langle \mathcal{R}(\vec{x}) \ldots \mathcal{R}(\vec{x}) \rangle}{(\sigma^2)^{n/2}} = \frac{\langle \mathcal{R}^n \rangle}{\sigma^n}, \] (3.4)

as the connected part of the \( n \)-point correlator (evaluated at the same point) normalized by the \( n \)-th power of the standard deviation. The connected part of the \( n \)-point correlator is obtained by subtracting from the full expression of the \( n \)-point correlator the gaussian contribution (which is zero for the 3-point correlator but, as we shall see in the next section, does not vanish for \( n \geq 4 \), see e.g. eq. (2.4)). Given the definition in eq. (3.4), we have \( C^{(3)} = S \).

The integration in \( \mathbb{R}^3 \) is limited by the tetrahedral region illustrated below (with semi-perimeter \( K \equiv (k_1 + k_2 + k_3)/2 \))

\[ V_B \equiv \left\{ \begin{array}{l} k_1 = K(1 - \alpha + \beta)/2 \\ k_2 = K(1 + \alpha + \beta)/2 \\ k_3 = K(1 - \beta) \end{array} \right\} \]

with slices

\[ k_1 = k_2 = K, k_3 = 0 \quad \beta = 1 \]
\[ \alpha = -1 \quad k_1 = k_3 = K, k_2 = 0 \]
\[ \alpha = 1 \quad k_2 = k_3 = K, k_1 = 0 \]

(3.5)

\(^5\)The delta function corresponds to invariance under translations, and the fact that the bispectrum depends only on the lengths of the three sides of the triangle formed by the momenta corresponds to invariance under rotations. Homogeneity and isotropy, therefore, reduce the nine parameters of the triad \( \vec{k}_{1,2,3} \) down to three.
and ranges $K \in [0, \infty)$, $\alpha \in [-(1 - \beta), 1 - \beta]$ and $\beta \in [0, 1]$. Qualitatively, if we write (assuming $\mathcal{P}_R$ and $\mathcal{B}_R$ are scale-independent) $(\sigma^2)^{3/2} \sim \mathcal{P}_R^{3/2}$ and $\mathcal{B}_R \sim \mathcal{B}_R \mathcal{P}_R^{1/2}$, we see that one expects $S \sim \mathcal{B}_R \mathcal{P}_R^{1/2}$. As a rule of thumb, we expect an almost gaussian perturbation if $\mathcal{B}_R \ll \mathcal{P}_R^{−1/2}$. This means that $\mathcal{B}_R$ is an interesting quantity to compute to have at least a first idea about the impact of the three-point correlator on non-gaussianities. To perform the integral in eq. (3.3) some care is needed for the angular part since the triangular condition enforced by the delta function involves also angles. We use the exponential integral form of the integral in eq. (3.3) some care is needed for the angular part since the triangular condition.

By using eq. (3.7), the goal is to compute the skewness as function of the PBH mass. To this end, we need to compute the three-point function $\langle \mathcal{R}_{k_1} \mathcal{R}_{k_2} \mathcal{R}_{k_3} \rangle$ in Fourier space. Notice that, in order to keep the full shape in momentum space of $\mathcal{B}_R(k_1, k_2, k_3)$ into account.

We compute $\langle \mathcal{R}_{k_1} \mathcal{R}_{k_2} \mathcal{R}_{k_3} \rangle$ by means of the standard background+perturbations splitting and using the standard in-in formalism. We report directly the results and refer to appendix A.1 for details. One can isolate two dominant contributions in the calculation of three-point function. The first one (the sum of expressions in and end below) originates from the following operator in the cubic action of $\mathcal{R}$: $\mathcal{C}_3 \propto − \int d^3 x \; a^2 \; \epsilon \eta' R^2 \mathcal{R}'$. The second contribution (dubbed red) originates by a convenient redefinition of $\mathcal{R}$, see eq. (A.7).
We consider first the analytical model discussed in section 2. Our result for $B_R(k_1, k_2, k_3)$ is

$$B_R(k_1, k_2, k_3) = B_R^{(\text{in})}(k_1, k_2, k_3) + B_R^{(\text{end})}(k_1, k_2, k_3) + B_R^{(\text{red})}(k_1, k_2, k_3),$$

(3.9)

where

$$k^6_{\text{in}}B_R^{(\text{in})} = \frac{-\text{Im}}{16\xi^3} \frac{H^4}{T^4} \frac{\Gamma (\nu_{IIj})^3}{\pi^{3/2}} \frac{e^{3(\eta_{II} - \eta_{II})}}{2^{1/2 - 3\eta_{II}}} \left( x_{12}x_{23} \eta_{II} - \frac{3}{2} \eta_{II} \mathcal{R}^{3}_{II} \mathcal{F}_{\text{in}}(x_{1}, x_{2}, x_{3}) \right),$$

(3.10)

$$\mathcal{R}^3_{II} \equiv \prod_{i=1}^{3} \left[ \alpha_{k_i} \frac{\beta_{k_i}}{x_{23} \epsilon_{3} (\eta_{II} + \frac{1}{2}) - \beta_{k_i} \frac{\beta_{k_i}}{x_{23} \epsilon_{3} (\eta_{II} + \frac{1}{2})} \right],$$

(3.11)

$$k^6_{\text{in}}B_R^{(\text{end})} = \frac{-\text{Im}}{16\xi^3} \frac{H^4}{T^4} \frac{\Gamma (\nu_{IIj})^3}{\pi^{3/2}} \frac{e^{3(\eta_{II} - \eta_{II})}}{2^{1/2 - 3\eta_{II}}} \left( x_{12}x_{23} \eta_{II} - \frac{3}{2} \eta_{II} \mathcal{F}_{\text{end}}(x_{1}, x_{2}, x_{3}) \right),$$

(3.12)

$$k^6_{\text{in}}B_R^{(\text{red})} = \frac{-4\pi^4\eta_{II}}{(x_{12}x_{23})^3} \left[ x_{3}^3 P_R(k_1)P_R(k_2) + x_{1}^3 P_R(k_2)P_R(k_3) + x_{2}^3 P_R(k_1)P_R(k_3) \right].$$

(3.13)

In the expressions above we defined $x_i \equiv k_i / k_{\text{in}}$, with $k_{\text{in}}$ the comoving wavenumber associated to the start of the USR phase at conformal time $\tau_{\text{in}}$ by the relation $k_{\text{in}} \tau_{\text{in}} = -1$. The two functions $\mathcal{F}_{\text{in}}(x_{1}, x_{2}, x_{3})$ and $\mathcal{F}_{\text{end}}(x_{1}, x_{2}, x_{3})$ are defined by the general expression

$$\mathcal{F}(x_{1}, x_{2}, x_{3}) \equiv \frac{i 8 \xi^{3/2} k_{\text{in}}^{7/2}}{H^3} \left( \mathcal{R}_{k_1}^s \mathcal{R}_{k_2}^s \frac{d \mathcal{R}_{k_3}^s}{d \tau} + \mathcal{R}_{k_1}^s \mathcal{R}_{k_3}^s \frac{d \mathcal{R}_{k_2}^s}{d \tau} + \mathcal{R}_{k_2}^s \mathcal{R}_{k_3}^s \frac{d \mathcal{R}_{k_1}^s}{d \tau} \right) \bigg|_{\tau = \tau_s},$$

(3.14)

and we have

$$\mathcal{F}_{\text{in}}(x_{1}, x_{2}, x_{3}) = \frac{e^{-i(x_1 + x_2 + x_3)}}{(x_{12}x_{23})^{3/2}} \left[ x_{2}^2 (x_{1}x_{2} - ix_{1} - ix_{2} - 1) + x_{1}^2 (x_{2}x_{3} - ix_{2} - ix_{3} - 1) \right.$$ \left. + x_{2}^2 (x_{1}x_{3} - ix_{1} - ix_{3} - 1) \right],$$

(3.15)

while the corresponding expression for $\mathcal{F}_{\text{end}}(x_{1}, x_{2}, x_{3})$ is more complicated and must be evaluated numerically for non-integers $\eta_{II}$.

Armed with these expressions, we are ready to compute the reduced bispectrum in eq. (3.2) and tackle the integral in eq. (3.7).

In figure 4 we show the bispectrum as function of the PBH mass. The results obtained in the context of the analytical model are shown in the left panel of figure 4. Notice that the ratio $\mathcal{S}_R = \delta^3_R / \sigma^2_R$ is proportional to the factor $A_s^{1/2} = H / \sqrt{8\pi^2} \epsilon_1$. On the left-side of the y-axis, we fix this value by means of the procedure outlined in the introduction of section 2 (see table 1 and related discussions). On the right-side y-axis, on the contrary, we indicate the size of the ratio $\mathcal{S}_R / A_s^{1/2}$ such that this quantity is not affected by the details of the fit of the power spectrum at large scales (where our analytical model is not accurate since $\epsilon_1$ is taken to be constant).

We find that $B_R^{(\text{in})}$ is negligible while $B_R^{(\text{end})}$ gives the dominant contribution compared to the local term $B_R^{(\text{red})}$. In the analytical model, the two contributions $B_R^{(\text{in})}$ and $B_R^{(\text{end})}$ come, respectively, from the two delta function transition in the time evolution of $\eta$

$$\frac{d \eta}{d \tau} = \eta_{II} \delta (\tau - \tau_{\text{in}}) + (-\eta_{II} + \eta_{III}) \delta (\tau - \tau_{\text{end}}).$$

(3.16)
Figure 4. Smoothed skewness in eq. (3.7) as function of the PBH mass computed for the model in which \( f_{\text{PBH}} = 1 \) in the gaussian approximation (see figure 3). The yellow line represents the sum of the three contributions in eq. (3.9). Separately, we also show the contribution from \( B_{\mathcal{R}}^{(\text{end})} \) and \( B_{\mathcal{R}}^{(\text{red})} \) (that is the local contribution); \( B_{\mathcal{R}}^{(\text{in})} \) gives a negligible contribution not shown in the plot. The right-side of the y-axis shows the value of \( S_\mathcal{R} \) in units of \((A_s)^{1/2}\) with \( A_s = H^2/8\pi^2\epsilon I \). On the left-side y-axis, \( A_s \) is fixed by the normalization chosen in figure 3. Right panel. Smoothed skewness in eq. (3.7) as function of the PBH mass computed for the numerical model in ref. [32]. In the numerical model, only the contribution \( B_{\mathcal{R}}^{(\text{red})} \) survives (solid yellow line).

Modeling the transition at \( \tau = \tau_{\text{end}} \) as a sharp step function can be misleading. This was already noticed in ref. [47] (even though in the context of non-attractor inflation models) and further analyzed in ref. [48]. The outcome of the analysis carried out in ref. [47] is that if one takes — in place of the sharp step function — a smooth transition at \( \tau_{\text{end}} \), the contribution \( B_{\mathcal{R}}^{(\text{end})} \) gets drastically reduced.

The evolution of \( \eta \) was already shown in figure 2 where it is indeed clear that the step-function approximation at \( \tau_{\text{end}} \) is the crude approximation of a smoother transition. When the smooth transition is implemented, we indeed find that \( B_{\mathcal{R}}^{(\text{end})} \) is completely negligible in the computation of the bispectrum. In short, this is because in the presence of a smooth transition \( B_{\mathcal{R}}^{(\text{end})} \) receives — instead of the instantaneous contribution of the delta function — two integrated contributions (respectively, before and after the transition at \( \tau_{\text{end}} \)) with opposite signs that almost cancel between each other (at the \( \% \) level). Let us give here just the idea of the proof (which is different compared to the one given in ref. [47]). Instead of the step-function approximation, we introduce the hyperbolic tangent parametrization

\[
\eta(N) = \frac{1}{2} \left[ -\eta_{\text{in}} + \eta_{\text{II}} \tanh \left( \frac{N - N_{\text{in}}}{\delta N} \right) \right] + \frac{1}{2} \left[ \eta_{\text{II}} + (\eta_{\text{II}} - \eta_{\text{in}}) \tanh \left( \frac{N - N_{\text{end}}}{\delta N} \right) \right],
\]

where the parameter \( \delta N \) controls the width of the two transitions at \( N_{\text{in}} \) and \( N_{\text{end}} \). This parametrization can be considered as a proxy for the actual numerical result (see figure 2). On the other hand, the limit \( \delta N \to 0 \) reproduces the step-function approximation. Notice that we also introduced an hyperbolic tangent to model the transition at \( N = N_{\text{in}} \), although not strictly needed for the validity of our considerations at \( N = N_{\text{end}} \). The hyperbolic tangent

\[
\eta_{\text{II}} = 5.74, \quad \eta_{\text{III}} = -0.75, \quad \Delta N = 1.79
\]
parametrization in eq. (3.17) is shown in the left panel of figure 5 for different values of $\delta N$. Consider the transition at $N = N_{\text{end}}$. In the cubic action $-\int d\tau d^3 x \, a^2 \epsilon \eta \, R^2 R'$ we now have to integrate over time instead of just picking up a delta function term at $N = N_{\text{end}}$. One ends up, as anticipated before, with the sum of two integrals (respectively, $I_{\text{left}}$ and $I_{\text{right}}$) covering the two regions indicated in the left panel of figure 5. For small $\delta N$, when the transition is very sharp, the two terms $I_{\text{left}}$ and $I_{\text{right}}$ have the same sign, and their sum reconstructs the large bispectrum found in eq. (3.12). For larger $\delta N$, $I_{\text{right}}$ flips sign and the cancellation occurs with increasing precision. In the right panel of figure 5 we show the quantity $(I_{\text{left}} + I_{\text{right}})/I_{\text{left}}$ as function of the width of the transition $\delta N$. For $\delta N \gtrsim 0.4$, the sum $I_{\text{left}} + I_{\text{right}}$ quickly gives a cancellation with the $\%$ accuracy.

