# Intrinsic non-Gaussianity of ultra slow-roll inflation

Guillermo Ballesteros[1,2], Jesús Gambín Egea[2], Thomas Konstandin[3],
Alejandro Pérez Rodríguez[1,2], Mathias Pierre[3], and Julián Rey[3]

[1]*Departamento de Física Teórica, Universidad Autónoma de Madrid (UAM),
Campus de Cantoblanco, 28049 Madrid, Spain*

[2]*Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain*

[3]*Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*

### Abstract

We study the non-Gaussian tail of the curvature fluctuation, $\zeta$, in an inflationary scenario with a transient ultra slow-roll phase that generates a localized large enhancement of the spectrum of $\zeta$. To do so, we implement a numerical procedure that provides the probability distribution of $\zeta$ order by order in perturbation theory. The non-Gaussianities of $\zeta$ can be shown to arise from its non-linear relation to the inflaton fluctuations and from the *intrinsic* non-Gaussianities of the latter, which stem from its self interactions. We find that intrinsic non-Gaussianities, which have often been ignored to estimate the abundance of primordial black holes in this kind of scenario, are important. The relevance of the intrinsic contribution depends on the rapidity with which the transient ultra slow-roll phase occurs, as well as on its duration. Our method cannot be used accurately when the perturbative in-in formalism fails to apply, highlighting the relevance of developing fully non-perturbative approaches to the problem.

guillermo.ballesteros@uam.es, j.gambin@csic.es, thomas.konstandin@desy.de, alejandro.perezrodriguez@uam.es,
mathias.pierre@desy.de, julian.rey@desy.de

# Contents

# 1 Introduction

In recent years, considerable effort has been devoted to understanding inflation beyond the slow-roll regime. Particular attention has been paid to models featuring a phase of ultra slow-roll (USR) [1], which is realized in single-field inflation models whose potential has a non-stationary inflection point, which makes the inflaton slow down. In this scenario, the Fourier modes of scalar perturbations exiting the horizon around the USR phase are enhanced, leading to a peak in the primordial power spectrum at those scales, which can eventually trigger the formation of a significant amount of primordial black holes (PBH) (see e.g. [2–4] for concrete examples of models). These PBH may account for the dark matter (see [5] for a recent review on PBH bounds). Concurrently, the breakdown of slow-roll around the inflection point leads to non-Gaussianities that are also enhanced around the same scales. Since the PBH abundance is highly sensitive to the details of the probability distribution function (PDF) of curvature fluctuations at large values, it is important to determine it as accurately as possible. Indeed, there is growing interest on the characterization of the tail of said PDF and, more in general, on studying inflationary perturbations in the non-perturbative regime (see e.g. [6–9] for results from various perspectives).

In this paper, we focus on the non-Gaussianity arising in single-field inflation models with a phase of USR and consider its manifestation on the tail of the PDF of curvature fluctuations. Much of the literature on the topic, specifically in the context of PBH formation, has followed one of two approaches:

- The computation of high-order (three or larger) correlators of $\zeta$ –the curvature perturbation on slices of uniform density–, which encode information about its non-Gaussian statistics around $\zeta = 0$.

These calculations are usually performed in the $\zeta$-gauge, in which the only scalar degree of freedom is placed in the spatial part of the metric, see in particular [10,11] for PBH formation from USR. The possibility of extracting information about the tail of the PDF for large $\zeta$ from these correlators is hampered by their own definition. See [8] for a clear introductory discussion illustrating this difficulty. However, there are several works which have studied the effects of perturbative non-Gaussianities on the PBH abundance [12–15].

- The computation of the non-Gaussian PDF of $\zeta$ by performing a non-linear transformation of the PDF of $\delta\phi$ (perturbation of the inflaton field), which is *assumed* to be Gaussian. The non-linear relation between the two variables can be obtained from the so-called $\delta N$ formalism [16,17], which provides the gauge transformation between $\zeta$ and $\delta\phi$ (in the $\delta\phi$-gauge, where the single scalar degree of freedom is placed in the matter field) to all orders in perturbations and second order in gradients, see also [18].[1] Examples of this approach are [23–29].

In this paper, we asses the validity of assuming a Gaussian $\delta\phi$ in the latter approach. This can be done by computing the bispectrum of $\delta\phi$ and higher-order correlators of this variable and using them to construct a non-Gaussian input for the non-linear relation between $\zeta$ and $\delta\phi$ in the $\delta N$ formalism.[2] For simplicity, we restrict our analysis to including the effect of the bispectrum of $\delta\phi$, but our analysis can be generalized to include higher-order correlators as well. Our approach can be decomposed in three steps:

1. We compute the tree-level power spectrum and bispectrum for $\delta\phi$ in the $\delta\phi$-gauge using the in-in formalism. This only requires knowing the Lagrangian for fluctuations of $\delta\phi$ at cubic order. As we mentioned above, our procedure can be generalized to include higher-order correlators. It can also accommodate loop corrections.

2. On a spatial lattice, we construct a random variable endowed with the spectrum and bispectrum computed for $\delta\phi$ in step 1. We sample the distribution of this random field (at the end of inflation) to get a proxy for the non-Gaussian PDF of $\delta\phi$.

3. We use the non-linear relation between $\delta\phi$ and $\zeta$ (encoded in the gauge transformation provided by the $\delta N$ formalism) to compute the non-Gaussian PDF of $\zeta$ from the one of $\delta\phi$ (obtained in the previous step).

We can distinguish two formally different contributions to the non-Gaussianity of $\zeta$. The first is the non-Gaussianity due to the *non-linear* relation between $\delta\phi$ and $\zeta$ (assuming Gaussian $\delta\phi$). The second is the correction over the former due to the non-Gaussianity of $\delta\phi$, coming from its self interactions, which we account for perturbatively. This correction will henceforth be referred to as *intrinsic*.[3] Of course, this distinction is artificial; if we were to compute the PDF of $\zeta$ in the $\zeta$-gauge, we would simply use the self interactions of $\zeta$. However, the distinction between the two contributions in the $\delta\phi$-gauge is convenient for two reasons. First, we are indeed interested in checking the commonly used assumption of Gaussianity for

---

[1]This non-linear relation from the $\delta N$ formalism can also be implemented in the context of stochastic inflation [6,7,19], rendering an equivalent result for the PDF of $\zeta$ from an USR phase [20] (due to the presence of an attractor and the fact that the power spectrum is peaked). See also [21,22].

[2]This is different from what was done in [30], where the bispectrum of $\delta\phi$ was used as input for a *linear* relation with $\zeta$ to compute the bispectrum of the latter.

[3]In the literature on PBH formation, the word *intrinsic* has been used earlier to refer to the non-Gaussianity in density fluctuations arising from their non-linear relation to curvature fluctuations, see e.g. [14,31,32]. That is a different notion from the one we are interested in in this work.

$\delta\phi$ in the derivation of the PDF of $\zeta$. Second, using the $\delta\phi$-gauge the calculations are simpler (especially for higher-order correlators and at higher loop order, see [18]).

We consider an USR phase preceded by a standard slow-roll (SR) phase, followed by a constant-roll (CR) phase, characterized by a constant second slow roll parameter $\eta = \eta_{\rm CR} < 0$ (which ensures that $\epsilon$ grows and inflation eventually ends). We parametrize the slow-roll parameters in such a way that the duration and smoothness of the USR phase can be controlled. We find that the bispectrum of $\delta\phi$ can introduce sizable corrections to the PDF of $\zeta$, depending on the duration of the phase and smoothness of the transitions in and out of it. However, in such a regime the approach may not allow to derive accurate predictions for the tail of the PDF. This is somewhat analogous to what happens with the one-loop corrections to the scalar power spectrum, which are also controlled by the cubic term in the Lagrangian for $\delta\phi$ [18].

The paper is structured as follows. In Section 2, we review the in-in formalism, discuss the cubic interactions of $\delta\phi$ in the $\delta\phi$-gauge and the gauge transformation to the $\zeta$-gauge. In Section 3, we compute the bispectrum and use Maldacena's consistency relation as a check that we have included all relevant interactions. In Section 4, we implement a numerical procedure to generate a random field in a lattice with the correct power spectrum and bispectrum, and determine the resulting PDF of $\zeta$. The conservation of the bispectrum of $\zeta$ on superhorizon scales is discussed in Appendix A. In Appendix B we discuss how the Gram-Charlier series, which has sometimes been used to include perturbative non-Gaussian effects on the PDF of $\zeta$, fails in the present case.

## 2  Fluctuations in ultra slow-roll inflation

### 2.1  The inflationary setup

Throughout the paper, we set $\hbar = c = 1$, and we denote as $M_P = (8\pi\,G)^{-1/2}$ the reduced Planck mass. The starting point for our analysis is the canonical action for a minimally coupled scalar (inflaton) field $\Phi(t, \boldsymbol{x})$ with potential $V$,

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

where $R$ denotes the Ricci scalar. For the background (homogeneous) inflaton field $\phi(t)$, the metric reduces to FLRW, with scale factor $a(t)$. The (homogeneous) background equations of motion for the field and metric are

$$\phi'' = -(3 - \epsilon)(\phi' + M_P^2\,\partial_\phi V/V), \qquad M_P^2\,H^2(3 - \epsilon) = V\,, \tag{2}$$

where we use the amount of expansion measured in e-folds, $N$ (defined via $\mathrm{d}N = H\,\mathrm{d}t = \dot{a}/a\,\mathrm{d}t$) as time variable. Primes denote the corresponding derivatives, $' \equiv \mathrm{d}/\mathrm{d}N$, and dots indicate standard (cosmic) time derivatives, $\dot{} \equiv \mathrm{d}/\mathrm{d}t$. The quantity $\epsilon$ appearing in Eq. (2) is the first slow-roll parameter, whose definition is

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{1}{2M_P^2} \frac{\dot{\phi}^2}{H^2}\,. \tag{3}$$

For convenience, we also introduce the second slow-roll parameter

$$\eta \equiv -\frac{1}{2} \frac{\dot{\epsilon}}{H\epsilon}\,, \tag{4}$$

4

which measures the acceleration of the background inflaton field.

We are interested in inflationary models with a potential featuring a non-stationary inflection point. In these models, the inflaton goes first through a SR phase (with $\epsilon, |\eta| \ll 1$).[4] This stage ends as the inflaton approaches the local minimum, and is followed by an USR phase (with $\epsilon \ll 1$ and $\eta \geq 3/2$) in which the inflaton slows down as it climbs up the local maximum of the potential. Finally, the dynamics undergo a CR phase, characterized by a constant $\eta$, as the inflaton rolls down the local maximum towards the absolute minimum of the potential. During this last phase, the first slow-roll parameter $\epsilon$ grows until inflation ends. The slow-roll parameter $\epsilon$ remains small throughout the entire evolution and becomes of $\mathcal{O}(1)$ only at the very end of inflation.

We study perturbations to the background solution using the ADM formalism [33]. The metric is parametrized as

$$\mathrm{d}s^2 = -\mathcal{N}^2 \mathrm{d}t^2 + h_{ij}\big(\mathrm{d}x^i + \mathcal{N}^i \mathrm{d}t\big)\big(\mathrm{d}x^j + \mathcal{N}^j \mathrm{d}t\big), \tag{5}$$

where the lapse $\mathcal{N}$ and shift $\mathcal{N}_i$ are Lagrange multipliers that must be determined via their (non-dynamical) equations of motion and inserted back into the action. To perform the calculation, it is convenient to fix a gauge. Let us split the field into a homogeneous piece and a perturbation

$$\Phi(t, \boldsymbol{x}) = \phi(t) + \delta\phi(t, \boldsymbol{x}). \tag{6}$$

We can either work in the $\zeta$-gauge, defined by the conditions

$$\delta\phi(t, \boldsymbol{x}) = 0, \qquad h_{ij} = a^2 e^{2\zeta}\delta_{ij}, \tag{7}$$

or in the $\delta\phi$-gauge, defined by

$$\zeta(t, \boldsymbol{x}) = 0, \qquad h_{ij} = a^2 \delta_{ij}. \tag{8}$$

In this way, by using a coordinate transformation, we place the only scalar degree of freedom in the theory either in the metric or in the matter sector, respectively. Throughout the rest of this work, we neglect the effect of vector and tensor perturbations. The full action for the perturbations in either gauge contains, prior to any approximation, multiple interaction terms at any order, due to the non-linearities introduced by the lapse and shift. However, since the first slow-roll parameter $\epsilon$ is small during ultra slow-roll –in fact, it remains small until inflation approaches its end– and the second parameter $\eta$ is of $\mathcal{O}(1)$ at most and many interaction terms are proportional to powers of $\epsilon$, the action simplifies considerably for the kind of models we are interested in. This is more evident in the $\delta\phi$-gauge, which is also convenient, since our goal is to explore the relevance of the non-Gaussianity of $\delta\phi$.

## 2.2 The in-in formalism

We will use the in-in formalism [34] –see also [35,36] and Appendix A of [18] for reviews of various lengths– to compute the bispectrum of $\delta\phi$. Here we summarize a few basic equations of the formalism that will be useful later on.