In conclusion, we find that the analytical toy model in which the evolution of $\eta$ in the presence of USR is modeled by means of a step-function approximation is not best-suited for the computation of non-gaussianities since it misses the important cancellation that is at work when the transition between adjacent regions is implemented in a more physical way. For this reason, we switch from now on to the numerical model introduced in section 2. By direct computation, we confirm that $B_{\text{R}}^{(\text{end})}$ is negligible in the numerical model, where we verified the same cancellation discussed above for the hyperbolic tangent parametrization. This is because the evolution of $\eta$ in the numerical model can be indeed approximated very well by eq. (3.17) with width parameter $\delta N \simeq 0.44$ (for which the cancellation is optimal, see figure 5). As a final comment before proceeding our analysis, notice that the cancellation of $B_{\text{R}}^{(\text{end})}$ is not an exact mathematical identity but, as discussed before, its precision depends on the specific value of $\delta N$. In the numerical model introduced in section 2, the latter happens to be large enough that one can safely consider $B_{\text{R}}^{(\text{end})} \simeq 0$ when compared to $B_{\text{R}}^{(\text{red})}$. However, if one takes a model in which, say, $\delta N \simeq 0.3$ then our result shows that the cancellation works only at the $\sim 50\%$ level; this means that $B_{\text{R}}^{(\text{end})}$ gets reduced by, say, a factor of 2 but remains very sizable. Since the momentum structure of $B_{\text{R}}^{(\text{end})}$ is not of local type, in such a situation

Figure 5. Left panel. Hyperbolic tangent parametrization of $\eta$ in eq. (3.17) for different $\delta N$ (from darker to lighter red, $\delta N = 0.1 \div 0.5$ in steps of 0.1). The black line is the step-function limit $\delta N \to 0$. The two integrals $I_{\text{left}}$ and $I_{\text{right}}$ refer to the intervals $[N_{\text{end}} - 1.5, N_{\text{end}}]$ and $[N_{\text{end}}, N_{\text{end}} + 1.5]$, respectively. Right panel. Sum of the two integrals $I_{\text{left}} + I_{\text{right}}$ (normalized to $I_{\text{left}}$, and expressed in percentage) as function of $\delta N$. When the ratio $(I_{\text{left}} + I_{\text{right}})/I_{\text{left}}$ is above 100%, the two integrals $I_{\text{left}}$ and $I_{\text{right}}$ have the same sign and they sum constructively. Below this value, a cancellation occurs with increasing precision.
it would become interesting to include it. The relevant question becomes the following: is a model with $\delta N \simeq 0.3$ realistic? To answer this question, we can reverse-engineering the inflaton equation of motion and get

$$V(N) = V(N_{\text{ref}}) \exp \left\{ -2 \int_{N_{\text{ref}}}^{N} dN' \left[ \frac{\epsilon(3 - \eta)}{3 - \epsilon} \right] \right\},$$

(3.18)

$$\phi(N) = \phi(N_{\text{ref}}) \pm \int_{N_{\text{ref}}}^{N} dN' \sqrt{2 \epsilon},$$

(3.19)

as far as inflaton potential and field profile are concerned. These two equations are exact. We can now use the parametrization of $\eta$ given in eq. (3.17) and derive the time-evolution of $\epsilon$. Using $\eta \simeq -(1/2)d \log \epsilon/dN$, we find the following (cumbersome but analytical) expression

$$\epsilon(N) = \epsilon_I e^{-\eta_{\text{mI}}(N-N_{\text{ref}})} \left[ \cosh \left( \frac{N - N_{\text{end}}}{\delta N} \right) \cosh \left( \frac{N - N_{\text{mI}}}{\delta N} \right) \right]^{-\delta N \eta_{\text{mII}}/2} \left[ \cosh \left( \frac{N_{\text{ref}} - N_{\text{end}}}{\delta N} \right) \cosh \left( \frac{N_{\text{ref}} - N_{\text{mI}}}{\delta N} \right) \right]^{\delta N (\eta_{\text{mI}} - \eta_{\text{mII}})/2} \left[ \cosh \left( \frac{N_{\text{ref}} - N_{\text{end}}}{\delta N} \right) \sech \left( \frac{N_{\text{ref}} - N_{\text{mI}}}{\delta N} \right) \right]^{\delta N (-2 \eta_{\text{mI}} + \eta_{\text{mII}})/2},$$

(3.20)

where $\epsilon_I \ll 1$ is the value of $\epsilon$ at some reference $N_{\text{ref}}$ in region I. We now integrate eqs. (3.18), (3.19) and get the inflaton potential as function of $\phi$ for different $\delta N$. We show our result in figure 6. We find that the width of the transition region of $\eta$ is related to the deformation of the stationary inflection point of the potential. Small values $\delta N \to 0$ deepen the local minimum and lead to a situation in which the inflaton risks being trapped into it (thus preventing a graceful exit). This is clear as far as the limit $\delta N \to 0$ is concerned. One is tempted to conclude that the same result holds true when comparing the cases with $\delta N = 0.44$ and $\delta N = 0.3$. However, any intuition must be taken with a grain of salt since we know that small changes in the way one deforms the inflection point from its stationary configuration have a huge impact on the amplitude of the power spectrum (see discussion in ref. [32]). For this reason, we postpone a quantitative estimate of the relation between the power spectrum and $\delta N$ to the end of this section, and we move on with our discussion of non-gaussianities in the context of the numerical model in its benchmark configuration where $B_{R}^{(\text{end})}$ can be neglected.

The smoothed skewness in the context of the numerical model is shown in the right panel of figure 4. The contribution $B_{R}^{(\text{red})}$ survives also in the numerical model, and it is shown in yellow in the right panel of figure 4. In the numerical model, the contribution $B_{R}^{(\text{red})}$ is evaluated at some later time after the end of the USR phase when the relevant modes cease evolving (that is region III in the analytical model). We indicate with $\eta_0$ (which is de facto equivalent to $\eta_{\text{mII}}$) the (negative) value of $\eta$ at this transition time. Since $\eta_0$ enters as an overall multiplicative factor in the computation of the local contribution of the bispectrum (the dominant one, as explained above), in the right panel of figure 4 we plot the smoothed skewness in units of $-\eta_0$, that is $S_R/(-\eta_0)$. As far as this contribution is concerned, the numerical model gives a result that is in good agreement with the analytical model (despite the poor description of the latter at large scales).
In conclusion, we find that the presence of USR leaves a sizable imprint on the bispectrum of comoving curvature perturbations (and, consequently, on the density contrast) at the typical scales which are relevant for PBH formation. Non-gaussianities are predominantly of local type, meaning that the bispectrum of comoving curvature perturbations in Fourier space has the form (see eq. (3.13))

$$B_{\text{loc}}^R(k_1, k_2, k_3) = (6f_{NL}/5)[\Delta R(k_1)\Delta R(k_2) + \Delta R(k_1)\Delta R(k_3) + \Delta R(k_2)\Delta R(k_3)].$$

(3.21)

We introduced the amplitude parameter $f_{\text{NL}}$ with its historical normalization that became conventional, which for the class of models that we are studying reads $f_{\text{NL}} = -5/6\eta_0$.

The presence of USR dynamics is crucial to generate potentially sizable non-gaussianities. Conceptually, the reason is the following. The modes that contribute to the formation of the peak in the power spectrum of comoving curvature perturbations are those that are (in modulus) enhanced (or suppressed, as far as the modes that contribute to the dip of the power spectrum are concerned) by the negative friction phase that takes place during USR. These modes cease evolving only after the end of USR, a handful of $e$-folds before the end of inflation. During this last part of the inflationary dynamics, the Hubble parameter $\eta$ takes $O(1)$ negative values which are needed to connect the end of the USR phase with the end of inflation. In fact the inflaton almost stops on the top of the local maximum that forms when an approximate stationary inflection point is present in the potential, and accelerates afterwards to reach the subsequent absolute minimum where inflation ends. Since we are forced by the presence of the USR dynamics to evaluate correlators for comoving scales $k \gtrsim 10^{12}\text{Mpc}^{-1}$ in a region where $|\eta| \sim O(1)$, the usual slow-roll suppression which affects non-gaussianities at CMB scales is not at work anymore, thus opening the possibility of having potentially large effects.

---

Figure 6. Solution of eqs. (3.18), (3.19) with $\eta$-evolution given in eq. (3.17) and $\epsilon$-evolution given in eq. (3.20) for different $\delta N$ with $\phi_{\text{ref}} \equiv \phi(N_{\text{ref}})$.

---

6Locality here is just the statement that if the curvature is evaluated at some point $\vec{x}$ then the non-gaussianity is localized at same position. To be clear, a non-local type of non-gaussianity takes the form $R(\vec{x}) = R(\vec{0}) + \int d^3\vec{y}d^3\vec{z}K(\vec{x}; \vec{y}, \vec{z})[R(\vec{y})R(\vec{z}) - \langle R(\vec{y})R(\vec{z}) \rangle]$ with the local case corresponding to the kernel $K(\vec{x}; \vec{y}, \vec{z}) = 3f_{\text{NL}}/5\delta^{(3)}(\vec{y} - \vec{x})\delta^{(3)}(\vec{z} - \vec{x})$. 
Let us now close this section with a final comment related to the cancellation of $B_{R}^{\text{end}}$. Consider the scalar potential in eq. (2.14) that we write in the Einstein frame in the form

$$V(\phi) = \frac{\lambda \phi^4}{4(1 + \xi \phi^2)^2} \left[ 3 + \xi^2 \phi_0^4 - 8(1 + c_3) \frac{\phi_0}{\phi} + 2(1 + c_2)(3 + \xi \phi_0^2) \frac{\phi_0^2}{\phi^2} \right],$$

where $\phi_0 = 1$ indicates the position of the inflection point and $c_{2,3}$ describe deformations from its stationary configurations (that corresponds to $c_{2,3} = 0$). Changing the value of the non-minimal coupling (gray region in the plot above) keeping $c_{2,3} \neq 0$ fixed does not alter the shape of the potential around the inflection point. This means that all these solutions will be described by a similar value of $\delta N$. However, we find that a % variation of $\xi$ changes the peak of the power spectrum by many orders of magnitude. The smaller $\xi$ the lower $P_R(\mathbf{k}_{\text{peak}})$. In practice, in the context of the numerical model, the peak of the power spectrum is controlled by the size of the non-minimal coupling $\xi$. A small tuning of $\xi$ leaves $\delta N$ almost unchanged but allows to select the desired value of $P_R(\mathbf{k}_{\text{peak}})$. We will see this explicitly in section 5.

On the other hand, it is clear that changing the values of $c_{2,3}$ directly alters, by construction, the shape of the approximate stationary inflection point (red curves in the plot above, see ref. [32]). For instance, we found good inflationary solutions (that is solutions of the inflaton equation of motion in which the inflaton successfully overcomes the inflection point and reaches the absolute minimum at the origin) with $\delta N \approx 0.38$. According to the right panel of figure 5, this is enough to get a cancellation that works only at the 10% level. In such a case, neglecting $B_{R}^{\text{end}}$ is not justifiable. Of course, more work is needed to validate these solutions against CMB data and bound from Hawking evaporation (in particular, all solutions found with $\delta N \approx 0.38$ tend to generate PBHs that are too light to be dark matter).

The point that we want to make here is that the condition $B_{R}^{\text{end}} \approx 0$ is not guaranteed but one should always confront the $\eta$-evolution of a particular solution with our result in the right panel of our figure 5 before drawing any conclusion about the size of non-gaussianities generated by the transition at $N = N_{\text{end}}$. The recipe to follow is simple. Compute the time-evolution of $\eta$ by solving the inflaton equation of motion and compare it with eq. (3.17) to extract the value of $\delta N$ that best suits the shape of the transition at $N = N_{\text{end}}$. From the right panel of figure 5, one then reads the accuracy of the cancellation.

In the rest of this work we will focus on cases with $B_{R}^{\text{end}} \approx 0$.