Given a general Hermitian operator[5] $\hat{\mathcal{O}}(t)$ in the Heisenberg picture, its vacuum expectation value is:

$$\left\langle \hat{\mathcal{O}}(t) \right\rangle = \langle 0| \, \hat{F}^{-1}(t, -\infty_+)\hat{\mathcal{O}}_I(t)\hat{F}(t, -\infty_-) \, |0\rangle \bigg|_{\text{no bubbles}}, \tag{9}$$

---

[4]The large-scale fluctuations observed imprinted in the CMB correspond to this phase in these models.

[5]We will denote quantum operators and quantized fields with "hats", ˆ.

where the subscript $_I$ carried by the operator $\hat{\mathcal{O}}_I(t)$ denotes the interaction picture and $|0\rangle$ represents the interaction vacuum.[6] The time evolution operator in the interaction picture is:

$$\hat{F}(t,t_0) = T\exp\left(-i\int_{t_0}^t \mathrm{d}t'\hat{H}_I(t')\right) \quad \text{and} \quad \hat{F}^{-1}(t,t_0) = \overline{T}\exp\left(i\int_{t_0}^t \mathrm{d}t'\hat{H}_I(t')\right), \tag{10}$$

where $T$ and $\overline{T}$ denote time-ordering and anti time-ordering, respectively. We stress that $\left\langle \hat{\mathcal{O}}(t)\right\rangle$ does not receive bubble contributions [37], so all insertions of $\hat{H}_I$ must be connected to the operator $\hat{\mathcal{O}}_I$. In practice, we will only need to calculate expectation values with a single Hamiltonian insertion (since we are interested in the bispectrum of $\delta\phi$), so that

$$\left\langle \hat{\mathcal{O}}(t)\right\rangle = \langle 0|\,\hat{\mathcal{O}}_I(t)\,|0\rangle + 2\,\mathrm{Im}\left\{\int_{-\infty_-}^t \mathrm{d}t'\,\langle 0|\,\hat{\mathcal{O}}_I(t)\hat{H}_I(t')\,|0\rangle\right\} + \mathcal{O}\left(\hat{H}_I^2\right). \tag{11}$$

The prescription $i\omega$ (with an infinitesimal $\omega > 0$ that eventually will be sent to 0) in the lower limit of time integrals, $-\infty_\pm \equiv -\infty(1\pm i\omega)$, ensures that the system projects onto the interaction picture vacuum $|0\rangle$ at times $t\to-\infty$. The operators $\hat{\mathcal{O}}_I$ and $\hat{H}_I$, in the interaction picture, are composed of fields that evolve following the dynamics governed by the free Hamiltonian. Therefore, these fields can be decomposed into creation and annihilation operators as usual. For the fluctuations of the inflaton in this picture, we have

$$\hat{\delta\phi}(t,\boldsymbol{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}e^{i\boldsymbol{kx}}\hat{\delta\phi}_{\boldsymbol{k}}(t), \quad \text{where} \quad \hat{\delta\phi}_{\boldsymbol{k}}(t) = \delta\phi_k(t)\hat{a}_{\boldsymbol{k}} + \delta\phi_k^*(t)\hat{a}_{-\boldsymbol{k}}^\dagger. \tag{12}$$

The creation and annihilation operators, $\hat{a}_{\boldsymbol{k}}^\dagger$ and $\hat{a}_{\boldsymbol{k}}$, satisfy $\left[\hat{a}_{\boldsymbol{k}},\,\hat{a}_{\boldsymbol{p}}^\dagger\right] = (2\pi)^3\delta\left(\boldsymbol{k}-\boldsymbol{p}\right)$. The Fourier modes $\delta\phi_k(t)$ satisfy the free equation of motion which, in this case, at first order in the SR parameter $\epsilon$, is

$$\ddot{\delta\phi}_k + 3H\dot{\delta\phi}_k + \left(\frac{k^2}{a^2} + \partial_\phi^2 V\right)\delta\phi_k = 0, \tag{13}$$

where, neglecting terms suppressed by $\epsilon$, the second derivative of the potential can be written as

$$\partial_\phi^2 V \simeq -\frac{H^2}{2}\left(3\eta + \frac{\eta^2}{2} + \frac{\dot{\eta}}{H}\right). \tag{14}$$

Besides, $\delta\phi_k(t)$ satisfies Bunch-Davies initial conditions (which ensure canonical commutation relations):

$$\lim_{\tau\to-\infty}\delta\phi_k(\tau) = \frac{1}{a(\tau)}\frac{e^{-ik\tau}}{\sqrt{2k}}, \tag{15}$$

where the factor $1/a$ comes from the canonical normalization of $\delta\phi$, and $\tau$ is the conformal time defined according to $\mathrm{d}t = a(\tau)\mathrm{d}\tau$.

The $n$-point correlation functions in Fourier space are defined as

$$\left\langle \hat{\delta\phi}(t,\boldsymbol{x}_1)\cdots\hat{\delta\phi}(t,\boldsymbol{x}_n)\right\rangle \equiv \int \prod_{j=1}^n\left(\frac{\mathrm{d}^3k_j}{(2\pi)^3}e^{i\boldsymbol{k}_j\boldsymbol{x}_j}\right)\left\langle \hat{\delta\phi}_{\boldsymbol{k}_1}(t)\cdots\hat{\delta\phi}_{\boldsymbol{k}_n}(t)\right\rangle. \tag{16}$$

---

[6]In the Heisenberg picture, the vacuum state does not evolve in time, but the operators, such as $\hat{\mathcal{O}}(t)$ do evolve. In the interaction picture the fields evolve in time according to the free Hamiltonian (without interactions) and satisfy standard commutation relations; and the corresponding vacuum is the one of the free Hamiltonian. In the derivation of the in-in formula (9) one starts with the Heisenberg picture and then expresses fields and vacuum in that picture in terms of the ones in the interaction picture, see Appendix A of [18] for details.

The connected two-point correlation is related to the dimensionless power spectrum as follows:

$$\mathcal{P}_{\delta\phi}(t, \boldsymbol{k}) \, (2\pi)^3 \, \delta(\boldsymbol{k} + \boldsymbol{k}') \equiv \frac{k^3}{2\pi^2} \langle \delta\hat{\phi}_{\boldsymbol{k}}(t) \, \delta\hat{\phi}_{\boldsymbol{k}'}(t) \rangle_c \,. \tag{17}$$

The subscript $_c$ refers to the connected part of the correlation, defined as the contribution having a single Dirac delta that guarantees the momentum conservation.

## 2.3 The $\delta N$ formula

We now discuss the relation between the scalar degrees of freedom in the $\delta\phi$-gauge and the $\zeta$-gauge. For this purpose, it is necessary to consider a transformation connecting both gauges. Neglecting gradients and at late times,[7] the perturbation $\zeta$ (defined in the $\zeta$-gauge, see Eq. (7)) can be expressed as a perturbation in the local expansion rate [18]

$$\zeta = \log\left[\frac{a(t + \delta t)}{a(t)}\right], \tag{18}$$

where the infinitesimal time transformation $\delta t$ is the one connecting the $\delta\phi$-gauge (left-hand side) and the $\zeta$-gauge (right-hand side):

$$\phi(t + \delta t) + \delta\phi(t + \delta t) = \phi(t) \,. \tag{19}$$

Using Eq. (19), we can find $\delta t$ as a function of $\delta\phi$, and then use (18) to determine $\zeta$ as a function of $\delta\phi$ up to gradients and order by order in perturbations

$$\zeta = -\frac{H}{\dot{\phi}}\delta\phi - \frac{\eta}{2}\left(-\frac{H}{\dot{\phi}}\delta\phi\right)^2 + \frac{1}{3}\left(\eta^2 + \frac{\dot{\eta}}{2H}\right)\left(-\frac{H}{\dot{\phi}}\delta\phi\right)^3 + \mathcal{O}(\delta\phi^4). \tag{20}$$

To get to this expression, we have used the fact that, neglecting gradients and up to second order in perturbations, the equation of motion for $\delta\phi$ reads [18]

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\left(-\frac{H}{\dot{\phi}}\delta\phi\right) - \frac{\eta}{2}\left(-\frac{H}{\dot{\phi}}\delta\phi\right)^2\right] = \mathcal{O}(\delta\phi^3). \tag{21}$$

Incidentally, this equation makes manifest the conservation of $\zeta$ on super-Hubble scales ($k \ll aH$) up to second order in perturbations (see [38] for a proof to first order, and [16, 39] for a proof up to all orders).

It should be noted that (18) and (19) are nothing but the $\delta N$ formula, see [16], which non-linearly relates (on super-Hubble scales) the curvature perturbation on uniform density slices with the difference in local expansion due to a field perturbation on spatially-flat slices. This $\delta N$ *formalism* allows to compute (20) to all orders in $\delta\phi$ by just solving the background equations of motion for $\phi$ and $H$ with shifted initial conditions.

For the particular case in which $\eta$ is approximately constant (as it is the case in the CR phase after USR, when the Fourier modes in the peak of the curvature power spectrum freeze out), the background equations of motion can be solved analytically, and the $\delta N$ formula has a closed form [19, 20, 23, 24]

$$\eta_{\mathrm{CR}}\zeta = \log\left(1 - \eta_{\mathrm{CR}}H\frac{\delta\phi}{\dot{\phi}}\right), \tag{22}$$

where $\eta_{\mathrm{CR}}$ is the constant value of $\eta$ along the CR attractor. As a consistency check, notice that expanding (22) as a Taylor series around $\delta\phi = 0$, one recovers (20) for constant $\eta$.

---

[7]Neglecting gradients implies that the time $t$ must be chosen once the USR phase is over and the Fourier modes of interest for the problem we are studying (which defines a certain coarse-graining scale) are super-Hubble.

If one *assumes* $\delta\phi$ to be Gaussian, the PDF of $\zeta$ obtained from (22) has the usual exponential tail reported for USR models (see [20] for a recent discussion),

$$P(\zeta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[ -\frac{1}{2\sigma^2\eta_{\mathrm{CR}}^2} \left(1 - e^{\eta_{\mathrm{CR}}\zeta}\right)^2 + \eta_{\mathrm{CR}}\zeta \right], \tag{23}$$

where

$$\sigma^2 = \frac{H^2}{\dot{\phi}^2} \int \mathcal{P}_{\delta\phi}(k)\, \mathrm{d}\log k \tag{24}$$

is the variance of $\zeta$ in linear perturbation theory. In the integrand, $\mathcal{P}_{\delta\phi}$ is the dimensionless power spectrum of $\delta\phi$ defined in (17) which, at tree level, is

$$\mathcal{P}_{\delta\phi}(k) = \frac{k^3}{2\pi^2} |\delta\phi_k|^2 . \tag{25}$$

To complete the description of the PDF of $\zeta$, it is necessary to analyze which are the integration limits that appear in Eq. (24). These limits, or cutoffs, are dependent on the problem under consideration and are closely related to the observable we are interested in. In our case, as we have been discussing, we are interested in PBH formation in post-inflationary epochs; a process that occurs when a $\zeta$-mode enters the horizon with sufficient amplitude to generate a matter fluctuation that eventually collapses forming a PBH. To characterize the PDF of $\zeta$ in such a way that it captures the information associated with the PBH formation, we must choose the cutoffs in a way that resolves the physics related to the USR phase during inflation, i.e. we must capture the peak of the power spectrum in the variance, Eq. (24). This is because at these scales the probability of generating a fluctuation of $\zeta$ large enough to end up forming a PBH is maximal, since it is precisely on these scales that the power spectrum reaches its highest value.[8]

We are interested in trying to determine the deviation from (23) induced by the fact that $\delta\phi$ is not exactly a free field. In the following sections, we will study this issue by considering only the three-point function of $\delta\phi$.

## 2.4 Relevant interactions

We now determine the interaction terms relevant for the calculation of the bispectrum of $\delta\phi$ in the eponymous gauge. It is convenient to define the rescaled inflaton perturbation[9]

$$\varphi \equiv -\frac{H}{\dot{\phi}}\delta\phi . \tag{26}$$

The interaction Lagrangian at cubic order in the $\delta\phi$-gauge is [39]

$$\frac{\mathcal{L}_I^{(3)}}{M_P^2} = -c_0\, a^3 H^2 \varphi^3 - c_1\, a^3 \varphi \dot{\varphi}^2 - c_2\, a^3 \dot{\varphi}\, \partial_i\varphi\, \partial^i\left(\partial^{-2}\dot{\varphi}\right) - c_3\, a\, \varphi\, (\partial_i\varphi)^2$$

---

[8]There are different prescriptions to compute the PBH abundance –see e.g. [40] for a recent discussion– but all involve setting cutoffs via window functions. The specific prescription that is used (e.g. Press-Schechter vs peak theory) leads to quantitative differences in the PBH abundance, but it is not relevant for our purpose, i.e. illustrating the relevance of non-Gaussianities in the PDF of $\zeta$. These different approaches will not give qualitatively significant differences as long as the peak of the power spectrum lies inside the chosen cutoffs.