---

\textsuperscript{7}Notice that here $\phi$ is not canonically normalized but this fact does not alter the validity of the qualitative argument below. See ref. [32] for details. We also neglect here for simplicity the presence of higher-dimensional operators.
The trispectrum and the kurtosis

The four-point correlation function in momentum space has the structure
\[
\langle R_{k_1} R_{k_2} R_{k_3} R_{k_4} \rangle = \left[ (2\pi)^6 \delta^{(3)}(\vec{k}_1 + \vec{k}_2) \delta^{(3)}(\vec{k}_3 + \vec{k}_4) \Delta_R(k_1) \Delta_R(k_3) + \text{two permutations} \right] +
\]
\[
(2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 + \vec{k}_4) T_R(k_1, k_2, k_3, k_4, k_{12}, k_{14})
\]
\[
(4.1)
\]

The trispectrum \(T_R\) depends on the six scalars specifying the quadrilateral formed by the wave-vectors, as illustrated in the schematic picture above where the two dashed lines correspond to \(k_{12} = |\vec{k}_1 - \vec{k}_2|\) and \(k_{14} = |\vec{k}_1 - \vec{k}_4|\). Notice that in the computation of \(\langle R_{k_1} R_{k_2} R_{k_3} R_{k_4} \rangle\) we need to isolate the connected part since the disconnected part is already present at the gaussian level, see eq. (2.4). In analogy with the reduced bispectrum introduced in eq. (3.2), it is useful to define the reduced trispectrum \(T_R\) which is given by
\[
T_R(k_1, k_2, k_3, k_4, k_{12}, k_{14}) = T_R \left\{ \Delta_R(k_1) \Delta_R(k_2) \left[ \Delta_R(|\vec{k}_1 + \vec{k}_4|) + \Delta_R(|\vec{k}_1 + \vec{k}_3|) \right] + 11 \text{ permutations} \right\}.
\]
\[
(4.2)
\]

As done in the previous section, we define the smoothed kurtosis as
\[
\mathcal{K}_R = \frac{1}{(\sigma_R^2)^2} \int \prod_{i=1}^4 \frac{d^3 \vec{k}_i}{(2\pi)^3} W(k_i R) \frac{16(1 + \omega)^2}{(5 + 3\omega)^2} \left( \frac{k_i}{aH} \right)^2 \langle R_{k_1} R_{k_2} R_{k_3} R_{k_4} \rangle.
\]
\[
(4.3)
\]

The exact computation of \(\mathcal{K}_R\) is beyond the scope of this paper. However, an important remark is in order. When non-gaussianities are of local type, they generate — in addition to the bispectrum of the form defined by eq. (3.21) and obtained in eq. (3.13) as a consequence of our explicit computation — also a contribution to the reduced trispectrum in eq. (4.2). For the sake of simplicity, we ignore in the following the smoothing procedure and the relation between density contrast and curvature perturbations, and we just consider the curvature perturbation \(R\). It is a simple exercise to show that if we take \(R = R_\varphi + (3 f_{\text{NL}}/5) (R_\varphi^2 - \langle R_\varphi^2 \rangle)\) we find a trispectrum of the form given in eq. (4.2) with \(T_R = 18 f_{\text{NL}}^3 / 25\) [49, 50]. This can be understood in the context of the \(\text{in-in}\) formalism as shown in appendix A.2 (see derivation of eq. (A.37)).

We do not attempt here to compute \(\mathcal{K}_R\) carefully but we just limit the discussion to the simple (and well-known) observation that a non-zero local-type bispectrum of order \(O(f_{\text{NL}})\) also implies a non-zero contribution to the reduced trispectrum of order \(O(f_{\text{NL}}^3)\). Statistically speaking, this means that we expect a non-zero fourth-order cumulant related to the third-order one. This comment actually applies also to higher-order cumulants as well.

Of course, we stress again that in addition to this \(O(f_{\text{NL}}^3)\) contribution other terms are expected which are generated, for instance, by a pure quartic interaction. In the following,
we shall work under the assumption that all these additional terms are negligible, and that the dominant contribution to the four-point correlator is the one generated by local cubic interactions.

5 On the PBH abundance in the presence of non-gaussianities: the local case

We are now in the position, using eq. (2.16), to analyze the impact of non-gaussianities for the PBH abundance. To make our discussion more clear, let us start by considering the PDF for the fractional overdensity \( \delta \) whose integral above the threshold gives the mass fraction \( \beta \) (see eq. (2.8)). We indicate with \( P_{NG}(\delta) \) the generic non-gaussian PDF, and with \( P_G(\delta) \) the PDF in the gaussian approximation. As anticipated, eq. (2.16) follows from the Gram-Charlier series of the PDF

\[
P_{NG}(\delta) = P_G(\delta) \sum_{n=0}^{\infty} \frac{1}{n!} C_R^{(n)}(0,0,...,0) H_n \left( \frac{\nu}{\sqrt{2}} \right),
\]

(5.1)

with \( \nu \equiv \delta/\sigma \) and where we introduced the \( n \)-th complete exponential Bell polynomial

\[
B_n(x_1, ..., x_n) = \sum_{k=1}^{n} B_{n,k}(x_1, ..., x_{n-k+1}),
\]

(5.2)

with \( B_{n,k} \) the partial exponential Bell polynomials. \( C_R^{(n)} \) is the \( n \)-th cumulant defined in eq. (3.8). In eq. (5.1), we carried out explicitly the expansion up to the sixth order to show that, from this order on, the coefficient of the \( n \)-th term is not just given by \( C_R^{(n)} \). This is a well-known result in the theory of statistics [51, 52]. Without entering to deep into the mathematical detail, eq. (5.1) can be thought as a Taylor expansion of \( P_{NG}(\delta) \) around a reference PDF, taken to be the gaussian distribution \( P_G(\delta) \). For completeness, if we integrate eq. (5.1) to compute \( \beta_{NG} = \gamma \int_{\delta_{th}}^{\infty} d\delta P_{NG}(\delta) \) and use the property

\[
\frac{1}{\sqrt{2\pi}\sigma^2} \int_{\delta_{th}}^{\infty} d\delta e^{-\frac{\delta^2}{2\sigma^2}} H_n \left( \frac{\delta}{\sqrt{2}\sigma} \right) = \frac{1}{\sqrt{\pi}} e^{-\delta_{th}^2/2\sigma^2} H_{n-1} \left( \frac{\delta_{th}}{\sqrt{2}\sigma} \right),
\]

(5.4)

we find precisely eq. (2.16). From this simple discussion, we therefore learn that eq. (2.16) follows from a series expansion around the gaussian PDF.

In principle, all the cumulants are needed to compute the PBH abundance with eq. (2.16). However, to examine the relevance of non-gaussianities, we first start focusing only on the third order cumulant \( C_R^{(3)} \), which we computed in section 3 (see figure 4),

\[
\frac{d^n}{dx^n} e^{-x^2/2\sigma^2} = \frac{(-1)^n}{2^n n! \sigma^ne^{-x^2/2\sigma^2}} H_n \left( \frac{x}{\sqrt{2}\sigma} \right),
\]

(5.3)

so that the reader can immediately recognize in eq. (5.1) a derivative expansion.
and set the remaining ones to zero. A complete calculation is performed later in this section.

We show our results in the left panel of figure 7. The analysis is performed in the context of the numerical model, and we focus on the PBH mass corresponding to the peak of the mass distribution (i.e. the most abundant PBHs). The dotted line (labelled \( \gamma_C \)) shows the PBH abundance computed integrating eq. (5.1) above the threshold as a function of the power spectrum of curvature perturbations at the peak position. Let us stress again that in eq. (5.1) \( C_R^{(3)} \) does not enter only at third order in the expansion, but also affects higher ones. For definiteness, we include terms up to \( n = 15 \) in eq. (5.1). This means that \( C_R^{(3)} \) enters at \( n = 3, 6, 9, 12, 15 \). We do not attempt here to include more terms and test the convergence of eq. (5.1) since, as we shall discuss in a moment, we will eventually adopt a different approach for the computation of the abundance (see [31] for a critical discussion on the truncation of the Gram-Charlier serie). On the top x-axis we also show the relative change of the non-minimal coupling (with respect to the value \( \xi_{PBH} = 1 \) that gives the totality of dark matter in PBHs for the gaussian model) that is needed to modify the peak of the power spectrum according to the corresponding point on the bottom x-axis. The dashed black line corresponds instead to the gaussian approximation in eq. (2.8). Two main messages can be extracted from this plot. Non-gaussianities are relevant, and modify the abundance computed with gaussian statistics by several orders of magnitude. On the other hand, the abundance is very sensitive to the amplitude of the power spectrum, and it is enough to change the latter quantity by a small amount to obtain the same abundance predicted by the gaussian approximation.

As mentioned above, this calculation is incomplete, because it misses the contributions from higher orders in the expansion and from the cumulants \( C_R^{(n)} \) with \( n > 3 \) (which are non-zero because, as discussed in section 4 for the specific case of the fourth-order cumulant, they at least receive a non-zero “induced” contribution from the local three-point function). Baring cancellations among the different contributions, one would expect that, at qualitative level, the results presented above would still hold, once all the other terms are included. In other words, the purpose of the estimate performed above is to highlight the relevance of non-gaussianities for the PBHs abundance. On the other hand, for a proper calculation one could in principle use eq. (2.16), and compute all the cumulants needed to reach the convergence of the serie at the desired accuracy. However, exploiting the results of section 3, we follow a simpler strategy, which will also allow us to take into account the non-gaussianities arising from the non-linear relation between curvature and density perturbations. The idea is the following. In section 3 we have shown that the three-point correlator in the presence of USR generates non-gaussianities that are predominantly of local type, since they give a bispectrum of the form given in eq. (3.21). This means that we can write in position space

\[
\mathcal{R} = \mathcal{R}_G + \frac{3f_{NL}}{5} \left( \mathcal{R}_G^2 - \langle \mathcal{R}_G^2 \rangle \right),
\]

where \( \mathcal{R}_G \) is gaussian and the term \( \langle \mathcal{R}_G^2 \rangle \) is added to ensure that \( \langle \mathcal{R} \rangle = 0 \). The statistical properties of the non-gaussian variable \( \mathcal{R} \) are therefore completely specified in terms of those of the gaussian variable \( \mathcal{R}_G \) and the parameter \( f_{NL} \), which controls the amount of primordial non-gaussianities. As explained in section 2, in threshold statistics the abundance of PBHs is obtained computing the probability that the density perturbation exceeds the threshold value \( \delta_{th} \). Here, the strategy is to relate the density perturbation to the comoving curvature

\[\delta_{th} \to \delta_{th}^f \quad \text{in the limit of small \( f_{NL} \).} \]
perturbation $\mathcal{R}$ and its spatial derivatives $\partial_i \mathcal{R}$ and $\Delta \mathcal{R}$, where $\partial_i$ denotes the derivative with respect to the spatial direction $i$. Using eq. (5.5), these non-gaussian variables can be expressed in terms of the analogous ones for the gaussian variable $\mathcal{R}_G$, whose statistical properties are known. In practice, exploiting the relation in eq. (5.5), we bypass the need of computing all the cumulants in eq. (2.16).

To see explicitly how this procedure can be implemented, let us first define the following quantity:

$$\sigma_j^2 = \int \frac{dk}{k} k W(kR)^2 P_R(k) k^{2j}, \quad \gamma = \frac{\sigma_1^2}{\sigma_0 \sigma_2}. \quad (5.6)$$

For ease of notation, we also define the following gaussian variables:

$$\nu = \frac{\mathcal{R}_G}{\sigma_0}, \quad \eta_i = \partial_i \mathcal{R}_G, \quad x = -\frac{\Delta \mathcal{R}_G}{\sigma_2}, \quad (5.7)$$

The joint PDF of the gaussian variables $\nu$, $\vec{\eta}$ and $x$ is [26]:

$$P(\nu, \vec{\eta}, x) d^3 \nu d^3 \vec{\eta} dx = A e^{-Q d\nu d^3 \vec{\eta} dx}, \quad \text{with} \quad Q = \frac{\nu^2}{2} + \frac{(x-x_*)^2}{2(1-\gamma^2)} + \frac{3 \vec{\eta} \cdot \vec{\eta}}{2\sigma_1^2}, \quad (5.8)$$

and

$$A = \frac{6\sqrt{3}}{8\pi^{5/2} \sqrt{2(1-\gamma^2)} \sigma_1^2}, \quad x_* = \gamma \nu. \quad (5.9)$$

To proceed, we should relate the comoving curvature perturbation with the density perturbation. As anticipated, this map is non-linear, and therefore it hides some unavoidable non-gaussianities. The full relation can be written in the form [53]

$$\delta(x, t) = -\frac{4(1+\omega)}{5+3\omega} \left( \frac{1}{aH} \right)^2 e^{-5\mathcal{R}(\vec{x})/2} \Delta e^{\mathcal{R}(\vec{x})/2}$$

$$= -\frac{2(1+\omega)}{5+3\omega} \left( \frac{1}{aH} \right)^2 \left[ \Delta \mathcal{R}(\vec{x}) + \frac{1}{2} \partial_i \mathcal{R}(\vec{x}) \partial_i \mathcal{R}(\vec{x}) \right], \quad (5.10)$$

We have two sources of non-linearities (and thus non-gaussianities) that come from $i$ terms of order higher than two in the expansion of the exponential and $ii)$ the quadratic term in the square brackets. Notice that eq. (5.10) is written in comoving gauge (defined by requiring both comoving slicing and comoving threading) whose choice seems appropriate since often used to perform simulations that describe the formation of PBHs in the context of numerical relativity. Furthermore, eq. (5.10) is not exact but includes terms up to the second order in the so-called spatial gradient expansion [53]. Notice that in eq. (5.10) $\mathcal{R}$ denotes the non-gaussian comoving curvature perturbation. Using eq. (5.5) one can recast the expression above in terms of the gaussian variables $\nu$, $\vec{\eta}$, $x$ in eq. (5.7). Then, solving for the variable $x$ one finds:

$$x_\delta(\delta, \nu, \vec{\eta}) = \frac{9(aH)^2 e^{2\sigma_0 |\nu+(\nu^2-1)\sigma_0\nu|} \delta + (2 + 8\alpha + 8\nu\alpha \sigma_0 + 8\nu^2 \alpha^2 \sigma_0^2) \vec{\eta} \cdot \vec{\eta}}{4(1+2\nu \alpha \sigma_0) \sigma_2}. \quad (5.11)$$

Above we have defined $\alpha = 3f_{NL}/5$ and taken $w = 1/3$, corresponding to the radiation domination era.
Finally, the fraction of the Universe collapsing into PBHs is obtained integrating the joint PDF in eq. (5.8) rewritten in terms of $\nu$, $\vec{\eta}$ and $\delta$, and imposing that density perturbation is above the threshold $\delta_{th}$:

$$\beta_{NG} = \gamma \int_{\delta_{th}}^{\infty} d\delta \int d^3\eta \int_{-\infty}^{\infty} d\nu J \mathcal{P}(\nu, \vec{\eta}, x_{\delta}),$$

(5.12)

where $J$ is the Jacobian of the transformation in eq. (5.11):

$$J = \left| \frac{9(aH)^2 e^{2\sigma_0 [\nu + (\nu^2 - 1)\sigma_0 \alpha]} \sigma_2}{4(1 + 2\nu \alpha \sigma_0 \sigma_2)} \right|,$$

(5.13)

Eq. (5.12) generalizes the gaussian expression in eq. (2.8) including the effect of primordial non-gaussianities of local type, and of non-gaussianities induced by the non-linear relation in eq. (5.10). For the latter source of non-gaussianities, an analogous expression has been obtained in [15]. One can check that in the limit of $f_{NL} \to 0$ and using eq. (5.10) at linear order in $R$, eq. (5.12) reduces to eq. (2.8).

We are now in position to quantify the impact of the non-gaussianities on the PBH abundance. We show in the left panel of figure 7 our results for $\beta(M_{PBH})$ as a function of the power spectrum at peak position for both the gaussian and non-gaussian cases. As before, we consider the numerical model, for which $\eta_0 \simeq -0.75$ and therefore $f_{NL} = -5/6 \eta_0 \simeq 0.625$. To isolate the effect of primordial local non-gaussianities, we first perform the calculation employing the linear relation among density and the comoving curvature perturbations, eq. (2.6). The result is shown with a red line in figure 7. Local non-gaussianities (with positive $f_{NL}$) tend to enhance the abundance of PBHs, and their effect can be reabsorbed by decreasing the amplitude of the power spectrum by a factor $2 \div 3$. This can be achieved at the prize of a small re-tuning of the parameter that controls the height of the peak of the power spectrum. In the case discussed in this paper, such parameter is the non-minimal coupling to gravity, and the re-tuning that one needs to gauge away the effect of non-gaussianities is of $\sim O(10^{-5})$.

We notice that using the Gram-Charlier series with the only inclusion of the third order cumulant $C_R^{(3)}$ and few terms in the expansion, leads to an increase of the PBH abundance and therefore, at the qualitative level, it captures the main effect of primordial non-gaussianities (in the limit in which non-linearities are neglected). However, we reiterate that some care is needed in the truncation of the serie, and a sufficient number of terms should be included in order to obtain a good convergence. We do not discuss further the Gram-Charlier series, and we focus on the approach that led to eq. (5.12). A critical discussion about the validity of the approach based on the Gram-Charlier series can be found in ref. [31].

Then, we introduce in the calculation also the non-linearities among density and comoving curvature perturbations. Using eq. (5.12) we obtain the green line in figure 7. We find that their effect is opposite to the one of primordial local non-gaussianities, making the formation of PBHs harder, in agreement with what found in [15] (but differently from this reference we also include primordial non-gaussianities). The reason for that can be read from eq. (5.11): non-linearities tends to increase the threshold $x_{\delta}$ relevant for PBH formation. From figure 7 one can see that, once again, the effect of non-gaussianities on the abundance can be reabsorbed by rescaling (in this case decreasing) the height of the power spectrum by a modest amount.

One could wonder how these results change for different values of $\eta_0$. To address this question, we show in the right panel of figure 7, as a function of $-\eta_0$, the rescaling factor...
Figure 7. Left panel: PBH abundance as a function of the power spectrum of curvature perturbations at peak position. The dashed black line is for the gaussian approximation, eq. (2.8). The dotted black line refers to the abundance computed integrating the Gram-Charlier series in eq. (5.1), see text for details. The red line takes into account the effect of primordial non-gaussianities, and the green line further includes the non-linearities between density and curvature perturbations. Right panel: rescaling factor of the power spectrum needed, in presence of non-gaussianities, in order to get the same PBH abundance predicted by the gaussian approximation. This quantity is shown as a function of $-\eta_0$, the parameter controlling the amount of local non-gaussianities. The color code of the lines is the same as in the left panel. The calculations in both panels are for the numerical model used in section 3.

of the amplitude of power spectrum needed to get the same PBH abundance predicted by the gaussian approximation. For comparison, let us mention that the models of single field inflation with an USR phase in refs. [32, 54, 55], which are designed to obtained a sizable PBH abundance, predict $-\eta_0 \simeq 0.48 - 1.22$ (see also [56] for a survey of different models in the literature). Our results show that non-gaussianities do not play a dramatic role for the computation of the abundance in this class of models, meaning that their effect can be compensated by a small tuning of the model parameters.

Before concluding we shall mention a couple of additional ingredients that are relevant to quantify the impact of non-gaussianities. In our analysis we focused on how non-gaussianities, primordial or from the non-linear relation, change the statistics of the density field. However, they also modify the threshold $\delta_{\text{th}}$. This is because $\delta_{\text{th}}$ depends on the shape of the density perturbation, which is affected by the presence of non-gaussianities. These modifications has been studied in the literature following different approaches, focusing on the effect of non-linearities [14, 16, 17, 21, 22], or including also primordial non-gaussianities [19] (for the latter case the power spectrum was assumed to be a Dirac-delta distribution; our power spectrum is significantly different). A complete investigation of the role of non-gaussianities should include also these effects. We leave it for future work.

Finally, quantum fluctuations of the inflaton field can alter the classical trajectory, modifying therefore the inflationary dynamics. This effect, which can be accounted for in
the framework of stochastic inflation [57], introduces non-gaussianities in the PDF of the comoving curvature perturbations, which can substantially alter the prediction for the PBHs abundance. Investigation of these quantum effects for the PBHs phenomenology is underway in the literature, see e.g. [33, 58–66].

6 Summary and conclusions

Let us conclude with a brief summary of the main novelties and results of our work. We have focused on the study of non-gaussianities which are generated by the presence of USR dynamics during inflation. The latter occur when the Hubble parameter $\eta$ takes values $> 3/2$.

An USR phase is crucial for the production of a sizable number of PBHs, and previous studies often limited the computation of their abundance to the gaussian approximation.

We have computed the bispectrum of the comoving curvature perturbations $R$ in the context of the $in-in$ formalism. From the operators appearing in the cubic action of $R$ we have isolated two dominant terms. The contribution from one of those drastically depends on the transition between the USR phase and the conventional slow-roll phase. We have shown that unless this transition is very sharp, a cancellation occurs in the calculation, dramatically reducing the imprint of this operator on the bispectrum.

We have then offered a procedure, based on a simple analysis of the time evolution of $\eta$, to estimate how sharp the transition should be in order for this cancellation to apply. Our results extend the findings of ref. [47], which focused on a scenario of non-attractor inflation where the USR phase, taken with $\eta = 3$, is connected to a phase with negligible $\eta$. On the contrary, in the context of models for PBHs production, the USR phase (generically with $\eta \neq 3$) is followed by a phase where $\eta$ is negative and $O(1)$. This is a generic property in the presence of USR, required to connect the end of the latter with the end of inflation.

We have found that the remaining contributions to the bispectrum generate non-gaussianities predominantly of the local type, with

$$f_{NL} = 5/6 (-\eta_0),$$

where $\eta_0$ is the value of the parameter $\eta$ after the end of the USR phase. We worked out the details of our analysis both semi-analytically and numerically, in the context of the model introduced in eqs. (2.13), (2.14), in which the inflaton field is a real scalar singlet equipped with a generic polynomial potential and non-minimally coupled to gravity in the Jordan frame. For this class of models, we extend the results presented in [56]. Recently it has been claimed that at the end of an USR phase non-gaussianities are slow-roll suppressed and thus negligible [67]. However, as we have shown with an explicit example, this is not necessarily true. On the contrary, in the class of models that we have studied, and which are motivated by the PBHs production, $\eta_0$ is generically negative and $O(1)$.

Then, using the formalism of threshold statistics, we have investigated the impact of non-gaussianities found here on the PBH abundance. To evaluate their relevance, or in other words the validity of the gaussian approximation, we have considered eq. (2.16), which allows to compute the PBH abundance once the cumulants of non-gaussian distribution are known. Focusing only on the first non-gaussian correction, i.e. the third-order cumulant $\zeta^{(3)}_R$, we have found that primordial non-gaussianities sizeably modify the estimation of the PBH abundance based on gaussian statistics. Qualitatively, the same conclusion has been obtained in [46, 56], where an approximation of eq. (2.16) was adopted to quantify the impact of non-gaussianities. A quantitative comparison with these analysis and the approximations adopted
there will be presented in [31]. Then, we have taken a step forward and we have shown how to compute the PBH abundance in presence of local non-gaussianities, extending the gaussian threshold statistics in eq. (2.8). This leads to eq. (5.12). In our calculations we have kept into account the non-linear relation between the density and curvature perturbations, which introduces another source of non-gaussianities in the addition to the primordial ones. A similar formalism has been presented in [15], but focused only to the study of non-linearities. Using eq. (5.12), we have found that the impact of non-gaussianities can be reabsorbed by a change of the amplitude of the power spectrum of curvature perturbations of a factor $2 \div 3$ at most. Namely, after such shift one recovers the same PBH abundance obtained with the gaussian approximation. For the inflationary model that we have analyzed, this modification is achieved by a minor (below the $\%$ level) tuning of one parameter (the non-minimal coupling to gravity).

Finally, let us stress that these results do not necessarily imply that non-gaussianities are irrelevant for the phenomenology of PBHs. A notable example of the contrary are the primordial gravitational wave induced by scalar curvature perturbations. This signal is an important route to test PBHs with future gravitational-wave interferometers. The energy density of the gravitational waves scales quadratically with the amplitude of the power spectrum of curvature perturbations, and therefore there exists a correlation between the PBH abundance and the gravitational wave signal. One shall be interested in scenarios where PBHs account for all the dark matter, or such that their abundance is in agreement with observational constraints in the relevant mass range. The gravitational signal would then be a prediction, whose theoretical uncertainty is tied to the ones involved in the calculations of the PBHs abundance. The effects discussed here are therefore relevant in this context. The interplay between gravitational signal and PBH abundance will be studied in a future work.

**Acknowledgments**

The research of A.U. is supported in part by the MIUR under contract 2017 FMJFMW (“New Avenues in Strong Dynamics,” PRIN 2017) and by the INFN grant “SESAMO — SinergiE di SAPore e Materia Oscura.” M.T. acknowledges support from the INFN grant “LINDARK,” the research grant “The Dark Universe: A Synergic Multimessenger Approach No. 2017X7X85” funded by MIUR, and the project “Theoretical Astroparticle Physics (TAsP)” funded by the INFN.

**A Three-point correlators in the in-in formalism**

**A.1 Computation of the three-point correlator $\langle R_{k_1}R_{k_2}R_{k_3} \rangle$**

In this appendix we sketch the derivation of eqs. (3.10), (3.12), (3.13). Even though the formalism used is quite standard, we shall explain in detail our analysis in order to better highlight the difference between our result and the existing literature.

The obtain the power spectrum of the comoving curvature perturbation, computed in section 2, one has to solve the equation of motion obtained from the second-order action for the comoving curvature perturbation.

This is not enough for the computation of non-gaussianities (equivalently, for the computation of higher-order correlators) since the latter can be thought as an imprint left by interactions. More in detail, we can split the full Hamiltonian in three parts $\mathcal{H}_{\text{full}} = \mathcal{H}_{\text{cl}} + \mathcal{H} + \mathcal{H}_{\text{int}}$, where $\mathcal{H}_{\text{cl}}$ is the background Hamiltonian constructed with the classical background field, $\mathcal{H}$
where the subscript \( f \) denotes interaction picture operators, and the interaction Hamiltonian \( \hat{H}_{\text{int}} \) is also constructed with the interaction picture operators. The \( i\epsilon \) prescription (where \( \epsilon \), not to be confused with the Hubble parameter, is a small real positive number) is introduced in order to regularize the time integral in the early time limit. This prescription assures that the integrand vanishes at early times, and hence the integral remains finite. At this stage, it is important to keep in mind that the vacuum expectation value on the left-hand side is evaluated on the interaction-picture vacuum while the right-hand side is evaluated on the vacuum of the free theory. The interaction Hamiltonian \( \hat{H}_{\text{int}} \) evolves the vacuum of the free theory to the interaction vacuum at the time \( \tau \) at which we evaluate the three-point function.

Eq. (A.3) makes clear that non-gaussianities, encoded in three- and higher-point correlators, are a manifestation of interactions. Notice also that eq. (A.3) is valid at the first order in \( S \). We remark that our goal is to compute, using eq. (A.3), the three-point correlator at some late time \( \tau \) close to the end of inflation when we can safely take the super-horizon curvature perturbations as constants.