[9]This is a definition, there is no truncation in powers of $\delta\phi$. This variable coincides with the linear term (in powers of $\delta\phi$) of the comoving curvature perturbation in the flat gauge. It also coincides with the redefined field $\zeta_n$ introduced in [39] and used in [11]. We emphasize that (26) implies no change of gauge: $\varphi$ is the single scalar degree of freedom in the $\delta\phi$-gauge.

$$-c_4 \, a^3 \varphi \left(\partial_i \partial_j \left(\partial^{-2} \dot{\varphi}\right)\right)^2 + \frac{\mathrm{d}}{\mathrm{d}t}\left(c_5 \, a^3 H \varphi^3\right),\tag{27}$$

where the coefficients are given by

$$c_0 = \frac{1}{3}\epsilon \, \eta \, \epsilon_3 (\epsilon + 2\eta - \epsilon_3 - \epsilon_4 - 3)\tag{28}$$

and

$$c_1 = \frac{\epsilon^2}{2}(\epsilon - 2)\,,\quad c_2 = 2\epsilon^2\,,\quad c_3 = -\epsilon^2\,,\quad c_4 = -\frac{\epsilon^3}{2}\,,\quad c_5 = -\frac{\epsilon^2}{3}(\epsilon + 3\eta).\tag{29}$$

We define recursively $\epsilon_n = \dot{\epsilon}_{n-1}/(H\epsilon_{n-1})$, with $\epsilon_1 = \epsilon$ and $\epsilon_2 = -2\eta$. One can see that all the $c_n$ coefficients with $n \geq 1$ are suppressed by powers of $\epsilon \ll 1$ with respect to $c_0$. Similarly, during an USR phase, $\eta$ changes quickly and $\epsilon_3$ can grow significantly, so that the only relevant term in the presence of an USR phase is the inflaton self interaction $c_0 \, \varphi^3$. This term contains all contributions to $\mathcal{L}_I^{(3)}$ that are proportional to $\varphi^3$, and also those proportional to $\varphi^2 \dot{\varphi}$ after integrating by parts in time. Within these contributions, the main one is the one coming from the third order in the Taylor expansion of the potential, which is proportional to $\partial_\phi^3 V$, and becomes large around the inflection point of the potential. From now on, we neglect the rest of the interactions. Therefore, we approximate the interaction Hamiltonian in the interaction picture, in the $\delta\phi$-gauge, as

$$\hat{H}_I(t) = M_P^2 \int \mathrm{d}^3\boldsymbol{x}\, c_0(t) a(t)^3 H(t)^2 \hat{\varphi}(t, \boldsymbol{x})^3\,,\tag{30}$$

where the scalar fluctuation $\hat{\varphi}$ in the interaction picture is:

$$\hat{\varphi}(t, \boldsymbol{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} e^{i\boldsymbol{k}\boldsymbol{x}} \hat{\varphi}_{\boldsymbol{k}}(t)\,,\quad \text{where}\quad \hat{\varphi}_{\boldsymbol{k}}(t) = \varphi_k(t)\hat{a}_{\boldsymbol{k}} + \varphi_k^*(t)\hat{a}_{-\boldsymbol{k}}^\dagger\,.\tag{31}$$

The free modes satisfy the equation of motion

$$\varphi_k'' + (3 + \epsilon - 2\eta)\varphi_k' + \left(\frac{k}{aH}\right)^2 \varphi_k = 0\,,\tag{32}$$

with the initial conditions determined by Bunch-Davies, after canonical normalization.

# 3 Bispectrum of $\zeta$

In this section, we compute the lowest order contribution to the bispectrum of $\zeta$ in an expansion on $H/M_P$ in order to obtain some insight into whether the (perturbative) non-Gaussianity of $\delta\phi$ can yield an important contribution to that of $\zeta$ at values of $\zeta$ relevant for PBH formation. The intuition developed with this calculation will help us to interpret the results in Sec. 4 for the non-Gaussian PDF of $\zeta$ including corrections arising from the non-Gaussianity of $\delta\phi$.

## 3.1 Contributions to the bispectrum of $\zeta$

During the CR stage that follows the USR phase, the curvature perturbation is related to $\varphi$ as

$$\zeta = \varphi - \frac{\eta_{\mathrm{CR}}}{2}\varphi^2 + \frac{\eta_{\mathrm{CR}}^2}{3}\varphi^3 + \cdots.\tag{33}$$

This can be obtained by evaluating (20) for constant $\eta = \eta_{\mathrm{CR}}$, or, equivalently, expanding (22) around $\delta\phi = 0$. We now consider the Fourier modes $\hat{\zeta}_{\boldsymbol{k}}$ of the quantum operator $\hat{\zeta}$, related to $\hat{\varphi}$ by (33). Expanding the three-point correlation function of $\hat{\zeta}$ in powers of $\hat{\varphi}$, the two lowest order terms in $H/M_P$ are

$$\langle \hat{\zeta}_{\boldsymbol{p}} \hat{\zeta}_{\boldsymbol{q}} \hat{\zeta}_{\boldsymbol{k}} \rangle = \langle \hat{\varphi}_{\boldsymbol{p}} \hat{\varphi}_{\boldsymbol{q}} \hat{\varphi}_{\boldsymbol{k}} \rangle - \frac{\eta_{\mathrm{CR}}}{2} \left( \int \frac{\mathrm{d}^3 \ell}{(2\pi)^3} \langle \hat{\varphi}_{\boldsymbol{k}-\boldsymbol{\ell}} \hat{\varphi}_{\boldsymbol{\ell}} \hat{\varphi}_{\boldsymbol{q}} \hat{\varphi}_{\boldsymbol{p}} \rangle + \mathrm{perm.} \right). \tag{34}$$

Introducing the notation

$$\mathcal{I}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k})(2\pi)^3 \delta^{(3)}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{k}) = \langle \hat{\varphi}_{\boldsymbol{p}} \hat{\varphi}_{\boldsymbol{q}} \hat{\varphi}_{\boldsymbol{k}} \rangle_c \,, \tag{35}$$

$$\mathcal{N}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k})(2\pi)^3 \delta^{(3)}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{k}) = -\frac{\eta_{\mathrm{CR}}}{2} \left( \int \frac{\mathrm{d}^3 \ell}{(2\pi)^3} \langle \hat{\varphi}_{\boldsymbol{k}-\boldsymbol{\ell}} \hat{\varphi}_{\boldsymbol{\ell}} \hat{\varphi}_{\boldsymbol{q}} \hat{\varphi}_{\boldsymbol{p}} \rangle_c + \mathrm{perm.} \right), \tag{36}$$

we obtain that the bispectrum of $\zeta$ is

$$\langle \hat{\zeta}_{\boldsymbol{p}} \hat{\zeta}_{\boldsymbol{q}} \hat{\zeta}_{\boldsymbol{k}} \rangle_c = B(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k})(2\pi)^3 \delta^{(3)}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{k}) = \Big[ \mathcal{I}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k}) + \mathcal{N}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k}) \Big] (2\pi)^3 \delta^{(3)}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{k}). \tag{37}$$

If $\delta\phi$ had no interactions, $\mathcal{I}$ would be zero (since it corresponds the correlator of an odd number of fields). Its leading contribution arises from the tree-level diagram featuring a single insertion of the cubic interaction Hamiltonian:

$$\langle \hat{\varphi}_{\boldsymbol{p}}(t) \hat{\varphi}_{\boldsymbol{q}}(t) \hat{\varphi}_{\boldsymbol{k}}(t) \rangle_c = 2 \,\mathrm{Im} \left[ \int_{-\infty_-}^{t} \mathrm{d}t' \langle \hat{\varphi}_{\boldsymbol{p}}(t) \hat{\varphi}_{\boldsymbol{q}}(t) \hat{\varphi}_{\boldsymbol{k}}(t) \hat{H}_I(t') \rangle_c \right]. \tag{38}$$

Plugging in $\hat{H}_I$ from (30) and the field $\hat{\varphi}$ in the interaction picture described in Eq. (31), we have

$$\mathcal{I}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k}) = 3! M_P^2 \int_{-\infty_-}^{t} \mathrm{d}t' a(t')^3 c_0(t') H(t')^2 \, 2\,\mathrm{Im} \Big[ \varphi_p(t) \varphi_p^*(t') \varphi_q(t) \varphi_q^*(t') \varphi_k(t) \varphi_k^*(t') \Big], \tag{39}$$

where we stress that $\varphi_k$ are the free modes satisfying (32). The contribution $\mathcal{N}$ in (36) is non-zero even if $\delta\phi$ is non-interacting, since it is a correlator of an even number of fields. Therefore, its leading contribution is obtained by directly applying Wick's theorem, and corresponds to the *non-linear* contribution to the bispectrum. The result reduces to permutations of expectation values in the free vacuum of pairs of free fields [11],

$$\mathcal{N}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k}) = -\eta_{\mathrm{CR}} \Big( |\varphi_q|^2 |\varphi_k|^2 + |\varphi_p|^2 |\varphi_k|^2 + |\varphi_q|^2 |\varphi_p|^2 \Big). \tag{40}$$

To recap, we have obtained simple expressions for the lowest order contributions, $\mathcal{I}$ and $\mathcal{N}$, to the bispectrum of $\zeta$. On the one hand, $\mathcal{I}$ is given by the tree-level bispectrum of $\varphi$, and it accounts for the cubic self interactions of $\varphi$. In particular, it is zero for a free (i.e. Gaussian) $\varphi$; in other words, it gives an *intrinsic* non-Gaussian contribution to the bispectrum of $\zeta$. On the other hand, $\mathcal{N}$ arises because of considering a *non-linear* (in this truncated calculation, quadratic) relation between $\delta\phi$ and $\zeta$. It is non-zero even if $\varphi$ is a free (i.e. Gaussian) field. At the lowest order in $H/M_P$, the split between intrinsic non-Gaussiniaty and that coming from non-linearities is manifest for the bispectrum. Indeed, the total bispectrum of $\zeta$ is the sum of (39), arising solely from $\delta\phi$ interactions, and (40), due to non-linearity, which contains no information on $\delta\phi$ interactions. At higher orders in $H/M_P$, however, this separation is not so transparent. For example, if one computes loop corrections to (40), one needs to include insertions of the cubic (or higher order) interaction Hamiltonian to compute the correlator in (36), which partly accounts for the intrinsic non-Gaussianity in the bispectrum of $\zeta$.

A priori, it is not necessary to stop at the quadratic order in the non-linear relation between $\delta\phi$ and $\zeta$. The $\delta N$ formalism provides a non-linear relation between them, of which (33) is a truncated expansion (whose first non-linear term gave us $\mathcal{N}$). It is however instructive to use the formulae computed in this section to gain some intuition on the relative contribution to the non-Gaussianity of $\zeta$ of the non-Gaussianity of $\delta\phi$ by simply comparing the size of $\mathcal{I}$ and $\mathcal{N}$. One might naively think that the intrinsic contribution to the bispectrum of $\zeta$, i.e. $\mathcal{I}$, should always be smaller than the one coming from the change of variables, i.e $\mathcal{N}$, given that $\mathcal{I}$ is proportional to $c_0$ (which is suppressed by $\epsilon \ll 1$) and $\mathcal{N}$ comes with a factor $\eta_{\mathrm{CR}}$ (see Eq. (36)), which can be $\mathcal{O}(1)$. However, this is actually not the case. A quick estimate of the relative importance between $\mathcal{I}$ and $\mathcal{N}$ can be obtained from a more careful analysis of the quotient of the slow-roll parameters involved in Eqs. (35) and (36). Each $\varphi$ comes with a factor $\epsilon^{-1/2}$ from Eq. (36). Besides, in Eq. (35) $c_0 \sim \epsilon\, \eta\, \epsilon_3$ (from Eq. (28)), while the prefactor in Eq. (36) is directly $\eta_{\mathrm{CR}}$, as we already mentioned. Thus, we have $\mathcal{I}/\mathcal{N} \sim \epsilon_3 = (\log\epsilon_2)'$, which can be large depending on the sharpness of the transition. Moreover, the integration in time in Eq. (35) (made explicit in Eq. (39)) can alter the naive prediction on the hierarchy between $\mathcal{I}$ and $\mathcal{N}$ obtained by just counting slow-roll parameters. We will do a numerical comparison of the sizes of $\mathcal{I}$ and $\mathcal{N}$ in Section 3.3, discussing the impact of the smoothness of the transitions between SR, USR and CR.