The cubic action, in terms of the cosmic time, is given by \( S_3 = \int d\tau d^3\vec{x} L_3 \), with [4]

\[
\mathcal{L}_3 = a^3 \epsilon^2 \mathcal{R} \dot{\mathcal{R}}^2 + a \epsilon^2 \mathcal{R} (\partial \mathcal{R})^2 - 2a \epsilon \dot{\mathcal{R}} (\partial \mathcal{R})(\partial \chi) + \frac{a^3 \epsilon^2}{2} \mathcal{R}^2 \dot{\mathcal{R}}
\]

\[
+ \frac{\epsilon}{2a} (\partial \mathcal{R})(\partial \chi) \partial^2 \chi + \frac{\epsilon}{4a} (\partial^2 \mathcal{R})(\partial \chi)^2 + 2 f(\mathcal{R}) \frac{\delta S_2}{\delta \mathcal{R}} , \quad (A.4)
\]

\[
f(\mathcal{R}) = \frac{\epsilon_2}{4} \mathcal{R}^2 + \frac{1}{4} \mathcal{R} \dot{\mathcal{R}} + \text{terms quadratic in } \mathcal{R} \text{ with spatial derivatives acting on } \mathcal{R} , \quad (A.5)
\]

where \( \epsilon_2 = \epsilon/H\epsilon = 2\epsilon - 2\eta \), \( \partial^2 \chi = a^2 \epsilon \dot{\mathcal{R}} \) and \( \delta S_2/\delta \mathcal{R} \) the variation of the quadratic action with respect to the perturbation. Notice that the cubic action \( S_3 \) is exact for arbitrary \( \epsilon \) and \( \eta \). As shown in ref. [4], the last term in eq. (A.4) can be absorbed by the field redefinition \( \mathcal{R} \to \mathcal{R}_n + f(\mathcal{R}_n) \) so that, after this redefinition, the interaction Hamiltonian in conformal

\[\text{For the sake of clarity, and to connect with section 3, let us be clear with our notation. In the following we use}\]

\[
\hat{R}_I(\tau, \vec{k}) = \int \frac{d^3\vec{k}'}{(2\pi)^3} \hat{R}_I(\tau, \vec{k}') e^{i\vec{k}' \cdot \vec{x}} , \quad \hat{R}_I(\tau, \vec{k}) = R_k(\tau) a_{\vec{k}}(\tau) + R^*_{k}(\tau) a_{\vec{k}}(\tau) . \quad (A.1)
\]

The subscript \( \tau \) is used here to distinguish between the left- and right-hand side of eq. (A.3). Furthermore, we have the contraction

\[
\langle 0 | \hat{R}_I(\tau_1, \vec{k}_1) \hat{R}_I(\tau_2, \vec{k}_2) | 0 \rangle = (2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2) R_{k_1}(\tau_1) R_{k_2}(\tau_2) . \quad (A.2)
\]
time $\tau$ is
\begin{align}
\mathcal{H}_{\text{int}}(\tau) &= - \int d^3\vec{x} \left[ a e^2 \mathcal{R}_n (\mathcal{R}'_n)^2 + a e^2 \mathcal{R}_n (\partial \mathcal{R}_n)^2 - 2 e \mathcal{R}'_n (\partial \mathcal{R}_n) (\partial \mathcal{R}_n) + \frac{a e^2}{2} \mathcal{R}_n^2 \mathcal{R}'_n \\
&\quad + \frac{e}{2a} (\partial \mathcal{R}_n) (\partial \mathcal{R}_n) (\partial^2 \mathcal{R}_n) + \frac{e}{4a} (\partial^2 \mathcal{R}_n) (\partial^2 \mathcal{R}_n)^2 \right],
\end{align}
\tag{A.6}
with $\partial^2 \mathcal{R}_n = a^2 e \mathcal{R}_n$ and $'$ denotes the derivative with respect to $\tau$. This is not the only effect that is introduced by the field redefinition since it also gives some extra terms in the three-point correlation function of $\mathcal{R}$. In Fourier space, we have
\begin{align}
\langle \mathcal{R}(\tau, \vec{k}_1) \mathcal{R}(\tau, \vec{k}_2) \mathcal{R}(\tau, \vec{k}_3) \rangle &= \langle \mathcal{R}_n(\tau, \vec{k}_1) \mathcal{R}_n(\tau, \vec{k}_2) \mathcal{R}_n(\tau, \vec{k}_3) \rangle \\
&\quad + \frac{e^2}{4} \left[ \langle \mathcal{R}^2_n(\tau, \vec{k}_1) \mathcal{R}_n(\tau, \vec{k}_2) \mathcal{R}_n(\tau, \vec{k}_3) \rangle + \text{two perms} \right] \\
&\quad + \text{higher order terms},
\end{align}
\tag{A.7}
where one should keep in mind that in eq. (A.7) $\mathcal{R}^2_n(\tau, \vec{k}_1)$ is not just the square of $\mathcal{R}_n(\tau, \vec{k}_1)$ but it originates from the Fourier transform of the square of $\mathcal{R}_n(\tau, \vec{x})$ (since the field redefinition is defined at the level of the action in position space) meaning that, by means of the convolution theorem, we actually have
\begin{align}
\mathcal{R}^2_n(\tau, \vec{k}_1) &= \int \frac{d^3\vec{q}}{(2\pi)^3} \mathcal{R}_n(\tau, \vec{q}) \mathcal{R}_n(\tau, \vec{k}_1 - \vec{q}),
\end{align}
\tag{A.8}
and similarly for the other two permutations. In eq. (A.7) we neglected terms with spatial derivatives in eq. (A.5) since we are interested in super-horizon modes and the contribution from the second term in eq. (A.5) since we are interested in the computation of the three-point function at late time. All in all, the strategy of the computation is the following.

1) We compute the three-point function $\langle \mathcal{R}_n(\tau, \vec{k}_1) \mathcal{R}_n(\tau, \vec{k}_2) \mathcal{R}_n(\tau, \vec{k}_3) \rangle$, that is the first term on the right-hand side of eq. (A.7), by means of eq. (A.3) using the interaction Hamiltonian given in eq. (A.6);

2) we shift back from the bispectrum of $\mathcal{R}_n$ to that of $\mathcal{R}$ by including the leading correction given by the second term on the right-hand side of eq. (A.7).

For notation simplicity, we drop from now on the subscript $n$. We start from point 1). In the interaction Hamiltonian most of the terms are $\epsilon$-suppressed. The only exception which is relevant for us is the term that contains $\epsilon^2$. In our model, we have
\begin{align}
\frac{d \eta}{d \tau} &= \eta_{\text{II}} \delta(\tau - \tau_m) + (-\eta_{\text{II}} + \eta_{\text{II}}) \delta(\tau - \tau_{\text{end}}),
\end{align}
\tag{A.9}
meaning that in the integral in eq. (A.3) we have two contributions corresponding to the transition at $\tau = \tau_m$ and $\tau = \tau_{\text{end}}$. Let us compute explicitly only the first contribution. The integral on the right-hand side of eq. (A.3) gives (after promoting classical fields to operators)
\begin{align}
2 \text{Im} \eta_{\text{II}} a(\tau_m)^2 e_1 \int d^3\vec{x} \langle \hat{\mathcal{R}}_1(\tau, \vec{k}_1) \hat{\mathcal{R}}_1(\tau, \vec{k}_2) \hat{\mathcal{R}}_1(\tau, \vec{k}_3) \hat{\mathcal{R}}_1(\tau, \vec{x}) \hat{\mathcal{R}}_1(\tau, \vec{x}) \rangle \\
&\quad = 2 \text{Im} \eta_{\text{II}} a(\tau_m)^2 e_1 \int d^3\vec{x} \int \frac{d^3\vec{q}_1}{(2\pi)^3} \frac{d^3\vec{q}_2}{(2\pi)^3} \frac{d^3\vec{q}}{2\pi} \langle \hat{\mathcal{R}}_1(\tau, \vec{k}_1) \hat{\mathcal{R}}_1(\tau, \vec{k}_2) \rangle \\
&\quad \times \hat{\mathcal{R}}_1(\tau, \vec{k}_3) \hat{\mathcal{R}}_1(\tau, \vec{q}) \hat{\mathcal{R}}_1(\tau, \vec{q}_1 - \vec{q}) \hat{\mathcal{R}}_1(\tau, \vec{q}_2) \rangle e^{i\vec{x}(\vec{q}_1 + \vec{q}_2)},
\end{align}
\tag{A.10}
where in the last step we switched to Fourier space and we used the convolution theorem. The expectation value on the vacuum \(|0\rangle\) of the free theory can be computed by means of the Wick theorem. We only have three relevant contractions\(^\text{12}\) each of them that counts twice since, for instance, the two contractions

\[
\int d^3 \vec{x} \int d^3 \vec{q}_1 \int d^3 \vec{q}_2 \int d^3 \vec{q} \frac{d^3 \vec{q}_1}{(2\pi)^3} \frac{d^3 \vec{q}_2}{(2\pi)^3} \frac{d^3 \vec{q}}{(2\pi)^3} \hat{\langle} \hat{R}_I(\tau, \vec{k}_1) \hat{R}_I(\tau, \vec{k}_2) \hat{R}_I(\tau, \vec{k}_3) \hat{R}_I(\tau, \vec{q}) \hat{R}_I(\tau, \vec{q}_1 - \vec{q}) \hat{R}_I(\tau, \vec{q}_2) \rangle e^{i \vec{q} \cdot (\vec{q}_1 + \vec{q}_2)}
\]

\[
= \int d^3 \vec{x} e^{-i \vec{x} \cdot (\vec{k}_1 + \vec{k}_2 + \vec{k}_3)} \hat{R}_{k_1}(\tau) \hat{R}_{k_2}(\tau) \hat{R}_{k_3}(\tau) \hat{R}_{\vec{q}}(\tau) \hat{R}_{\vec{q}_1}(\tau) \hat{R}_{\vec{q}_2}(\tau) \hat{R}_{\vec{q}_1}(\tau),
\]

(A.11)

and

\[
\int d^3 \vec{x} \int d^3 \vec{q}_1 \int d^3 \vec{q}_2 \int d^3 \vec{q} \frac{d^3 \vec{q}_1}{(2\pi)^3} \frac{d^3 \vec{q}_2}{(2\pi)^3} \frac{d^3 \vec{q}}{(2\pi)^3} \hat{\langle} \hat{R}_I(\tau, \vec{k}_1) \hat{R}_I(\tau, \vec{k}_2) \hat{R}_I(\tau, \vec{k}_3) \hat{R}_I(\tau, \vec{q}) \hat{R}_I(\tau, \vec{q}_1 - \vec{q}) \hat{R}_I(\tau, \vec{q}_2) \rangle e^{i \vec{q} \cdot (\vec{q}_1 + \vec{q}_2)},
\]

(A.13)

give the same result. If we factor out \((2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3)\), the contribution of eq. (A.10) to the bispectrum becomes

\[
B_R^{(in)}(k_1, k_2, k_3) = 4 \text{Im} \left\{ \eta_{11} a(\tau_{in})^2 \epsilon_1 \hat{R}_{k_1}(\tau) \hat{R}_{k_2}(\tau) \hat{R}_{k_3}(\tau) \right. \times \left. \left[ \hat{R}^*_{k_1}(\tau_{in}) \hat{R}^*_{k_2}(\tau_{in}) \hat{R}^*_{k_3}(\tau_{in}) + \hat{R}^*_{k_1}(\tau_{in}) \hat{R}^*_{k_2}(\tau_{in}) \hat{R}^*_{k_3}(\tau_{in}) \right. \right.
\]

\[
\left. \left. + \hat{R}^*_{k_2}(\tau_{in}) \hat{R}^*_{k_3}(\tau_{in}) \hat{R}^*_{k_1}(\tau_{in}) \right] \right\},
\]

(A.14)

taken at late time \(\tau \to 0^-\) (so that the term \(\hat{R}_{k_1}(\tau) \hat{R}_{k_2}(\tau) \hat{R}_{k_3}(\tau)\) has to be computed in region III). Using eq. (2.15) and the expansion of the Hankel functions for small argument, we find eq. (3.10) used in the text.

The computation of the contribution at \(\tau = \tau_{end}\) follows the same steps but now we pick up the second delta function in eq. (A.9). We find the result

\[
B_R^{(end)}(k_1, k_2, k_3) = 4 \text{Im} \left\{ (-\eta_{11} + \eta_{12}) a(\tau_{end})^2 \epsilon_1 (\tau_{end}) \hat{R}_{k_1}(\tau) \hat{R}_{k_2}(\tau) \hat{R}_{k_3}(\tau) \times \left. \left[ \hat{R}^*_{k_1}(\tau_{end}) \hat{R}^*_{k_2}(\tau_{end}) \hat{R}^*_{k_3}(\tau_{end}) + \hat{R}^*_{k_1}(\tau_{end}) \hat{R}^*_{k_2}(\tau_{end}) \hat{R}^*_{k_3}(\tau_{end}) \right. \right. \right.
\]

\[
\left. \left. + \hat{R}^*_{k_2}(\tau_{end}) \hat{R}^*_{k_3}(\tau_{end}) \hat{R}^*_{k_1}(\tau_{end}) \right] \right\}.
\]

(A.15)

\(^{12}\text{Notice that in principle there are 15 ways of contracting pairs of operators (divided in three classes: pairings of external fields with fields at the interaction vertex, pairings of external fields among themselves and vertex fields among themselves). However, we are only interested in the connected contractions, i.e. contractions where one external operator in the string \(\hat{R}_I(\tau, \vec{k}_1) \hat{R}_I(\tau, \vec{k}_2) \hat{R}_I(\tau, \vec{k}_3)\) is contracted with one operator of the interaction vertex \(H_{int}\); this reduces the number of contractions down to 6. Physically, the disconnected contractions which are not included are proportional to tadpole diagrams but, since we are assuming that the background solution is stable (or, equivalently, that \(\langle \hat{R} \rangle = 0\)) all such amplitudes vanish.\)
Finally, we move to consider the contribution $ii$) that comes from the field redefinition in eq. (A.7). If we neglect $\epsilon$, we have $\epsilon_2 = -2\eta_{III}$ since this term must be evaluated towards the end of inflation in region III. The correction in eq. (A.7) reads

$$-\frac{\eta_{III}}{2} \int \frac{d^3 \vec{q}}{(2\pi)^3} \left[ \langle \hat{R}_I(\tau, \vec{q}) \hat{R}_I(\tau, \vec{k}_1 - \vec{q}) \hat{R}_I(\tau, \vec{k}_2) \hat{R}_I(\tau, \vec{k}_3) \rangle + \text{two permutations} \right]$$  \hspace{1cm} (A.16)

Consider the first term in the square brackets. If we apply the Wick theorem, we have again the contraction

$$-\frac{\eta_{III}}{2} \int \frac{d^3 \vec{q}}{(2\pi)^3} \langle \hat{R}_I(\tau, \vec{q}) \hat{R}_I(\tau, \vec{k}_1 - \vec{q}) \hat{R}_I(\tau, \vec{k}_2) \hat{R}_I(\tau, \vec{k}_3) \rangle$$

$$= -\frac{\eta_{III}}{2} (2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_3 + \vec{k}_3) |\mathcal{R}_{k_2}(\tau)|^2 |\mathcal{R}_{k_3}(\tau)|^2$$  \hspace{1cm} (A.17)

counts twice since it gives the same result of

$$-\frac{\eta_{III}}{2} \int \frac{d^3 \vec{q}}{(2\pi)^3} \langle \hat{R}_I(\tau, \vec{q}) \hat{R}_I(\tau, \vec{k}_1 - \vec{q}) \hat{R}_I(\tau, \vec{k}_2) \hat{R}_I(\tau, \vec{k}_3) \rangle.$$  \hspace{1cm} (A.18)

Repeating the same argument for the other two permutations in eq. (A.7) and factoring out $(2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3)$ from the definition of the bispectrum, we find

$$B_R^{(\text{red})}(k_1, k_2, k_3) = -\eta_{III} \left( |\mathcal{R}_{k_2}(\tau)|^2 |\mathcal{R}_{k_3}(\tau)|^2 + |\mathcal{R}_{k_1}(\tau)|^2 |\mathcal{R}_{k_3}(\tau)|^2 + |\mathcal{R}_{k_1}(\tau)|^2 |\mathcal{R}_{k_2}(\tau)|^2 \right)$$

$$= -\frac{4\pi^4 \eta_{III}}{(k_1 k_2 k_3)^3} \left[ k_1^3 \mathcal{P}_R(k_2) \mathcal{P}_R(k_3) + k_2^3 \mathcal{P}_R(k_1) \mathcal{P}_R(k_3) + k_3^3 \mathcal{P}_R(k_1) \mathcal{P}_R(k_2) \right] ,$$  \hspace{1cm} (A.19)

which reproduces eq. (3.13) (and agrees, if one takes $\mathcal{P}_R(k) = H^2/8\pi^2 \epsilon$, with the standard result obtained in single-field slow-roll inflation, see e.g. ref. [73]).

Before moving to the comparison of our results with the literature, let us mention that ref. [74] claims that the field redefinition in eq. (A.7) is the first order term of a more general expression, which leads to non-gaussianities of local form. Using the $\delta N$ formalism and working with a quadratic expansion of the inflaton potential around its local maximum, ref. [74] presents a resummation of the higher order corrections. Since the potential studied here differs from the simple approximation mentioned above, these results can not be directly applied to our case. Still, it would be interesting to extend such calculation for our scenario. Moreover, as mentioned in section 4, it is important to investigate the role of higher order contributions leading to non-local non-gaussianities.

We are now in the position to compare our result with refs. [47, 75]. Ref. [75] computes the amount of non-gaussianities in the case of non-attractor inflation models. In the specific realization considered, a canonically normalized scalar field rolls under a potential

$$V(\phi) = \begin{cases} 
  V_0 & \text{for } \phi < \phi_c \\
  V_2(\phi) & \text{for } \phi > \phi_c 
\end{cases}$$  \hspace{1cm} (A.20)

which is constant below some critical field value $\phi_c$ and where $V_2(\phi)$ supports a second phase of slow-roll inflation as in conventional models. In this model, which assumes the Minkowski
vacuum deep inside the horizon, $\eta$ transits from $\eta = 3$ (dubbed non-attractor phase) to $\eta = 0$ during the second slow-roll phase. Since $\eta = 0$ during the second inflationary phase, the only contribution to the three-point function comes from the delta function at transition time for $\eta$. Let us indicate this transition time with $\tau_e$. Ref. [75] also assumes that, immediately after the end of the non-attractor phase, the super-horizon curvature perturbations cease evolving (meaning that the comoving time $\tau$ at which eq. (A.3) has to be evaluated is taken to be the transition time $\tau_e$). As a result, ref. [75] finds a local bispectrum (see eq. (5.5)) with amplitude which, in the squeezed limit, is given by $f_{\text{NL}} = 5/2$. Notice that, based on this result, ref. [46] evaluated the amount of non-gaussianities expected in the case of PBH formation in single-field inflationary models featuring a USR phase. However, the model of ref. [75], designed for the case of non-attractor inflation, cannot be applied to the case relevant for PBH formation.

Ref. [47] improves the analysis of ref. [75] by assuming that the non-attractor inflation model consists of at least three different phases: the non-attractor phase (with $\eta = 3$), the transition phase (which is modeled in two ways dubbed smooth and sharp), and the final slow-roll phase (where $\eta$ eventually relaxes to $\eta = 0$). We reproduce, for ease of reading, in figure 8 the qualitative behavior of $\eta$ as function of time (with $N_e$ the $e$-fold time at which the non-attractor phase ends) for the two cases discussed in ref. [47] (smooth and sharp transition). In the case of a smooth transition, ref. [47] finds that the non-gaussianity computed in ref. [75] by means of the instantaneous approximation is almost cancelled. In the case of a sharp transition, ref. [47] finds that the amplitude of local non-Gaussianity is mainly determined by the value of the inflaton velocity at the end of the non-attractor phase.

Without going into detail, we remark here that none of these models, designed for the case of non-attractor inflation, can be considered as an optimal proxy for the case of PBH production. This is evident if we compare the time evolution of $\eta$ shown in figure 2 with those in figure 8. In realistic single-field inflationary models of PBH production which are compatible with observations, the field starts from slow-roll conditions ($\eta \approx 0$), goes through a USR phase ($\eta > 3$) and approaches the end of inflation via a third phase during which $\eta < 0$. This dynamics leaves a specific imprint on the three-point correlator which departs, both qualitatively and quantitatively, from the previous works, as discussed in this appendix.
Figure 9. Inflaton dynamics for the Coleman-Weinberg model studied in ref. [11]. In the left panel we show the time evolution of the inflaton field and in the right panel the time evolution of the Hubble parameter $\eta$.

### A.2 Computation of the three-point correlator $\langle R_{k_1}R_{k_2}R_{k_3} \rangle$: further remarks

Let us add few additional comments to the computation performed in appendix A.1.

a) We use the cubic action in eq. (A.4) as opposed to the expression [4, 73]

$$
S_3 = \int dt d^3 x \left\{ -a \epsilon R (\partial R)^2 - a^3 \epsilon R^3 + 3 a^3 \epsilon R \dot{R}^2 \right. \\
+ \frac{1}{2a} \left( 3 \dot{R} - \frac{\dot{R}}{H} \right) \left[ (\partial_i \partial_j \psi)(\partial_i \partial_j \psi) - (\partial^2 \psi)(\partial^2 \psi) \right] - \frac{2}{a} (\partial_i \psi)(\partial_i R)(\partial^2 \psi) \right\},
$$

(A.21)

with $\psi = -R/H + \alpha^2 \epsilon \partial^2 R$ because all terms in eq. (A.4) are explicitly proportional to the Hubble parameters and, therefore, more easy to handle. Furthermore, in eq. (A.4) the cubic vertex is simpler since the last term can be eliminated via a field redefinition [4]. The action obtained by means of the Lagrangian density in eq. (A.4) can be derived from the one in eq. (A.21) after integrating by parts, discarding total derivatives, and using the field equation that follows from the quadratic action [4]. Notice that eq. (A.21) was used in ref. [11] to compute non-gaussianities related to PBH production in the context of the chaotic new inflation model. Ref. [11] finds that non-gaussianities give small corrections to the PBH abundance. However, the model presented in ref. [11] differs — both qualitatively and quantitatively — from the class of models studied in this paper. The inflaton follows a Coleman-Weinberg potential (see inset plot in the left panel of figure 9) starting from large field values and rolling down towards the origin. Initial conditions and model parameters are tuned in such a way that the inflaton crosses the positive potential minimum and the barrier at the origin, falls in the negative potential minimum, bounces back, crosses a second time the origin, bounces back again (this time on the positive side of the potential), climbs again the central barrier and it has just enough inertia to spend $N \gtrsim 10$ e-folds moving slowly in the neighborhood of the origin after the oscillation. For ease of reading, we reproduce the classical dynamics of the inflaton in the left panel of figure 9 where the oscillation and the subsequent second phase of inflation is evident. The latter ends when the inflaton falls again in the negative potential minimum, where it finally settles down. In
this model, PBHs are produced during the second phase of inflation when the inflaton moves slowly close to the origin. In the right panel of figure 9, we computed the time evolution of the $\eta$ for the same model parameters and initial conditions used in the left panel. The solid (dashed) line indicates positive (negative) values. The time profile of $\eta$ is considerably more complicated than the one expected in our class of models mostly because of the presence of the oscillating phase during which $\eta$ changes twice from large positive to large negative values before taking for $\Delta N \simeq 1$ e-folds the constant value $\eta \simeq 3$ close to the origin (where the potential is flat). This is a clear example of a model that differs from the class of models that we analyze in this paper. However, we do not proceed further in the explicit analysis of this setup since one gets $n_s \simeq 0.94$ and $r \simeq 0.32$ for, respectively, the spectral index and the tensor-to-scalar ratio. These values are now grossly excluded by the most recent CMB measurements.

b) Consider the field redefinition $\mathcal{R} \rightarrow \mathcal{R}_n + f(\mathcal{R}_n)$ and the quadratic action for $\mathcal{R}$

$$S_2[\mathcal{R}] = \int d\tau d^3 \vec{x} a^2 \epsilon \left[ (\mathcal{R}')^2 - (\partial \mathcal{R})^2 \right]. \quad (A.22)$$

Since $f(\mathcal{R}_n)$ is a quadratic function in $\mathcal{R}_n$, the shift in the quadratic action generates cubic and quartic terms for $\mathcal{R}_n$ but not additional quadratic ones; schematically, we have $S_2[\mathcal{R}] \rightarrow S_2[\mathcal{R}_n] + \hat{S}_3[\mathcal{R}_n] + O(\mathcal{R}_n^4)$, and the quadratic action for $\mathcal{R}_n$ has, therefore, exactly the same form as for $\mathcal{R}$. This is the reason why even if we formally work with $\mathcal{R}_n$ in eqs. (A.6), (A.7) we can still use the free-field solutions. For the very same reason, the shift does not change $\mathcal{L}_3$ in eq. (A.4) but it just relabels $\mathcal{R} \rightarrow \mathcal{R}_n$; however, the utility of the field redefinition introduced in [4] is that what cancels the last term in eq. (A.4) is the cubic correction $\hat{S}_3[\mathcal{R}_n]$ introduced before and originating from the quadratic action, as one can explicitly check.

c) Consider now some of the interaction terms in eq. (A.6) that we neglected. Let us focus on the interaction term

$$\mathcal{H}^{(c_1)}(\tau) \equiv - \int d^3 \vec{x} a^2 \epsilon (\mathcal{R}')^2, \quad (A.23)$$

which generates a contribution to the bispectrum of the form

$$B^{(c_1)}(\mathcal{R}) = - 4 \text{Im} \mathcal{R}_{k_1}(\tau)\mathcal{R}_{k_2}(\tau)\mathcal{R}_{k_3}(\tau) \times \int_{-\infty(1-i\epsilon)}^{\tau} d\bar{\tau} a(\bar{\tau})^2 \epsilon(\bar{\tau})^2 \left[ \mathcal{R}_{k_1}^*(\bar{\tau}) \mathcal{R}_{k_2}^*(\bar{\tau}) \mathcal{R}_{k_3}^*(\bar{\tau}) + \text{two perms} \right]. \quad (A.24)$$

At late time the integrand eventually goes to zero because the perturbations settle to constant values. However, the integral is computed along the whole inflationary history, and, in particular, it crosses the USR phase during which the perturbations grow exponentially. It is simple to see that the contribution to the bispectrum from eq. (A.24) scales according to

$$B^{(c_1)}(\mathcal{R}) = \epsilon \left( \frac{H^4}{\epsilon f_k_{\text{in}}} \right) \mathcal{F}_{c_1}(x_1, x_2, x_3), \quad (A.25)$$

where $\mathcal{F}_{c_1}$ is a shape function which depends on the integral in eq. (A.24). As expected, eq. (A.25) pays a suppression factor $\epsilon \ll 1$ if compared to the contributions we included.
in our computation. For this reason, we neglected this contribution. Of course, in more realistic models in which it is well possible to reach $\epsilon_f = O(1)$ during the preUSR dynamics, the same argument can be invalidated. To proceed further, one has to evaluate the time integral and compute the shape function explicitly. In our model, we can split the integral in three parts. In region I, the integral is analytical. The contribution at $\tau = -\infty(1 - i\epsilon)$ vanishes thanks to the $\epsilon$ prescription (physically, this means that interaction are switched off in the Bunch-Davis vacuum), and only the contribution at $\tau = \tau_{in}$ survives. In region II and region III, the integral can be computed numerically. In region III perturbations with $k \lesssim k_{peak}$ settle exponentially fast to their final constant value, and the integral gives a small contribution since the integrand quickly vanishes. On the contrary, as discussed in ref. [32], perturbations with $k \gtrsim k_{peak}$ evolves in region III until their horizon crossing time $N_k > N_{end}$ (after which they become constant) and they contribute to the integral in eq. (A.24). We checked numerically that the contribution in eq. (A.25) to the bispectrum is largely subdominant.