## 3.2   Consistency relation

Before moving on, we indicate a quick consistency check that one can make to detect whether any term neglected in (30) (and therefore in the calculation of $\mathcal{I}$) was actually non-negligible. We use the consistency relation for the bispectrum in the squeezed limit [39, 41]. If we take $k \ll p \simeq q$ in Eq. (39) and use the fact that $\varphi_k$ freezes in the super-Hubble limit, we find[10]

$$\frac{\mathcal{I}(k,p)}{|\varphi_k(t)|^2|\varphi_p(t)|^2} = 12M_P^2\,\mathrm{Im}\left[\frac{\varphi_p(t)^2}{|\varphi_p(t)|^2}\int_{-\infty_-}^{t} \mathrm{d}t'\, a(t')^3 c_0(t') H(t')^2 \varphi_p^*(t')^2\right]. \tag{41}$$

Similarly,

$$\frac{\mathcal{N}(k,p)}{|\varphi_k(t)|^2|\varphi_p(t)|^2} = -\eta_{\mathrm{CR}}\left(2 + \frac{k^3}{p^3}\frac{\mathcal{P}_\zeta(p)}{\mathcal{P}_\zeta(k)}\right) \simeq -2\eta_{\mathrm{CR}}, \tag{42}$$

where $\mathcal{P}_\zeta(k) = (k^3/2\pi^2)|\varphi_k|^2$ is the tree level power spectrum of $\zeta$. The following relation must be obeyed by virtue of the consistency relation:

$$\frac{\mathcal{I}(k,p) + \mathcal{N}(k,p)}{|\varphi_k(t_e)|^2|\varphi_p(t_e)|^2} = -\frac{\mathrm{d}\log\mathcal{P}_\zeta(p)}{\mathrm{d}\log p}, \tag{43}$$

where $t_e$ denotes the end of inflation. If, while using our simplified $\hat{H}_I$ to compute $\mathcal{I}$, Eq. (43) did not hold, it would mean that some non-negligible term from (27) was not accounted for in (30). We do this check in Figure 1 for the model described in the next subsection. As shown in the figure, the equality of both sides of (43) holds. Although this check is encouraging, we note that some interactions become negligible precisely in the squeezed limit (see Appendix A) and this check is not sensitive to these contributions. See [42] and [43] for previous checks of the consistency relation in related settings. It is

---

[10]One might worry that, since the time integral starts at $t' \to -\infty_-$, there is some integration range in which $\varphi_k(t')$ is not constant, no matter how early it exited the horizon. However, the fact that $\delta\phi$ self interactions are localized around the USR phase (in other words, that $c_0$ quickly decays away from the USR phase) make it safe to restrict the integration domain to a neighbourhood on said phase, which makes it a good approximation to take $\varphi_k(t')$ as constant as long as it exited the horizon sufficiently before USR.
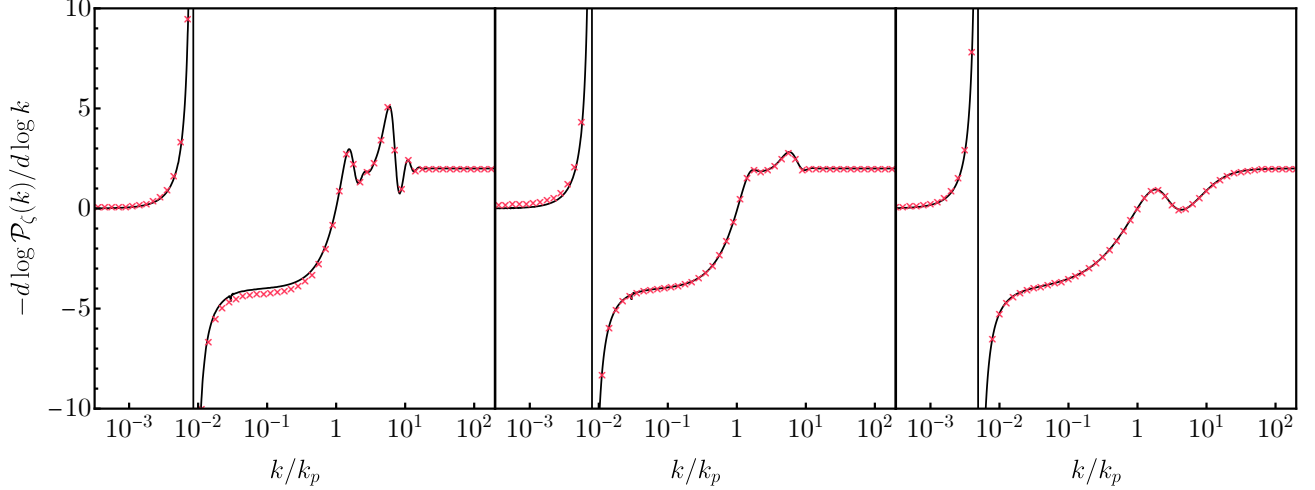
Figure 1: *Consistency relation for the sharp (left), intermediate (center) and smooth (right) cases in Table 1, obtained by comparing the left-hand side of Eq. (43) (red crosses) and the right-hand side involving the derivative of the power spectrum (solid black line). The agreement implies that we have considered all interactions that are relevant (at least) in the squeezed limit.*

also worth recalling that earlier works pointed out that the consistency relation does not hold if the power spectrum and the bispectrum are evaluated at a time at which $\zeta$ has not frozen in the super-Hubble limit, see for instance [44, 45] for the specific case of USR inflation. This can be expected, given that the most straightforward derivation of the consistency relation relies explicitly on the conservation of $\zeta$, see [39]. A generalization of the consistency relation for the case in which the modes of $\zeta$ do not freeze in the super-Hubble limit was proposed in [46]. Later works explored the impact on the bispectrum (and on the consistency relation) of a transition from USR to SR, see in particular [47, 48] and also reference [11], which we have already mentioned. As we have explained, we use the consistency relation as a tool aiming to test that we are including the necessary interactions in our computation of the bispectrum and therefore we check it at late times, when it has to apply in the type of model we consider.

### 3.3 Numerical model

In this section we study the dependence of $\mathcal{I}$ and $\mathcal{N}$ on the sharpness of the SR-to-USR and the USR-to-CR transitions. To do this, we consider a simplification of a parametrization introduced in [11], where a specific time-dependence (in terms of the number of e-folds, $N$) is imposed on the slow-roll parameter $\eta$:

$$\eta = \frac{1}{2}\left[ -\Delta\eta + (3 + \Delta\eta)\tanh\left(\frac{N}{\delta}\right) - (3 + 2\Delta\eta)\tanh\left(\frac{N - \Delta N}{\delta}\right) \right], \tag{44}$$

where we set $N = 0$ at the beginning of the USR phase. Here $\delta$ is a positive number that controls the smoothness of the transitions in and out of USR (the smaller $\delta$, the sharper the transitions). The parameter $\Delta N$ controls, together with $\delta$, the duration of USR. Indeed this is approximately equal to $\Delta N - \delta$. For this reason, as we increase $\delta$, we also need to increase $\Delta N$ to keep the duration of USR (and therefore the heigh of the peak of the spectrum) more or less constant. The number $\Delta\eta$ characterizes the
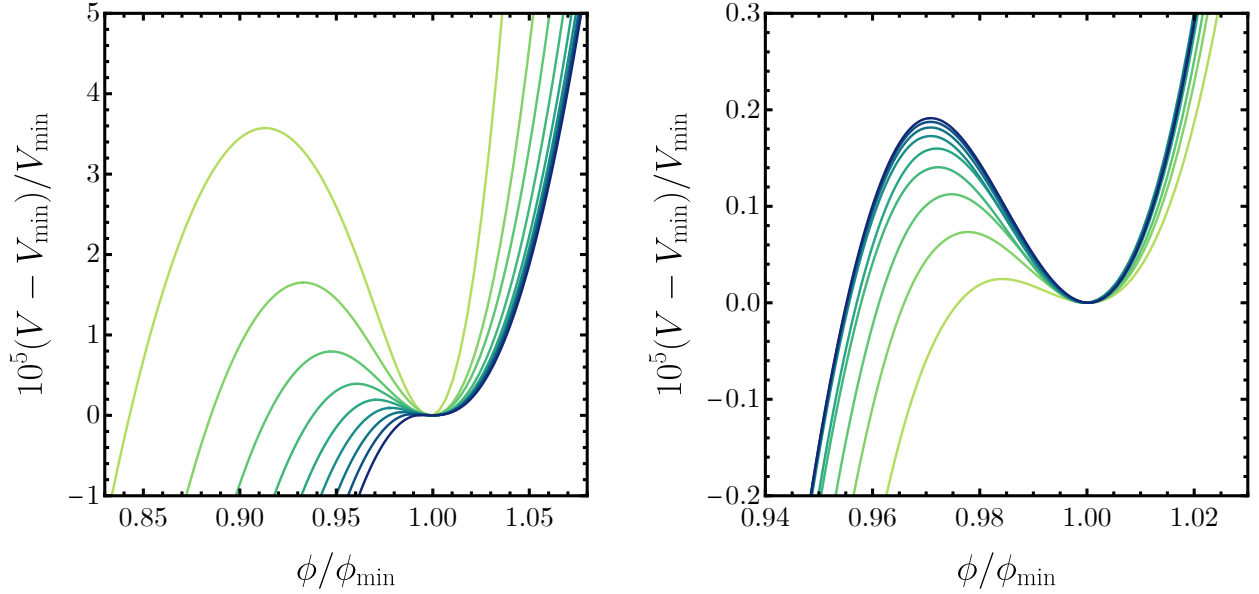
Figure 2: *Left panel: Potentials obtained by varying $\delta$ between $0.1$ (green) and $0.9$ (blue), for $\Delta N = 2$. Right panel: Potentials obtained by varying $\Delta N$ between $1.1$ (green) and $2$ (blue), for $\delta = 0.5$. We have set $\eta_{\mathrm{USR}} = 4$ and $\eta_{\mathrm{CR}} = -1$ for both panels.*

approximate values of $\eta$ in the USR and CR phases: $\eta_{\mathrm{USR}} = 3 + \Delta\eta$ and $\eta_{\mathrm{CR}} = -\Delta\eta$, respectively. This ensures that Wands duality [49] is satisfied[11] unless if $\delta$ is large enough (in which case $\eta$ never reaches $\eta_{\mathrm{USR}}$). The shape of $\eta$ from (44) mimics (with smooth transitions) a piecewise constant function, with $\eta_{\mathrm{USR}}$ the value of $\eta$ in the USR phase and $\eta_{\mathrm{CR}}$ the one in the subsequent CR phase. Integrating $\eta$ in time, we recover $\epsilon$; the corresponding boundary condition is fixed by choosing a reasonably small value of $\epsilon$ in the SR phase. In particular, we choose $\epsilon = 10^{-3}$ at very early times. Integrating $\epsilon$, we recover $H$; in this case, for which the boundary condition is fixed by the amplitude of perturbations on very large scales, $\mathcal{P}_\zeta \simeq (H/M_P)^2/(8\pi^2\epsilon) \approx 2.2 \times 10^{-9}$. The kind of potentials obtained by varying the parameters in this model are shown in Figure 2.

We choose three representative sets of parameters to illustrate our results (see Table 1). The three examples are selected such that the power spectrum reaches $\mathcal{P}_\zeta(k_p) \simeq 10^{-2}$ at the peak,[12] regardless of the sharpness of the transition. The consistency relation (43) is illustrated for the three cases under consideration in Figure 1, obtaining a good agreement. As we mentioned earlier, the fact that the consistency relation is satisfied is an indication that we are not neglecting relevant interactions when computing $\mathcal{I}$ including only the $c_0\,\varphi^3$ interaction.

We now move to the calculation of $\mathcal{I}$ and $\mathcal{N}$ using (39) and (40). As explained above, integrating (44) successively, we obtain $\epsilon$ and $H$. And differentiating (44), we obtain $\epsilon_3$ and $\epsilon_4$, which allows to compute $c_0$, see Eq. (28). Also, having the slow-roll parameters, we can compute the time evolution of the Fourier modes $\varphi_k$ in the free theory through Eq. (32).

---

[11]The validity of (22) relies on $\dot\eta$ being negligible after the inflaton overcomes the local maximum of the potential (which is the definition of CR). This is ensured by Wands duality holding or, equivalently, the potential being accurately approximated by a parabola around the local maximum (see [20] for a recent discussion).

[12]This peak occurs for scales crossing the horizon at the USR-to-CR transition, $k_p \simeq a(\Delta N)H(\Delta N)$, as illustrated in Fig. 4 on the right panel.

13

|  | smooth | intermediate | sharp |
|---|---|---|---|
| $\delta$ | 0.90 | 0.33 | 0.22 |
| $\Delta N$ | 2.19 | 1.72 | 1.63 |

Table 1: *Parameters for the three illustrative models considered in this work. In all three cases, $\eta_{\mathrm{USR}} = 4$ and $\eta_{\mathrm{CR}} = -1$ (in other words, $\Delta\eta = 1$).*
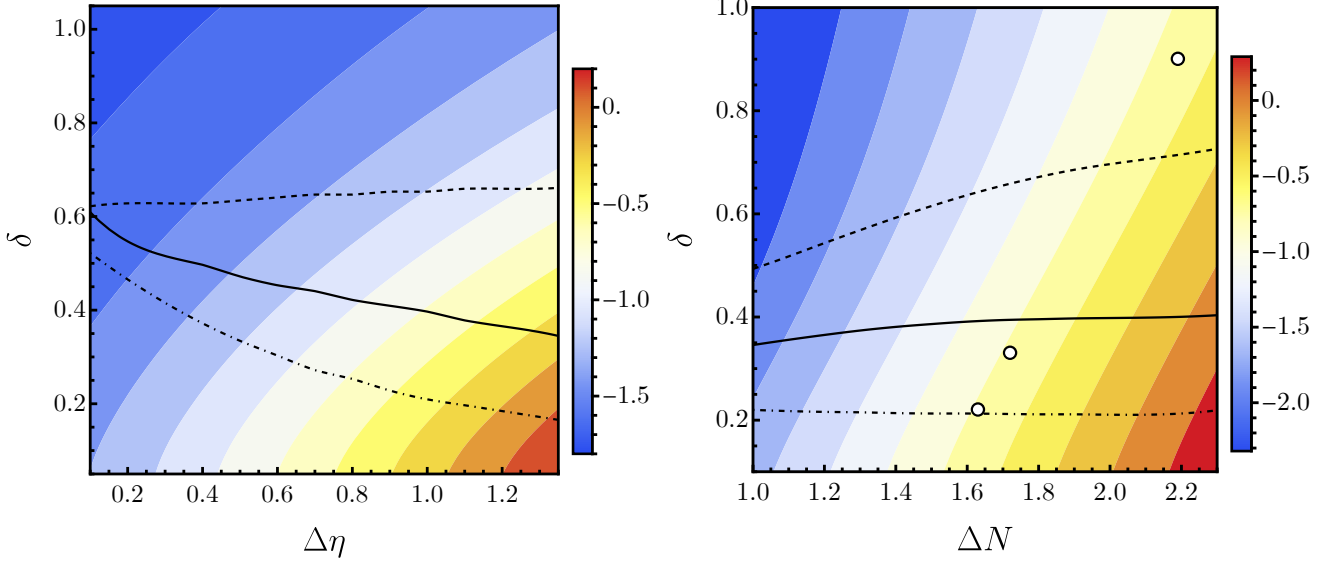


Figure 3: *Colored contours correspond to the peak value of $\log_{10}\mathcal{P}_\zeta$. Dashed, solid and dot-dashed black lines are the contours $\mathcal{I}/\mathcal{N} = -0.5, 0, 0.5$, respectively. Left panel shows the dependence of these two quantities on $\Delta\eta$ and $\delta$, fixing $\Delta N = 1.7$. Right panel shows their dependence on $\Delta N$ and $\delta$, fixing $\Delta\eta = 1$. The three white spots correspond to the three cases described in Table 1, which indeed have a similar peak height in their power spectrum (Figure 4, right panel). The assumption of Gaussian $\delta\phi$ is only valid sufficiently close to the solid black line.*

In order to compare the sizes of $\mathcal{I}$ and $\mathcal{N}$, we have to evaluate them at a certain momentum configuration. For illustrative purposes, we will consider the equilateral configuration of the bispectrum, where all the momenta coincide. We find that in this configuration the contributions $\mathcal{N}$ and $\mathcal{I}$ to the bispectrum are maximal on the scale associated with the peak of the power spectrum, see Figure 4. Therefore, a first assessment of the relative importance of the non-Gaussianity of $\delta\phi$ and the non-linear relation to $\zeta$ can be obtained with the ratio $\mathcal{I}(k_p)/\mathcal{N}(k_p)$. If $|\mathcal{I}(k_p)/\mathcal{N}(k_p)| \ll 1$, the non-Gaussianity of $\zeta$ is dominated by its non-linear relation to $\delta\phi$, making the non-Gaussianity of $\delta\phi$ unimportant. Therefore, in this case, one could expect Eq. (23), which assumes a Gaussian $\delta\phi$, to be a good approximation for the PDF of $\zeta$. However, if $\mathcal{I}(k_p)/\mathcal{N}(k_p)$ is not small, $\mathcal{I}$ cannot be neglected and one may expect a non-linear transformation (such as Eq. (22)) of a Gaussian $\delta\phi$ to lead to a poor estimate of the PDF of $\zeta$.

The results obtained from this comparison are displayed in Figure 3. We note that, in general, the contribution of $\mathcal{I}$ to the bispectrum is not negligible with respect to that of $\mathcal{N}$. However, due to a change of sign of $\mathcal{I}$ around $\delta \simeq 0.4$, there is a region in the parameter space where the contribution of $\mathcal{I}$ to the bispectrum is highly suppressed (see Figure 3). One can roughly estimate the value of $\delta$ for which $\mathcal{I}$ vanishes. First, one can notice that the factor $\mathrm{Im}\big[(\varphi_{k_p}(N)\varphi_{k_p}^*(N_e))^3\big]$ in Eq. (39) peaks around

$\Delta N$ and decreases for $N > \Delta N$ as $\varphi_{k_p}$ freezes to its asymptotic value. In practice, this term acts as a window function that restricts the integral in (39) to a narrow range of e-folds around the USR-to-CR transition, $N \simeq \Delta N$. Notice also that $c_0(N)$ crosses 0 during this transition. Thus, for $N$ around $\Delta N$, one can minimize $\mathcal{I}$ by requesting $c_0'(\Delta N) = 0$. We find numerically that the most relevant contribution to $c_0'(\Delta N)$ around $N \simeq \Delta N$ is

$$\frac{\mathrm{d}c_0}{\mathrm{d}N} \simeq \frac{1}{3}\,\epsilon\,\eta\,\epsilon_3\left(\epsilon + 2\,\eta - \epsilon_3 - \epsilon_4 - 3\right)', \tag{45}$$

which vanishes if the derivative of the term in brackets vanishes. In the limit $\delta \ll \Delta N$, this approximately corresponds to

$$c_0'(N) \simeq (2\Delta\eta + 3)^2[\delta(2\Delta\eta + 3) - 2] = 0 \implies \delta = \frac{2}{3 + 2\Delta\eta} = 0.4 \quad \text{for} \quad \Delta\eta = 1\,, \tag{46}$$

recovering the result represented in Figure 3. Notice that, as long as $\delta \ll \Delta N$, this result is essentially independent of $\Delta N$.

In [11], it was argued that the non-Gaussianity of $\delta\phi$ is essentially negligible if the USR-to-CR transition is sufficiently smooth, and the bispectrum of $\zeta$ is of the local form. It was found there that, as $\delta \to 0.44$, the contribution to the bispectrum of $\zeta$ due to the bispectrum of $\delta\phi$ vanishes, and the one coming from the field redefinition dominates. The conclusion they inferred was that, the smoother the transition, the more subdominant the intrinsic non-Gaussianity will become. However, we have seen that $\mathcal{I}$ vanishes for $\delta \approx 0.4$ because it crosses zero, not because it vanishes asymptotically. Indeed, if we keep increasing the value of $\delta$, $\mathcal{I}$ (the intrinsic non-Gaussianity) becomes sizable again, only with a different sign.[13]

In the next section, we will provide a more refined assessment of the relative contribution of $\mathcal{I}$ and $\mathcal{N}$ to the non-Gaussianity of $\zeta$.

## 4  Lattice simulation

In the last section it has been shown that the effect of the intrinsic non-Gaussianities of $\zeta$ –understood to be those coming from the interactions of $\delta\phi$– on its bispectrum can be at least as important as the contribution associated with the non-linearities between $\zeta$ and $\delta\phi$, depending on the parameters of the model. Therefore, a correct description of the PDF of $\zeta$ (and particularly its tail) may require including both sources of non-Gaussianities. However, there is no straightforward way to incorporate non-Gaussian corrections into the PDF starting from a Gaussian distribution.[14]

In this section we present a numerical method to generate a non-Gaussian field satisfying specific input perturbative statistics that we can choose at will, starting from a stochastic variable following a Gaussian distribution. In particular, we will construct a stochastic variable $\tilde{\zeta}$ –where from now on ˜ will be carried by stochastic variables– in such a way that its statistical moments coincide with the $n$-point correlation functions of the variable $\hat{\zeta}$ obtained from the in-in formalism, i.e.

$$\langle\tilde{\zeta}_{\boldsymbol{k}_1}\tilde{\zeta}_{\boldsymbol{k}_2}\cdots\tilde{\zeta}_{\boldsymbol{k}_n}\rangle_c = \langle\hat{\zeta}_{\boldsymbol{k}_1}\hat{\zeta}_{\boldsymbol{k}_2}\cdots\hat{\zeta}_{\boldsymbol{k}_n}\rangle_c\,. \tag{47}$$

---

[13]The analysis in [11] was done in the $\zeta$-gauge, using Maldacena's field redefinition $\mathcal{R} \to \mathcal{R}_n + f(\mathcal{R}_n)$ (see Eqs. (A4-A6) therein). Their $\mathcal{R}_n$ corresponds with our $\varphi$ in the $\delta\phi$-gauge.

[14]This is what the Gram-Charlier (or Edgeworth) series tries to do, but it does not converge in the present case. This series can even lead to non-positive results for the PDF upon truncation at certain orders, which clearly limits its usefulness to expand the PDF (see Appendix B for more details).
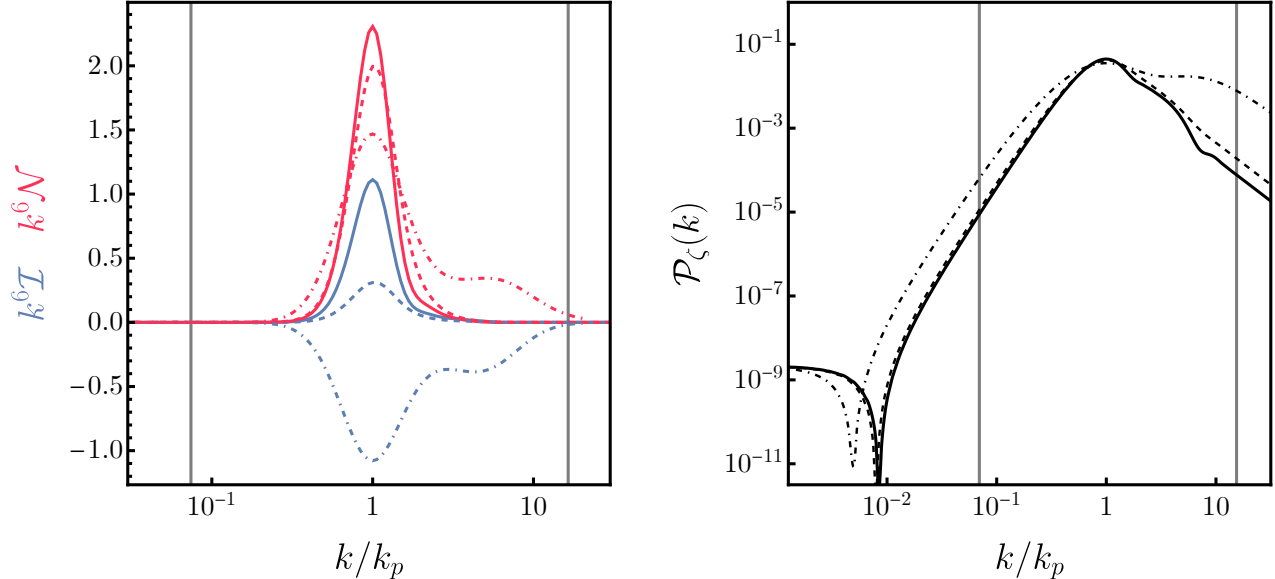
Figure 4: *Left panel: Contributions to the three-point function of $\zeta$ arising from the intrinsic non-Gaussianity of $\delta\phi$ (blue) and its nonlinear relation to $\zeta$ (red) rescaled by the factor $k^6$ for the sharp (solid), intermediate (dashed) and smooth (dot-dashed) examples in Section 3.3. The vertical lines denote the approximate range of momenta probed by the lattice simulation presented in Section 4. Right panel: Power spectra for the same examples as in the left panel. We denote the location of the peak in the spectra by $k_p$.*

The brackets $\langle\cdots\rangle$ refer to ensemble averages or correlation functions when applied to stochastic variables or quantum operators, respectively. We recall that the subscript $_c$ stands for the connected part of the averages and correlations, and it always comes accompanied by $\delta(\boldsymbol{k}_1 + \cdots + \boldsymbol{k}_n)$ that guarantees the momentum conservation. Imposing that the connected parts of the $n$-point ensemble averages and quantum correlation functions coincide, it is guaranteed that their disconnected parts also coincide, because the disconnected parts are obtained by the product of connected parts of lower orders [37].

Once we have ensured that $\tilde{\zeta}$ has the correct statistic, we can sample $\tilde{\zeta}$ to obtain its non-Gaussian probability distribution, thus completing the calculation of the PDF.

Since the relation between the curvature fluctuation $\zeta$ and the inflaton fluctuation $\delta\phi$ in the $\delta\phi$-gauge is exactly described by Eq. (22), we can describe $\tilde{\zeta}$ implicitly by means of a variable $\tilde{\varphi}$ that satisfies the correct statistics,

$$\langle\tilde{\varphi}_{\boldsymbol{k}_1}\tilde{\varphi}_{\boldsymbol{k}_2}\cdots\tilde{\varphi}_{\boldsymbol{k}_n}\rangle_c = \langle\hat{\varphi}_{\boldsymbol{k}_1}\hat{\varphi}_{\boldsymbol{k}_2}\cdots\hat{\varphi}_{\boldsymbol{k}_n}\rangle_c, \tag{48}$$

as discussed above. We stress that the choice of $\varphi$, rather than $\delta\phi$, is simply for convenience; but it is worth noting that $\varphi$ still describes the scalar fluctuations in the $\delta\phi$-gauge, Eq. (26).