Consider now the interaction term

$$\mathcal{H}^{(2)}_{int}(\tau) = 2 \int d^3\vec{x} \epsilon \mathcal{R}'(\partial_i \mathcal{R})(\partial_i \chi), \quad \chi = \epsilon a^2 \partial^{-2} \hat{R}, \quad (A.26)$$

with $\partial^{-2}$ the inverse laplacian. The corresponding contribution to the bispectrum reads

$$B^{(c2)}_R = 4 \text{Im} \mathcal{R}_{k_1}(\tau)\mathcal{R}_{k_2}(\tau)\mathcal{R}_{k_3}(\tau)$$

$$\times \int_{-\infty}^{\tau} d\bar{\tau} a(\bar{\tau})^2 \epsilon(\bar{\tau})^2 \left[ \frac{\vec{k}_2 \cdot \vec{k}_3}{k_3^2} \mathcal{R}^{*}_{k_1}(\bar{\tau})\mathcal{R}^{*}_{k_2}(\bar{\tau})\mathcal{R}^{*}_{k_3}(\bar{\tau}) + \text{six perms} \right], \quad (A.27)$$

where the scalar products can be expressed in terms of the magnitudes $k_i$ using the conservation enforced by the triangle identity, for instance $\vec{k}_2 + \vec{k}_3 = -\vec{k}_1 \implies \vec{k}_2 \vec{k}_3 = (k_1^2 - k_2^2 - k_3^2)/2$. As before, eq. (A.27) features the scaling

$$B^{(c2)}_R = \epsilon_1 \left( \frac{H^4}{\epsilon_f k_{in}} \right) \mathcal{F}_{c2}(x_1, x_2, x_3), \quad (A.28)$$

with the extra suppression given by $\epsilon_1 \ll 1$. The explicit computation of $\mathcal{F}_{c2}(x_1, x_2, x_3)$ is not conceptually different from that of $\mathcal{F}_{c1}(x_1, x_2, x_3)$ and leads to comparable results.

The few examples considered here confirm that the bispectrum computed from the time variation of $\eta$ captures the leading contribution coming from the interaction Hamiltonian. We also showed that organizing cubic interactions according to eq. (A.6) is indeed useful since it allows to exploit a power counting in terms of the Hubble parameter $\epsilon$ (assuming the latter remains small during the inflationary dynamics).

d) A comment on eq. (A.7) is needed. As anticipated, the three-point function is meant to be evaluated in the vacuum state of the full interacting theory. More formally, one should write

$$\langle \hat{R}(\tau, \vec{k}_1)\hat{R}(\tau, \vec{k}_2)\hat{R}(\tau, \vec{k}_3) \rangle = \langle \Omega(\tau) | \hat{R}_I(\tau, \vec{k}_1)\hat{R}_I(\tau, \vec{k}_2)\hat{R}_I(\tau, \vec{k}_3) | \Omega(\tau) \rangle, \quad (A.29)$$

with the vacuum of the interacting theory given by

$$|\Omega(\tau)\rangle = T e^{-i \int_{-\infty}^{\tau-i\epsilon} d\bar{\tau} a(\bar{\tau}) \mathcal{H}_{int}(\bar{\tau})} |0\rangle$$

where $T$ is the time-ordering operator. At initial time, when interactions are turned off (thanks to the $i\epsilon$ prescription) the initial
vacuum reduces to the Bunch-Davis vacuum $|0\rangle$. Notice that the three-point function 
\[ \langle \hat{R}(\tau, \vec{k}_1)\hat{R}(\tau, \vec{k}_2)\hat{R}(\tau, \vec{k}_3) \rangle \] 
would vanish if evaluated in $|0\rangle$ (since it contains an odd numbers of operators) so that its non-zero value is a genuine imprint left by interactions.

The same comment applies to all terms on the right-hand side of eq. (A.7) which should be evaluated in the vacuum state of the full interacting theory. This is indeed the way in which we computed the first contribution on the right-hand side, 
\[ \langle \hat{R}_n(\tau, \vec{k}_1)\hat{R}_n(\tau, \vec{k}_2)\hat{R}_n(\tau, \vec{k}_3) \rangle. \]
Consider now the other term
\[ \frac{\epsilon_2}{4} \left[ \langle \hat{R}^2_n(\tau, \vec{k}_1)\hat{R}_n(\tau, \vec{k}_2)\hat{R}_n(\tau, \vec{k}_3) \rangle + \text{two permutations} \right]. \tag{A.30} \]
We computed this term in the vacuum of the free theory. The difference is that in this case the operator is quartic in $\hat{R}_n$, and we have a non-zero leading contribution already in the vacuum $|0\rangle$ (the first term in the Dyson series of $|\Omega(\tau)\rangle$). This is the contribution that has been computed in eq. (A.19). Further corrections to this contribution can be discussed (dubbed “higher order terms” in eq. (A.7)). One possibility is given by the term
\[ \frac{\epsilon_3^3}{64} \left[ \langle \hat{R}^2_n(\tau, \vec{k}_1)\hat{R}^2_n(\tau, \vec{k}_2)\hat{R}^2_n(\tau, \vec{k}_3) \rangle + \text{permutations} \right], \tag{A.31} \]
which is also non-zero evaluated on the free-field vacuum. This term generates a contribution to the bispectrum which has the form
\[ B_{\hat{R}}^{(d_1)} = -\frac{\eta_{\text{II}}^3}{8} \int \frac{d\Omega_q}{(4\pi)} \int \frac{dq}{q} \frac{(2\pi^2)^2}{(k_1 + q)^3(k_3 - q)^3} P_R(k_1 + q)P_R(q)P_R(k_3 - q) + \text{permutations}. \tag{A.32} \]

If we approximate $P_R(q)$ as a spiky power spectrum so that the integration over $q$ picks up only the contribution at $k_{\text{peak}}$, a quick estimate of the previous term compared with the one in eq. (3.13) would be
\[ \frac{B_{\hat{R}}^{(d_1)}}{B_{\hat{R}}^{(\text{red})}} = O(\eta_{\text{II}}^3 P_R(k_{\text{peak}})) , \quad \text{with} \quad P_R(k_{\text{peak}}) \sim 10^{-2}. \tag{A.33} \]

The point to make here is that terms like the one in eq. (A.32) are usually tossed away in slow-roll inflation models since highly suppressed by $P_R \sim 10^{-9}$. In the presence of USR, one can not take lightly the same conclusion since the power spectrum can be much larger than the value suggested by CMB measurements. However, the suppression in eq. (A.33) might still allow to neglect these contributions. We can compute numerically the ratio in eq. (A.33) and check our estimate. It is worth doing this check explicitly since this is the largest correction to our approximation. We show our results, computed in the context of the numerical model, in figure 10. The suppression estimated in eq. (A.33) is confirmed.

e) Consider the effect of the field redefinition on the four-point correlator 
\[ \langle \hat{R}(\tau, \vec{k}_1)\hat{R}(\tau, \vec{k}_2)\hat{R}(\tau, \vec{k}_3)\hat{R}(\tau, \vec{k}_4) \rangle. \]
In Fourier space, the field redefinition generates,
Figure 10. Ratio in eq. (A.33) computed numerically. We consider the numerical model, and we fix $\eta_{H} = \eta_{0} = -0.75$. In the left panel, we show the ratio in the equilateral configuration. In the right panel, we use the parametrization in eq. (3.5) for fixed $K$ and three different choices for $\beta$.

at order $O(c_{2}^{2})$, the following 6 terms

\[ \frac{c_{2}^{2}}{16} \left[ (\hat{R}_{k_{1}}(\tau, k_{1})\hat{R}_{k_{2}}(\tau, k_{2})\hat{R}_{k_{3}}(\tau, k_{3})\hat{R}_{k_{4}}(\tau, k_{4}))) + (\hat{R}_{k_{1}}^{2}(\tau, k_{1})\hat{R}_{k_{2}}(\tau, k_{2})\hat{R}_{k_{3}}(\tau, k_{3})\hat{R}_{k_{4}}(\tau, k_{4})) + (\hat{R}_{k_{1}}(\tau, k_{1})\hat{R}_{k_{2}}^{2}(\tau, k_{2})\hat{R}_{k_{3}}(\tau, k_{3})\hat{R}_{k_{4}}(\tau, k_{4})) + (\hat{R}_{k_{1}}(\tau, k_{1})\hat{R}_{k_{2}}^{2}(\tau, k_{2})\hat{R}_{k_{3}}(\tau, k_{3})\hat{R}_{k_{4}}^{2}(\tau, k_{4})) + (\hat{R}_{k_{1}}^{2}(\tau, k_{1})\hat{R}_{k_{2}}^{2}(\tau, k_{2})\hat{R}_{k_{3}}(\tau, k_{3})\hat{R}_{k_{4}}^{2}(\tau, k_{4})) + (\hat{R}_{k_{1}}^{2}(\tau, k_{1})\hat{R}_{k_{2}}^{2}(\tau, k_{2})\hat{R}_{k_{3}}^{2}(\tau, k_{3})\hat{R}_{k_{4}}^{2}(\tau, k_{4})) \right] . \]

(A.34)

Each one of them can be computed in the vacuum of the free theory. Consider for simplicity only the first one, which we write in the form

\[ \frac{c_{2}^{2}}{16} \int \frac{d^{3}q_{1}}{(2\pi)^{3}} \frac{d^{3}q_{2}}{(2\pi)^{3}} \hat{R}_{I}(\tau, k_{1})\hat{R}_{I}(\tau, k_{2})\hat{R}_{I}(\tau, q_{1})\hat{R}_{I}(\tau, q_{2})\hat{R}_{I}(\tau, \bar{k}_{3} - \bar{q}_{1})\hat{R}_{I}(\tau, \bar{k}_{4} - \bar{q}_{2}) . \]

(A.35)

It is easy to see that there are 8 possible contractions, and one obtains

\[ \frac{c_{2}^{2}}{8} \left\{ \Delta(k_{1})\Delta(k_{2}) \left[ \Delta(|\bar{k}_{1} + \bar{k}_{3}|) + \Delta(|\bar{k}_{1} + \bar{k}_{4}|) + \Delta(|\bar{k}_{2} + \bar{k}_{3}|) + \Delta(|\bar{k}_{2} + \bar{k}_{4}|) \right] \right\} , \]

(A.36)

where the four terms in the curly brackets counted twice. Putting all together, eq. (A.34) reads

\[ \frac{c_{2}^{2}}{8} \left[ \Delta(k_{1})\Delta(k_{2})\Delta(|\bar{k}_{1} + \bar{k}_{4}|) + 23 \text{ permutations} \right] , \]

(A.37)

where the $4! = 24$ terms in the square brackets counts the permutations of the indices $1234$. This is the same structure that we introduced for the reduced trispectrum in eq. (4.2). If we identify $c_{2}/4 = 3f_{NL}/5$, the coefficient in front of eq. (A.37) becomes $18f_{NL}^{2}/25$. 

- 36 -
References

[1] Ya. B. Zel’dovich and I.D. Novikov, The Hypothesis of Cores Retarded during Expansion and the Hot Cosmological Model, Astronom. Zh. 43 (1966) 758.
[2] S. Hawking, Gravitationally collapsed objects of very low mass, Mon. Not. Roy. Astron. Soc. 152 (1971) 75 [insPIRE].
[3] B.J. Carr and S.W. Hawking, Black holes in the early Universe, Mon. Not. Roy. Astron. Soc. 168 (1974) 399 [insPIRE].
[4] J.M. Maldacena, Non-Gaussian features of primordial fluctuations in single field inflationary models, JHEP 05 (2003) 013 [astro-ph/0210603] [insPIRE].
[5] S.M. Leach and A.R. Liddle, Inflationary perturbations near horizon crossing, Phys. Rev. D 63 (2001) 043508 [astro-ph/0010082] [insPIRE].
[6] S.M. Leach, M. Sasaki, D. Wands and A.R. Liddle, Enhancement of superhorizon scale inflationary curvature perturbations, Phys. Rev. D 64 (2001) 023512 [astro-ph/0101406] [insPIRE].
[7] N.C. Tsamis and R.P. Woodard, Improved estimates of cosmological perturbations, Phys. Rev. D 69 (2004) 084005 [astro-ph/0307463] [insPIRE].
[8] W.H. Kinney, Horizon crossing and inflation with large eta, Phys. Rev. D 72 (2005) 023515 [gr-qc/0503017] [insPIRE].
[9] W.H. Kinney, A Hamilton-Jacobi approach to nonslow roll inflation, Phys. Rev. D 56 (1997) 2002 [hep-ph/9702427] [insPIRE].
[10] P. Ivanov, P. Naselsky and I. Novikov, Inflation and primordial black holes as dark matter, Phys. Rev. D 50 (1994) 7173 [insPIRE].
[11] R. Saito, J. Yokoyama and R. Nagata, Single-field inflation, anomalous enhancement of superhorizon fluctuations, and non-Gaussianity in primordial black hole formation, JCAP 06 (2008) 024 [arXiv:0804.3470] [insPIRE].
[12] A.A. Starobinsky, Spectrum of adiabatic perturbations in the universe when there are singularities in the inflation potential, JETP Lett. 55 (1992) 489 [insPIRE].
[13] A.M. Green, A.R. Liddle, K.A. Malik and M. Sasaki, A New calculation of the mass fraction of primordial black holes, Phys. Rev. D 70 (2004) 041502 [astro-ph/0403181] [insPIRE].
[14] C. Germani and R.K. Sheth, Nonlinear statistics of primordial black holes from Gaussian curvature perturbations, Phys. Rev. D 101 (2020) 063520 [arXiv:1912.07072] [insPIRE].
[15] V. De Luca, G. Franciolini, A. Kehagias, M. Peloso, A. Riotto and C. Ünal, The Ineludible non-Gaussianity of the Primordial Black Hole Abundance, JCAP 07 (2019) 048 [arXiv:1904.00970] [insPIRE].
[16] C.-M. Yoo, T. Harada, J. Garriga and K. Kohri, Primordial black hole abundance from random Gaussian curvature perturbations and a local density threshold, PTEP 2018 (2018) 123E01 [arXiv:1805.03946] [insPIRE].
[17] M. Kawasaki and H. Nakatsuka, Effect of nonlinearity between density and curvature perturbations on the primordial black hole formation, Phys. Rev. D 99 (2019) 123501 [arXiv:1903.02994] [insPIRE].
[18] S. Young, I. Musco and C.T. Byrnes, Primordial black hole formation and abundance: contribution from the non-linear relation between the density and curvature perturbation, JCAP 11 (2019) 012 [arXiv:1904.00984] [insPIRE].
[19] A. Kehagias, I. Musco and A. Riotto, Non-Gaussian Formation of Primordial Black Holes: Effects on the Threshold, JCAP 12 (2019) 029 [arXiv:1906.07135] [insPIRE].
[20] C.-M. Yoo, J.-O. Gong and S. Yokoyama, Abundance of primordial black holes with local non-Gaussianity in peak theory, JCAP 09 (2019) 033 [arXiv:1906.06790] [nSPIRE].