Let us emphasize that our procedure is not equivalent to performing a dynamical lattice simulation of the time evolution of the scalar fluctuations during inflation, as in [9]. Instead, we calculate the statistical properties of the scalar fluctuation $\zeta$ using the in-in formalism, and simulate its probability distribution *at the end of inflation*, where the modes of $\zeta$ with sufficiently long wavelength will be frozen at all orders in perturbations [16, 38, 39] (see also [50–53] for a discussion of this statement at the quantum level).[15]

---

[15]Although at the end of inflation there are still modes inside the horizon, and therefore not all the modes are frozen,

This allows us to incorporate the quantum dynamical effects arising from the in-in formalism. The price to pay is that, following our procedure, the probability distribution can only be found perturbatively, because the correlation functions are obtained by means of a loop (or Planck mass) expansion using the in-in formalism.

Although our procedure to obtain the PDF of $\tilde{\varphi}$ can be applied at any time during inflation, the change of gauge we use subsequently to get the PDF of $\tilde{\zeta}$ works only when the coarse-grained variable $\zeta$ is frozen, i.e. when all the Fourier modes of interest have become constant. In practice, for the kind of models we are interested in, this is equivalent to applying the change of gauge at the end of inflation, provided that the enhancement of $\zeta$ occurs for modes that become super-Hubble early enough.

## 4.1 Non-Gaussian field generation

To construct the stochastic variable $\tilde{\varphi}$, we will make use of an auxiliary Gaussian stochastic variable of unit variance $\tilde{\psi}_{\boldsymbol{k}}$, whose statistics is fully determined by the two-point correlation function:

$$\langle \tilde{\psi}_{\boldsymbol{k}} \tilde{\psi}_{\boldsymbol{p}} \rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{k} + \boldsymbol{p}) . \tag{49}$$

Since $\tilde{\psi}_{\boldsymbol{k}}$ is Gaussian, any $n$-point correlator can be obtained by using Wick's theorem. We decompose $\tilde{\varphi}$ as follows:

$$\tilde{\varphi}(\boldsymbol{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} e^{i\boldsymbol{k}\boldsymbol{x}} \tilde{\varphi}_{\boldsymbol{k}} , \quad \text{with} \quad \tilde{\varphi}_{\boldsymbol{k}} = \sum_n \tilde{\varphi}_{\boldsymbol{k}}^{(n)} , \tag{50}$$

where

$$\tilde{\varphi}_{\boldsymbol{k}}^{(n)} \equiv \frac{1}{n!} \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \cdots \frac{\mathrm{d}^3 k_{n-1}}{(2\pi)^3} \mathcal{G}_n(\boldsymbol{k}, -\boldsymbol{k}_1, \cdots, -\boldsymbol{k}_{n-1}) \frac{\tilde{\psi}_{\boldsymbol{k}_1}}{|\varphi_{k_1}|} \cdots \frac{\tilde{\psi}_{\boldsymbol{k}_{n-1}}}{|\varphi_{k_{n-1}}|} (2\pi)^3 \delta^{(3)}(-\boldsymbol{k} + \boldsymbol{k}_1 + \cdots + \boldsymbol{k}_{n-1}) . \tag{51}$$

We emphasize that we are describing the variable $\tilde{\varphi}$ at the end of inflation, $t_e$, so we omit the time dependence of all the objects that compose it assuming that they are evaluated at that time. Due to the freedom we have in choosing the functions $\mathcal{G}_n$, as long as they are totally symmetric under momenta permutations, we will be able to impose that $\tilde{\varphi}$ satisfies the appropriate statistics, as we previously discussed. The functions $\varphi_{k_i}$ are the modes, solution to the free equation of motion, which we find in the operator $\hat{\varphi}$ in Eq. (31). They are evaluated at the end of inflation and therefore we omit their time dependence. For convenience, we also define the connected part of the $n$-point correlation of $\hat{\varphi}$ at the end of inflation as:

$$\langle \hat{\varphi}_{\boldsymbol{k}_1} \cdots \hat{\varphi}_{\boldsymbol{k}_n} \rangle_c \equiv \mathcal{I}_n(\boldsymbol{k}_1, \cdots, \boldsymbol{k}_n)(2\pi)^3 \delta^3(\boldsymbol{k}_1 + \cdots + \boldsymbol{k}_n) . \tag{52}$$

We can thus identify $\mathcal{I}_3 = \mathcal{I}$ as the intrinsic contribution to the bispectrum in Eq. (39).

We will assume perturbation theory to be valid. This means that the correlators $\mathcal{I}_n$, Eq. (52), obtained through the in-in formalism will be tree-level dominated; with loop level corrections, generated by further insertions of the interaction Hamiltonian, see Section 2.2, being negligible. We will discuss the validity of this assumption for our examples later on.

Naively, the parameter controlling the validity of perturbation theory should be the ratio $H/M_P$. Thus, we can estimate the order of magnitude of $\mathcal{I}_n$ just by counting Planck masses. At $n$-th order, the interaction vertex in the $\delta\phi$-gauge is $\mathcal{L}_n \sim M_P^{2-n}$. Therefore, $\mathcal{I}_n \sim (H/M_P)^{2n-2}$, where the additional

---

since we are interested in the physics around the USR phase, it is legitimate to assume that in those times all the modes of interest will be frozen.

17

$M_P^{-n}$ factor comes from each of the $n$-external fields $\varphi$ coupled to the interaction vertex generated by $\mathcal{L}_n$, and where we have introduced $H$ because it is the only physical scale that can accompany $M_P$. As we will see below, $\mathcal{G}_n \sim I_n$ and finally $\tilde{\varphi}_{\boldsymbol{k}}^{(n)} \sim \mathcal{G}_n |\varphi_k|^{-(n-1)} \sim (H/M_P)^{n-1}$. Therefore, the expansion through which we define $\tilde{\varphi}$, Eq. (50), corresponds to a perturbative expansion in powers of $H/M_P$.

Let us now deduce the functions $\mathcal{G}_n$ imposing Eq. (48) order by order. Although the first correlation to be analyzed is the one-point function, associated to the zero mode of $\tilde{\varphi}$ and therefore indistinguishable from the background, we will impose in each simulation that it is absorbed by the background.[16]

Analyzing the connected two-point correlation at leading order, we obtain:

$$\langle \tilde{\varphi}_{\boldsymbol{k}} \tilde{\varphi}_{\boldsymbol{p}} \rangle_c = \langle \tilde{\varphi}_{\boldsymbol{k}}^{(2)} \tilde{\varphi}_{\boldsymbol{p}}^{(2)} \rangle_c = \frac{1}{4} \frac{\mathcal{G}_2(\boldsymbol{k}, \boldsymbol{p})^2}{|\varphi_k||\varphi_p|} (2\pi)^3 \delta^{(3)}(\boldsymbol{k} + \boldsymbol{p}) \,. \tag{53}$$

This connected average also receives corrections from $\tilde{\varphi}_{\boldsymbol{k}}^{(n)}$ with $n > 2$, but they will be suppressed by powers of $H/M_P$. Imposing the condition of Eq. (48), and taking into account that $\mathcal{I}_2(\boldsymbol{k}, \boldsymbol{p}) = |\varphi_k||\varphi_p|$, we get

$$\mathcal{G}_2(\boldsymbol{k}, \boldsymbol{p}) = 2|\varphi_k||\varphi_p| \,, \quad \text{and thus} \quad \tilde{\varphi}_{\boldsymbol{k}}^{(2)} = |\varphi_k| \, \tilde{\psi}_{\boldsymbol{k}} \,. \tag{54}$$

Repeating the same procedure with the three-point correlation gives

$$\mathcal{G}_3(\boldsymbol{k}, \boldsymbol{p}, \boldsymbol{q}) = \mathcal{I}_3(\boldsymbol{k}, \boldsymbol{p}, \boldsymbol{q}) \,, \tag{55}$$

while for the four-point correlation

$$\mathcal{G}_4(\boldsymbol{k}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\ell}) = \mathcal{I}_4(\boldsymbol{k}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\ell}) - \frac{4}{9} \left[ \frac{\mathcal{I}_3(\boldsymbol{k}, -\boldsymbol{p} - \boldsymbol{k}, \boldsymbol{p}) \mathcal{I}_3(\boldsymbol{q}, -\boldsymbol{\ell} - \boldsymbol{q}, \boldsymbol{\ell})}{|\varphi_{|\boldsymbol{p}+\boldsymbol{k}|}|^2} + (\boldsymbol{p} \to \boldsymbol{q}, \boldsymbol{\ell}) \right] \,. \tag{56}$$

In this last case, $\mathcal{G}_4$ receives a contribution not only from $\mathcal{I}_4$ but also from two insertions of $\mathcal{I}_3$. This is to be expected, since $\mathcal{I}_4$ at tree-level is composed of a contact diagram and diagrams in which a virtual scalar field propagates in an $s$, $t$ and $u$ channel.

The procedure can be extended in a straightforward manner to arbitrary orders, increasing the precision with which we describe the statistics of the variable $\tilde{\varphi}$. However, to illustrate how the corrections to the Gaussian case (described exclusively by $\mathcal{G}_2$) modify the PDF of $\tilde{\varphi}$ it is sufficient to just include $\mathcal{G}_3$, as we will explore in the next section. In general, the effect on the PDF of including increasingly higher-order corrections $\mathcal{G}_n$ will be a modification in the tail of the distribution, increasingly further away from the origin. Therefore, truncating the expansion up to a certain order will allow us to accurately describe the PDF up to a certain value of the variable $\tilde{\varphi}$, provided that perturbation theory works.

We stress that, as explained above, by imposing (48), the disconnected contributions to the $n$-point functions also coincide between averages and correlations,

$$\langle \tilde{\varphi}_{\boldsymbol{k}_1} \tilde{\varphi}_{\boldsymbol{k}_2} \cdots \tilde{\varphi}_{\boldsymbol{k}_n} \rangle = \langle \hat{\varphi}_{\boldsymbol{k}_1} \hat{\varphi}_{\boldsymbol{k}_2} \cdots \hat{\varphi}_{\boldsymbol{k}_n} \rangle \,, \tag{57}$$

which can be explicitly verified using the definition of $\tilde{\varphi}$ in Eq. (50) and the functions $\mathcal{G}_n$ described above.

---

[16]The one-point correlation is a special case, as its leading order is the tadpole, since there is no tree-level contribution. This means that, in practice, this contribution corresponds to a negligible shift on the background, which further justifies that we omit its precise treatment.
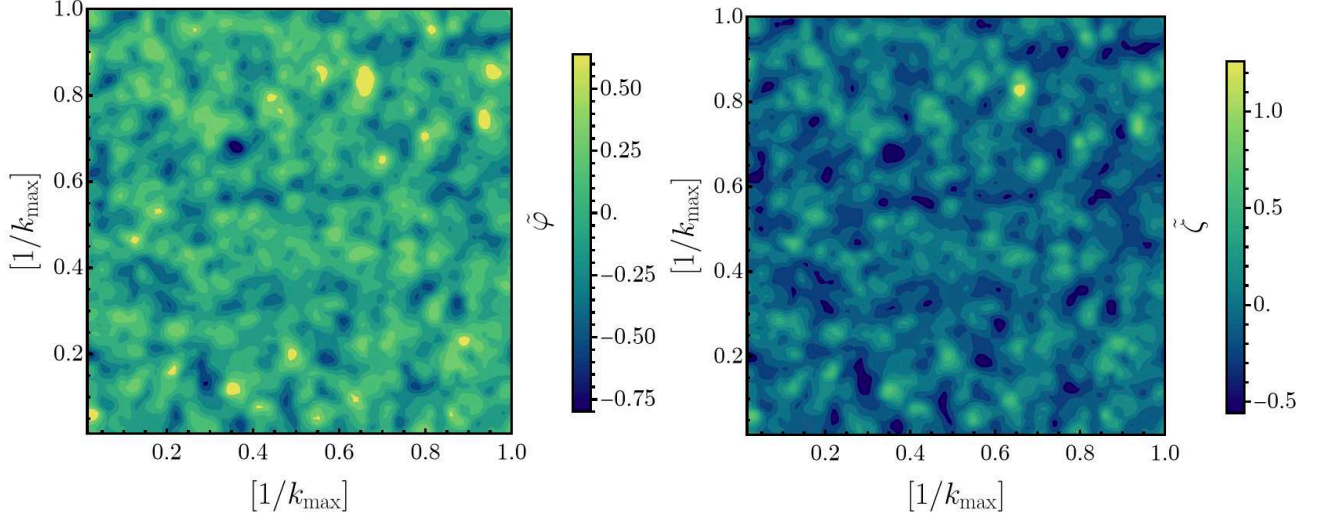
Figure 5: *Left panel: Two-dimensional slice of one realization of the intermediate transition example in Section 3.3 for the Gaussian contribution to $\tilde{\varphi}$ (i.e. $\tilde{\varphi}^{(2)}$). Right panel: Same realization as in the left panel for the full variable $\tilde{\zeta}$. Due to the strong non-Gaussian tail, hot spots are pushed to larger values.*

## 4.2 Validity of perturbation theory

Let us analyze in more detail the validity of the perturbative expansion through which we describe $\tilde{\varphi}$, Eq. (50). Although, as we have seen, the naive parameter controlling the expansion is $H/M_P$, the details of the interactions may mean that even though $H/M_P \ll 1$, in fact perturbation theory breaks. To guarantee the validity of perturbation theory, it has to be ensured that the different averages are dominated by the leading order. For example, in the two-point case,

$$\langle \tilde{\varphi}_{\boldsymbol{k}} \tilde{\varphi}_{\boldsymbol{p}} \rangle_c = \langle \tilde{\varphi}_{\boldsymbol{k}}^{(2)} \tilde{\varphi}_{\boldsymbol{p}}^{(2)} \rangle_c + \langle \tilde{\varphi}_{\boldsymbol{k}}^{(3)} \tilde{\varphi}_{\boldsymbol{p}}^{(3)} \rangle_c + \langle \tilde{\varphi}_{\boldsymbol{k}}^{(2)} \tilde{\varphi}_{\boldsymbol{p}}^{(4)} \rangle_c + \cdots , \tag{58}$$

it has to be satisfied that the latter two contributions are much smaller than the first one. Diagrammatically, these last two contributions are identified with the calculation of the two-point correlation at one loop. Therefore, we can state that our perturbative expansion will be valid, in general, when the loop-level corrections to the $n$-point correlation functions are suppressed with respect to the tree-level contribution. Including loop-level corrections would be possible by modifying $\mathcal{G}_n$, but it is beyond the scope of this paper.

## 4.3 Numerical results

In this section, for concreteness and to illustrate the method that we just described, we analyze the effect of the bispectrum of $\delta\phi$ on the non-Gaussian tail of the PDF of $\zeta$. Using the relations obtained in the previous section, we have

$$\tilde{\varphi}_{\boldsymbol{k}}^{(3)} = \frac{1}{3!} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \mathcal{I}(\boldsymbol{k}, \boldsymbol{k} - \boldsymbol{p}, -\boldsymbol{p}) \frac{\tilde{\psi}_{\boldsymbol{k}-\boldsymbol{p}}}{|\varphi_{|\boldsymbol{k}-\boldsymbol{p}|}|} \frac{\tilde{\psi}_{\boldsymbol{p}}}{|\varphi_p|} . \tag{59}$$

Let us rewrite this expression in a way that is numerically more tractable. We define the function

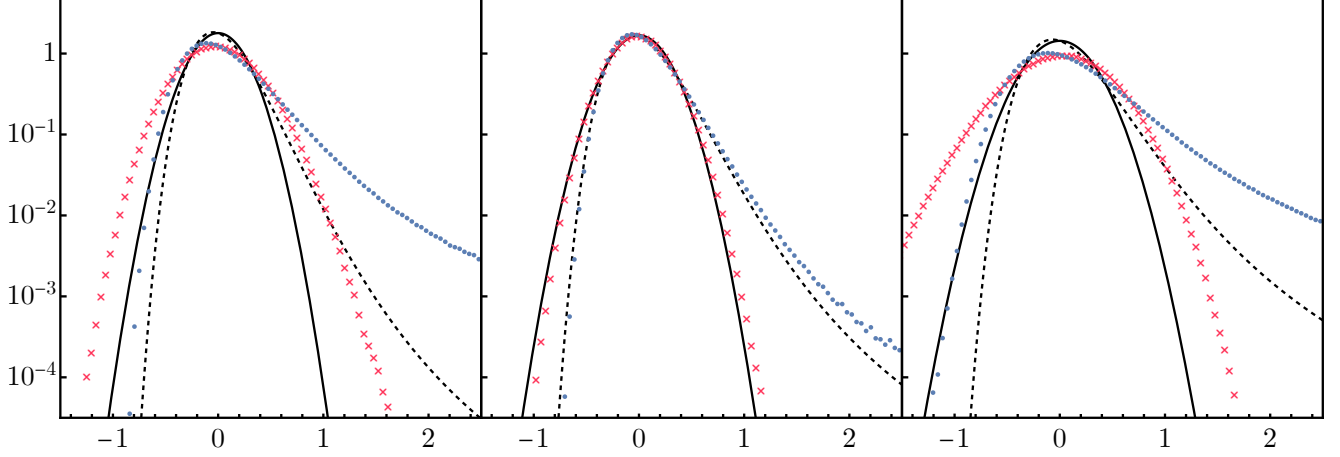$$\lambda_k(t) \equiv \varphi_k(t_e) \varphi_k^*(t) \tag{60}$$

19

Figure 6: *PDFs for the curvature perturbation $\tilde{\zeta}$ and the variable $\tilde{\varphi}$ for the three examples in Section 3.3. The panels correspond, from left to right, to a sharp, intermediate, and smooth transition in and out of USR. The solid and dashed black lines denote the Gaussian contribution to $\tilde{\varphi}$ (i.e. $\tilde{\varphi}^{(2)}$) and the corresponding distribution for $\tilde{\zeta}$, respectively. The red crosses denote the reconstructed non-Gaussian PDF for $\tilde{\varphi}$ (i.e. $\tilde{\varphi}^{(2)} + \tilde{\varphi}^{(3)}$), and the blue dots the corresponding PDF for $\tilde{\zeta}$ after using Eq. (22).*

in terms of the free modes $\varphi_k(t)$, see Eq. (31). Using Eq. (39), we find, at the end of inflation:

$$\tilde{\varphi}_{\boldsymbol{k}}^{(3)} = -iM_P^2 \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \int_{-\infty}^{t_e} \mathrm{d}t \, a(t)^3 c_0(t) H(t)^2 \Big[\lambda_p(t)\lambda_{|\boldsymbol{k}-\boldsymbol{p}|}(t)\lambda_k(t) - \text{h.c.}\Big] \frac{\tilde{\psi}_{\boldsymbol{k}-\boldsymbol{p}}}{|\varphi_{|\boldsymbol{k}-\boldsymbol{p}|}|} \frac{\tilde{\psi}_{\boldsymbol{p}}}{|\varphi_p|}, \qquad (61)$$

where, again, the modes $\varphi_k$, without an explicit time dependence, are evaluated at the end of inflation, $t_e$. Defining

$$\tilde{\Psi}_{\boldsymbol{k}}^+(t) \equiv \frac{\lambda_k(t)}{|\varphi_k|}\tilde{\psi}_{\boldsymbol{k}} \quad \text{and} \quad \tilde{\Psi}_{\boldsymbol{k}}^-(t) \equiv \frac{\lambda_k^*(t)}{|\varphi_k|}\tilde{\psi}_{\boldsymbol{k}} \qquad (62)$$

in terms of the auxiliary variable $\tilde{\psi}_{\boldsymbol{k}}$, we can rewrite $\tilde{\varphi}_{\boldsymbol{k}}^{(3)}$ as:

$$\tilde{\varphi}_{\boldsymbol{k}}^{(3)} = -iM_P^2 \int_{-\infty}^{t_e} \mathrm{d}t \, a(t)^3 c_0(t) H(t)^2 \Big\{ \lambda_k(t)\big[\tilde{\Psi}^+(t)^2\big]_{\boldsymbol{k}} - \lambda_k^*(t)\big[\tilde{\Psi}^-(t)^2\big]_{\boldsymbol{k}} \Big\}, \qquad (63)$$

where

$$\big[\tilde{\Psi}^\pm(t)^2\big]_{\boldsymbol{k}} = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \tilde{\Psi}_{\boldsymbol{p}}^\pm(t)\tilde{\Psi}_{\boldsymbol{k}-\boldsymbol{p}}^\pm(t) \qquad (64)$$

denotes the Fourier transform of $\tilde{\Psi}^\pm(t)^2$. We emphasize that writing the momentum integral of $\tilde{\varphi}_{\boldsymbol{k}}^{(3)}$ as a product in coordinate space makes the numerical calculation much more efficient.[17]

Once the stochastic variable $\tilde{\varphi}$ has been characterized through the auxiliary variable $\tilde{\psi}_{\boldsymbol{k}}$, we are in the position to analyze its statistics and to obtain its PDF by performing a sampling. As we discussed, given the PDF $\tilde{\varphi}$, obtaining the PDF of $\tilde{\zeta}$ is straightforward by their non-linear relation, described in Eq. (22).

In our simulation, we consider a lattice with $N^3 = 256^3$ points. In real space, it is defined within a certain volume $L^3$ and has a resolution of $l^3$. This naturally establishes cutoffs in momentum space,

---

[17]Writing $\tilde{\varphi}_{\boldsymbol{k}}^{(n)}$ as a product in coordinate space can be done at higher orders, $n > 3$, only with the contributions associated with the contact diagram. In those diagrams where there is a certain internal propagation, we will not be able to write the convolution and will be required to perform the momentum integral explicitly, increasing the computational cost.

$k_{\min} = 2\pi/L$ and $k_{\max} = 2\pi/l$. As we discussed at the end of Section 2.3, since we are interested in PBH formation, we must choose these cutoffs such that they resolve the physics involved in the USR phase, as shown in Figure 4.[18] The main steps of our simulation are summarized below:

1. Generate a random (real) Gaussian field with unit variance and Fourier-transform it to obtain $\tilde{\psi}_{\boldsymbol{k}}$. For each node in the lattice in momentum space, use Eq. (54) to compute $\tilde{\varphi}_{\boldsymbol{k}}^{(2)}$, and inverse Fourier-transform to obtain $\tilde{\varphi}^{(2)}(\boldsymbol{x})$.

2. Start with the variable $\tilde{\psi}_{\boldsymbol{k}}$ constructed in the previous step. For each node in the lattice in momentum space, use Eq. (63) to compute $\tilde{\varphi}_{\boldsymbol{k}}^{(3)}$.[19] This requires computing the convolution in (64), which is more efficiently done using the convolution theorem: transform $\tilde{\Psi}_{\boldsymbol{k}}^{+}(t)$ to real space, square it, and transform the result back to Fourier space. Finally, inverse Fourier transform $\tilde{\varphi}_{\boldsymbol{k}}^{(3)}$ to obtain $\tilde{\varphi}^{(3)}(\boldsymbol{x})$.

3. Adding up $\tilde{\varphi}^{(2)}(\boldsymbol{x}) + \tilde{\varphi}^{(3)}(\boldsymbol{x})$, we get a random field with the right power spectrum and bispectrum (up to $H/M_P$ suppressed corrections, as discussed above), whose PDF can be estimated numerically constructing a histogram. The PDF for $\tilde{\zeta}$ can then be obtained using the non-linear relation between $\zeta$ and $\delta\phi$.

The resulting PDF for $\tilde{\zeta}$ –composed by Fourier modes becoming super-Hubble around the USR phase, as defined by the lattice cutoffs– is shown in Figure 6 for the examples of Table 1.

Neglecting the non-Gaussianity of the inflaton perturbation $\delta\phi$ to describe the PDF of $\tilde{\zeta}$, as usually done in stochastic inflation or the $\delta N$ formalism, i.e. by using only Eq. (23), may not be a good approximation, depending on the duration of the USR phase and the abruptness of the transitions between SR and USR and USR and CR. In fact, the Gaussian approximation only works in a small region of parameter space in the model we have considered (see also Figure 3 and the corresponding discussion). We also recall that the region in which the intrinsic contribution to the bispectrum becomes negligible is related to a change of sign in the function $\mathcal{I}$, so that the validity of Eq. (23), when it happens, is due to an accidental cancellation. Incidentally, note that different momentum configurations of $\mathcal{I}$ enter the integral in (59) to compute $\tilde{\varphi}_{\boldsymbol{k}}^{(3)}$, aside from the equilateral limit studied in Sec. 3. As already emphasized, the reason to choose this limit in Sec. 3 was to obtain some simple intuition on the hierarchy between the different sources of non-Gaussianity before moving into the main calculation in the present section.

The examples we have labeled "smooth" and "sharp" (referring to the abruptness of the transitions) display very large changes in the PDF of $\zeta$ once the bispectrum of $\delta\phi$ is taken into account. However, perturbation theory breaks for these examples in the sense discussed in Section 4.2. Indeed, we have checked numerically that the variance of $\tilde{\varphi}^{(3)}$ is comparable to that of $\tilde{\varphi}^{(2)}$. In these cases, our method cannot be applied to derive the correct PDF and can only be used as an illustration that significant deviations from the purely Gaussian approximation can be expected. Truly non-perturbative methods would be required to deal with these examples and other choices of parameters for which perturbation

---

[18]For practical purposes, one can take $l = 1, L = N$, such that $k_{\max} = 2\pi$ and $k_{\min} = 2\pi/N$, and normalize the momenta when solving the equation of motion of $\varphi_k$ such that the peak of the power spectrum roughly coincides with $2\pi/\sqrt{N}$. Also, in our numerical calculation, we stick to the DFT conventions in the FFT module of Numpy. In particular, we define the inverse DFT with a prefactor of $1/N$, consistent with the factor of $1/(2\pi)^3$ in the inverse Fourier transform used for the in-in calculations.

[19]This implies computing one lattice for each time step in the integrand, and then integrating numerically over the samples of the integrand; in our case, we found a time step of 0.15 e-folds to be enough to accurately compute the integral using Simpson's method.

theory breaks. Indeed, in such cases the outcome cannot be improved by including the effects of a finite number of higher-order correlators of $\delta\phi$.