[21] C.-M. Yoo, T. Harada, S. Hirano and K. Kohri, Abundance of Primordial Black Holes in Peak Theory for an Arbitrary Power Spectrum, PTEP 2021 (2021) 013E02 [arXiv:2008.02425] [nSPIRE].

[22] I. Musco, V. De Luca, G. Franciolini and A. Riotto, Threshold for primordial black holes. II. A simple analytic prescription, Phys. Rev. D 103 (2021) 063538 [arXiv:2011.03014] [nSPIRE].

[23] C.-M. Yoo, T. Harada, S. Hirano and K. Kohri, Abundance of Primordial Black Holes in Peak Theory for an Arbitrary Power Spectrum, PTEP 2021 (2021) 013E02 [arXiv:2008.02425] [nSPIRE].

[24] I. Musco, J.C. Miller and L. Rezzolla, Computations of primordial black hole formation, Class. Quant. Grav. 22 (2005) 1405 [gr-qc/0412063] [nSPIRE].

[25] S. Young, The primordial black hole formation criterion re-examined: Parametrisation, timing and the choice of window function, Int. J. Mod. Phys. D 29 (2019) 2030002 [arXiv:1905.01230] [nSPIRE].

[26] J.M. Bardeen, J.R. Bond, N. Kaiser and A.S. Szalay, The Statistics of Peaks of Gaussian Random Fields, Astrophys. J. 304 (1986) 15 [nSPIRE].

[27] C. Germani and I. Musco, Abundance of Primordial Black Holes Depends on the Shape of the Inflationary Power Spectrum, Phys. Rev. Lett. 122 (2019) 141302 [arXiv:1805.04087] [nSPIRE].

[28] I. Musco, Threshold for primordial black holes: Dependence on the shape of the cosmological perturbations, Phys. Rev. D 100 (2019) 123524 [arXiv:1809.02127] [nSPIRE].

[29] D.W. Neilsen and M.W. Choptuik, Critical phenomena in perfect fluids, Class. Quant. Grav. 17 (2000) 761 [gr-qc/9812053].

[30] A. Kalaja et al., From Primordial Black Holes Abundance to Primordial Curvature Power Spectrum (and back), JCAP 10 (2019) 031 [arXiv:1908.03596] [nSPIRE].

[31] F. Riccardi, M. Taoso and A. Urbano, Solving peak theory in the presence of local non-gaussianities, to appear.

[32] G. Ballesteros, J. Rey, M. Taoso and A. Urbano, Primordial black holes as dark matter and gravitational waves from single-field polynomial inflation, JCAP 07 (2020) 025 [arXiv:2001.08220] [nSPIRE].

[33] G. Ballesteros, J. Rey, M. Taoso and A. Urbano, Stochastic inflationary dynamics beyond slow-roll and consequences for primordial black hole formation, JCAP 08 (2020) 043 [arXiv:2006.14597] [nSPIRE].

[34] B.J. Carr, K. Kohri, Y. Sendouda and J. Yokoyama, New cosmological constraints on primordial black holes, Phys. Rev. D 81 (2010) 104019 [arXiv:0912.5297] [nSPIRE].

[35] R. Laha, Primordial Black Holes as a Dark Matter Candidate Are Severely Constrained by the Galactic Center 511keV γ-Ray Line, Phys. Rev. Lett. 123 (2019) 251101 [arXiv:1906.09994] [nSPIRE].

[36] H. Niikura et al., Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations, Nature Astron. 3 (2019) 524 [arXiv:1701.02151] [nSPIRE].

[37] A. Katz, J. Kopp, S. Sibiryakov and W. Xue, Femtolensing by Dark Matter Revisited, JCAP 12 (2018) 005 [arXiv:1807.11495] [nSPIRE].

[38] F. Capela, M. Pshirkov and P. Tinyakov, Constraints on primordial black holes as dark matter candidates from capture by neutron stars, Phys. Rev. D 87 (2013) 123524 [arXiv:1301.4984] [nSPIRE].
[39] P. Pani and A. Loeb, *Tidal capture of a primordial black hole by a neutron star: implications for constraints on dark matter*, *JCAP* **06** (2014) 026 [arXiv:1401.3025] [inSPIRE].

[40] P. Pani and A. Loeb, *Tidal capture of a primordial black hole by a neutron star: implications for constraints on dark matter*, *JCAP* **06** (2014) 026 [arXiv:1401.3025] [inSPIRE].

[41] G. Defillon, E. Granet, P. Tinyakov and M.H.G. Tytgat, *Tidal capture of primordial black holes by neutron stars*, *Phys. Rev. D* **90** (2014) 103522 [arXiv:1409.0469] [inSPIRE].

[42] P. Montero-Camacho, X. Fang, G. Vasquez, M. Silva and C.M. Hirata, *Revisiting constraints on asteroid-mass primordial black holes as dark matter candidates*, *JCAP* **08** (2019) 031 [arXiv:1906.05950] [inSPIRE].

[43] P.W. Graham, S. Rajendran and J. Varela, *Dark Matter Triggers of Supernovae*, *Phys. Rev. D* **92** (2015) 063007 [arXiv:1505.04444] [inSPIRE].

[44] Planck collaboration, *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9 [arXiv:1905.05697] [inSPIRE].

[45] S. Matarrese, F. Lucchin and S.A. Bonometto, *A path integral approach to large scale matter distribution originated by non-Gaussian fluctuations*, *Astrophys. J. Lett.* **310** (1986) L21 [inSPIRE].

[46] G. Franciolini, A. Kehagias, S. Matarrese and A. Riotto, *Primordial Black Holes from Inflation and non-Gaussianity*, *JCAP* **03** (2018) 016 [arXiv:1801.09415] [inSPIRE].

[47] Y.-F. Cai, X. Chen, M.H. Namjoo, M. Sasaki, D.-G. Wang and Z. Wang, *Revisiting non-Gaussianity from non-attractor inflation models*, *JCAP* **05** (2018) 012 [arXiv:1712.09998] [inSPIRE].

[48] S. Passaglia, W. Hu and H. Motohashi, *Primordial black holes and local non-Gaussianity in canonical inflation*, *Phys. Rev. D* **99** (2019) 043536 [arXiv:1812.09998] [inSPIRE].

[49] E. Komatsu, *Hunting for Primordial Non-Gaussianity in the Cosmic Microwave Background*, *Class. Quant. Grav.* **27** (2010) 124010 [arXiv:1003.6097] [inSPIRE].

[50] L. Boubekeur and D.H. Lyth, *Detecting a small perturbation through its non-Gaussianity*, *Phys. Rev. D* **73** (2006) 021301 [astro-ph/0504046] [inSPIRE].

[51] A. Stuart and K. Ord, *Kendall’s Advanced Theory of Statistics. Volume I: Distribution Theory*, Oxford University Press, Oxford U.K. (1987).

[52] J.E. Kolassa, *Series Approximation Methods in Statistics*, Lecture Notes in Statistics, Springer, Berlin Germany (2006).

[53] T. Harada, C.-M. Yoo, T. Nakama and Y. Koga, *Cosmological long-wavelength solutions and primordial black hole formation*, *Phys. Rev. D* **91** (2015) 084057 [arXiv:1503.03934] [inSPIRE].

[54] G. Ballesteros and M. Taoso, *Primordial black hole dark matter from single field inflation*, *Phys. Rev. D* **97** (2018) 023501 [arXiv:1709.05565] [inSPIRE].

[55] I. Dalianis, A. Kehagias and G. Tringas, *Primordial black holes from α-attractors*, *JCAP* **01** (2019) 037 [arXiv:1805.09483] [inSPIRE].

[56] V. Atal and C. Germani, *The role of non-Gaussianities in Primordial Black Hole formation*, *Phys. Dark Univ.* **24** (2019) 100275 [arXiv:1811.07857] [inSPIRE].

[57] A.A. Starobinsky, *Stochastic de Sitter (inflationary) stage in the early universe*, *Lect. Notes Phys.* **246** (1986) 107 [inSPIRE].

[58] M. Biagetti, G. Franciolini, A. Kehagias and A. Riotto, *Primordial Black Holes from Inflation and Quantum Diffusion*, *JCAP* **07** (2018) 032 [arXiv:1804.07124] [inSPIRE].

[59] J.M. Ezquiaga and J. García-Bellido, *Quantum diffusion beyond slow-roll: implications for primordial black-hole production*, *JCAP* **08** (2018) 018 [arXiv:1805.06731] [inSPIRE].
D. Cruces, C. Germani and T. Prokopec, *Failure of the stochastic approach to inflation beyond slow-roll*, JCAP 03 (2019) 048 [arXiv:1807.09057] [inSPIRE].

H. Firouzjahi, A. Nassiri-Rad and M. Noorbala, *Stochastic Ultra Slow Roll Inflation*, JCAP 01 (2019) 040 [arXiv:1811.02175] [inSPIRE].

C. Pattison, V. Vennin, H. Assadullahi and D. Wands, *Stochastic inflation beyond slow roll*, JCAP 07 (2019) 031 [arXiv:1905.06300] [inSPIRE].

J.M. Ezquiaga, J. García-Bellido and V. Vennin, *The exponential tail of inflationary fluctuations: consequences for primordial black holes*, JCAP 03 (2020) 029 [arXiv:1912.05399] [inSPIRE].

C. Pattison, V. Vennin, D. Wands and H. Assadullahi, *Ultra-slow-roll inflation with quantum diffusion*, JCAP 04 (2021) 080 [arXiv:2101.05741] [inSPIRE].

A. De and R. Mahbub, *Numerically modeling stochastic inflation in slow-roll and beyond*, Phys. Rev. D 102 (2020) 123509 [arXiv:2010.12685] [inSPIRE].

D.G. Figueroa, S. Raatikainen, S. Rasanen and E. Tomberg, *Non-Gaussian tail of the curvature perturbation in stochastic ultra-slow-roll inflation: implications for primordial black hole production*, arXiv:2012.06551 [inSPIRE].

R. Bravo and G.A. Palma, *Unifying attractor and non-attractor models of inflation under a single soft theorem*, arXiv:2009.03369 [inSPIRE].

J.S. Schwinger, *Brownian motion of a quantum oscillator*, J. Math. Phys. 2 (1961) 407 [inSPIRE].

L.V. Keldysh, *Diagram technique for nonequilibrium processes*, Zh. Eksp. Teor. Fiz. 47 (1964) 1515 [inSPIRE].

R.D. Jordan, *Effective Field Equations for Expectation Values*, Phys. Rev. D 33 (1986) 444 [inSPIRE].

E. Calzetta and B.L. Hu, *Closed Time Path Functional Formalism in Curved Space-Time: Application to Cosmological Back Reaction Problems*, Phys. Rev. D 35 (1987) 495 [inSPIRE].

S. Weinberg, *Quantum contributions to cosmological correlations*, Phys. Rev. D 72 (2005) 043514 [hep-th/0506236] [inSPIRE].

X. Chen, M.-x. Huang, S. Kachru and G. Shiu, *Observational signatures and non-Gaussianities of general single field inflation*, JCAP 01 (2007) 002 [hep-th/0605045] [inSPIRE].

V. Atal, J. Garriga and A. Marcos-Caballero, *Primordial black hole formation with non-Gaussian curvature perturbations*, JCAP 09 (2019) 073 [arXiv:1905.13202] [inSPIRE].

M.H. Namjoo, H. Firouzjahi and M. Sasaki, *Violation of non-Gaussianity consistency relation in a single field inflationary model*, EPL 101 (2013) 39001 [arXiv:1210.3692] [inSPIRE].