The example we have labeled "intermediate" is interesting because it shows an appreciable difference between including the bispectrum of $\delta\phi$ or not, in a case in which the variances of $\tilde{\varphi}^{(2)}$ and $\tilde{\varphi}^{(3)}$ are, respectively, $\mathcal{O}(10^{-2})$ and $\mathcal{O}(10^{-3})$. We can get an idea of the importance of the effect by computing the integral of the PDF (with and without the inclusion of the bispectrum of $\delta\phi$) from some value of $\zeta$ up to the largest $\zeta$ for which we have data. This integral can be regarded as a first proxy to the calculation of the abundance of PBH, which is highly sensitive to the tail of the PDF. While an actual evaluation of the abundance is a more complex task, this integral is sufficient to illustrate in simple terms the relevance of including the bispectrum. For instance, integrating from $\zeta = 0.5$ we get that the ratio of the integrals is $\sim 1.5$ and if we integrate from $\zeta = 1$, we obtain $\sim 2$ for the same ratio; with the ratio changing linearly as a function of the lower limit of integration.

# 5 Summary and discussion

The assumption of Gaussian inflaton perturbations commonly used to derive the non-Gaussian PDF of $\zeta$ from its non-linear relation to $\delta\phi$ can be insufficient. This can happen for large enough $\zeta$ (which can be relevant for phenomena such as PBH formation) provided that self interactions of $\delta\phi$ are non-negligible. An interesting scenario in this context is that of a transient USR phase that enhances curvature fluctuations at specific comoving scales.

It is therefore important to develop methods that allow to compute the tail of the PDF including the effect of the *intrinsic* non-Gaussianities of $\delta\phi$, i.e. its self interactions. In this work we have approached this problem developing a method that allows to include these non-Gaussianities perturbatively by simulating them on a lattice. By construction, the method can be applied as long as perturbation theory does not break, but it can nonetheless provide interesting insights even when its regime of applicability starts to fail.

We have illustrated the method parametrizing the second slow-roll parameter, $\eta$, which allows to fix the sharpness of the transitions between the phases of inflation (SR, USR and CR) with a number parameter, $\delta$, which controls approximately the duration of said transitions. We determined the tree-level bispectrum of $\delta\phi$ using the in-in formalism and validated our choice of interaction terms checking Maldacena's consistency relation. In practice, we isolated the terms in the cubic interaction Hamiltonian (in the $\delta\phi$-gauge) providing the most relevant contribution to the bispectrum of $\delta\phi$. These turn out to be the self interactions of the form $\delta\phi^3$, coming from the third derivative of the potential of the inflaton. As a first estimate of the relevance of intrinsic non-Gaussianities, we compare the contribution to the bispectrum of $\zeta$ arising from the bispectrum of $\delta\phi$ and the contribution due to a (perturbative) quadratic relation between $\delta\phi$ and $\zeta$, which is non-zero even for a Gaussian $\zeta$. We find that the former contribution is only negligible with respect to the latter for a small range of values of $\delta$ (when the bispectrum of $\delta\phi$ crosses zero as it changes sign). This result already suggests that, in general, one should not neglect intrinsic non-Gaussianities when providing an input for a non-linear relation between $\delta\phi$ and $\zeta$.

To further account for this non-linear relation up to all non-linear orders (in the relation between $\zeta$ and $\delta\phi$) we used the $\delta N$ formalism (which is nothing but a change of gauge) and applied our numerical lattice procedure. We began by generating configurations of values of $\delta\phi$ (in position space) featuring the tree-level power spectrum and bispectrum, which is enough to illustrate the effects of the non-Gaussianities on the PDF of $\zeta$. We then sample this random field and use its PDF as an input for its non-linear relation to $\zeta$ provided by the $\delta N$ formalism. This gives the change in the PDF of the non-linear $\zeta$ obtained

by accounting for the non-zero bispectrum of $\delta\phi$. As anticipated by our first estimate –see the previous paragraph–, the deviations from the naive result assuming Gaussian $\delta\phi$ can be sizable, except for a narrow range of values of $\delta$.

Large variations of the inferred, *intrinsically* modified, PDF occur when perturbation theory starts to break, which underlines the need of developing fully non-perturbative strategies to probe the large-$\zeta$, strongly-interacting regime. In such cases, the PDF we can infer with our perturbative method is not an accurate characterization of the actual PDF. It is just evidence that the non-Gaussianity of $\delta\phi$ is non-negligible. Instead, within the naive regime in which perturbation theory is applicable (and thus our method too), milder modifications to the tail of the PDF are possible. Although we have not studied in detail the phenomenological implications in this regime (e.g. for the abundance of PBH), it can be relevant for model-building.

Finally, we remark that the procedure described in this paper can be generalized, as it allows to systematically include non-Gaussianities of $\delta\phi$ in the input for the non-linear relation with $\zeta$ up to an arbitrary order in perturbation theory, provided that the latter holds.

Our results underline the importance of developing non-perturbative methods to study the tail of the PDF of the curvature fluctuation.

# A    The bispectrum

In this appendix we provide the full expression for the bispectrum, accounting for all terms in the interaction Lagrangian at cubic order, Eq. (27), and we show that the bispectrum in the squeezed limit freezes after horizon crossing. The intrinsic contribution to the bispectrum can be expressed as

$$\mathcal{I}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{k}) = 2\,\mathrm{Im}\left[\int_{-\infty_-}^{t} \mathrm{d}t'\,\varphi_p(t)\varphi_q(t)\varphi_k(t)\sum_{n=0}^{4}C_n^*(t')\right], \tag{65}$$

where $-\infty_- = -\infty(1 - i\omega)$ is responsible for projecting onto the interaction vacuum and the coefficients $C_n$ can be expressed in terms of the coefficients $c_n$, defined in Eq. (29), via

$$C_0(t) \equiv 6M_P^2\,c_0(t)a(t)^3 H(t)^2\varphi_p(t)\varphi_q(t)\varphi_k(t), \tag{66}$$

$$C_1(t) \equiv 2M_P^2\,c_1(t)a(t)^3\varphi_p(t)\dot\varphi_q(t)\dot\varphi_k(t) + (p \leftrightarrow q, k), \tag{67}$$

$$C_2(t) \equiv M_P^2\, c_2(t) a(t)^3 \dot\varphi_p(t) \left[ \frac{\boldsymbol{q}\cdot\boldsymbol{k}}{k^2} \varphi_q(t)\dot\varphi_k(t) + (q \leftrightarrow k) \right] + (p \leftrightarrow q, k), \tag{68}$$

$$C_3(t) \equiv -2M_P^2\, c_3(t) a(t)(\boldsymbol{q}\cdot\boldsymbol{k})\varphi_p(t)\varphi_q(t)\varphi_k(t) + (p \leftrightarrow q, k), \tag{69}$$

$$C_4(t) \equiv 2M_P^2\, c_4(t) a(t)^3 \frac{(\boldsymbol{q}\cdot\boldsymbol{k})^2}{q^2 k^2} \varphi_p(t)\dot\varphi_q(t)\dot\varphi_k(t) + (p \leftrightarrow q, k), \tag{70}$$

where the prefactors correspond to the symmetry factors of each term. For the purpose of analyzing the consistency relation of the bispectrum in the squeezed limit [39], we take $k \ll p$. Then

$$\mathcal{I}(k,p) = 2\,\mathrm{Im}\left\{ \int_{-\infty_-}^{t} \mathrm{d}t' \left[ A(t')\varphi_k^*(t')\varphi_p^*(t')^2 + B(t')\varphi_k^*(t')\dot\varphi_p^*(t')^2 + C(t')\varphi_p^*(t')\dot\varphi_p^*(t')\dot\varphi_k^*(t') \right] \right\}, \tag{71}$$

with

$$\frac{A}{M_P^2} \equiv 6a^3 H^2 c_0 + 2p^2 a\, c_3, \qquad \frac{B}{M_P^2} \equiv 2a^3 c_1 + 2a^3 c_4, \qquad \frac{C}{M_P^2} \equiv 4a^3 c_1 - 2a^3 c_2 + 4\frac{(\boldsymbol{k}\cdot\boldsymbol{p})^2}{k^2 p^2} a^3 c_4. \tag{72}$$

Since $\varphi_k$ is already frozen in the time interval of interest (around the horizon crossing of $p$), one can use $\varphi_k(t') \simeq \varphi_k(t)$ and $\dot\varphi_k = 0$, and the integral simplifies as

$$\frac{\mathcal{I}(k,p)}{|\varphi_k(t)|^2 |\varphi_p(t)|^2} = 2\,\mathrm{Im}\left\{ \frac{\varphi_p(t)^2}{|\varphi_p(t)|^2} \int_{-\infty_-}^{t} \mathrm{d}t' \left[ A(t')\varphi_p^*(t')^2 + B(t')\dot\varphi_p^*(t')^2 \right] \right\}. \tag{73}$$

The bispectrum is related to the spectral index of the power spectrum via the consistency relation (43). Since the power spectrum freezes after the USR phase, the squeezed three-point function $\mathcal{I}(k,p)$ should freeze as well. From Eq. (73) it is not obvious that $\mathcal{I}(k,p)$ is time-independent. However, this can be shown analytically in a model-independent way.

One can integrate the term $B(t')\dot\varphi_p^*(t')^2$ by parts twice and use the equation of motion of $\varphi$ in addition to the quantization condition $\mathrm{Im}[\varphi_p(t)\dot\varphi_p^*(t)] = (4a^3\epsilon)^{-1}$ to eliminate the dependence on $\dot\varphi_p$. This gives

$$\frac{\mathcal{I}(k,p)}{|\varphi_k(t)|^2 |\varphi_p(t)|^2} = -\epsilon(t) - 2\eta(t) + 2M_P^2\,\mathrm{Im}\left\{ \frac{\varphi_p(t)^2}{|\varphi_p(t)|^2} \int_{-\infty_-}^{t} a^3 H^2 \left[ \mathcal{C}(t') - 4\epsilon^2 \left(\frac{p}{aH}\right)^2 \right] \varphi_p^{*2}(t')\,\mathrm{d}t' \right\}, \tag{74}$$

with

$$\mathcal{C} \equiv -2\epsilon\eta \left[ \epsilon^2 + \epsilon_3 \left( 3 - 2\eta + \epsilon_3 + \epsilon_4 \right) - \epsilon \left( 3 - 4\eta + 2\epsilon_3 \right) \right]. \tag{75}$$

Once the modes with momentum $p$ are well outside the horizon, $p \ll aH$, at a time $t' < t$, we obtain analytically and in general that

$$\mathrm{Im}\left[ \frac{\varphi_p(t)^2 \varphi_p^*(t')^2}{|\varphi_p(t)|^2} \right] = 2\,\mathrm{Im}\left[ \frac{\varphi_p(t)\varphi_p^*(t')}{|\varphi_p(t)|} \right] \mathrm{Re}\left[ \frac{\varphi_p(t)\varphi_p^*(t')}{|\varphi_p(t)|} \right] = 2\,\mathrm{Im}\left[ \varphi_p(t)\varphi_p^*(t') \right] + \mathcal{O}\left[ \frac{p}{a(t)H(t)} \right]. \tag{76}$$

By differentiating twice with respect to time and by using the equation of motion for $\varphi_p$ in the super-horizon limit, we arrive to the result

$$\mathrm{Im}\left[ \frac{\varphi_p(t)^2 \varphi_p^*(t')^2}{|\varphi_p(t)|^2} \right] = -\frac{1}{2M_P^2} \int_{t'}^{t} \left( \frac{1}{a^3\epsilon} \right) \mathrm{d}t'' + \mathcal{O}\left[ \frac{p}{a(t)H(t)} \right]. \tag{77}$$

The time integral in Eq. (74) can be performed by splitting it into two pieces: one up to $t_*$ and one from $t_*$ to $t$ such that the $p$-modes are well outside the horizon for all $t > t_*$. In the second interval one can use the relation (77), such that

$$\frac{\mathcal{I}(t|k,p)}{|\varphi_k(t)|^2 |\varphi_p(t)|^2} = \frac{\mathcal{I}(t_*|k,p)}{|\varphi_k(t_*)|^2 |\varphi_p(t_*)|^2} - \epsilon(t) + \epsilon(t_*) - 2\left[\eta(t) - \eta(t_*)\right] - \int_{t_*}^{t} a^3 H^2 \mathcal{C}\,\mathrm{d}t' \int_{t'}^{t} \left( \frac{1}{a^3\epsilon} \right) \mathrm{d}t'', \tag{78}$$

where we neglected terms suppressed by powers of $p/(aH)$. Noticing that $a^3 \mathcal{C} = \mathrm{d}\left[-2a^3\epsilon\eta H(\epsilon_3 - \epsilon)\right]/\mathrm{d}t$ and integrating by parts, this expression simplifies to

$$\frac{\mathcal{I}(t|k,p)}{|\varphi_k(t)|^2|\varphi_p(t)|^2} = \frac{\mathcal{I}(t_*|k,p)}{|\varphi_k(t_*)|^2|\varphi_p(t_*)|^2} + \left[a^3\epsilon\eta H(\epsilon_3 - \epsilon)\right]\Big|_{t_*} \int_{t_*}^t \left(\frac{1}{a^3\epsilon}\right) \mathrm{d}t'. \tag{79}$$

We emphasize that this result is general and that no slow-roll approximation has been made in the derivation. Its interpretation is the following: $\mathcal{I}(t|k,p)$ has a constant and a decaying solution once the modes are outside the horizon. The decaying solution of $\mathcal{I}(t|k,p)$ decays in the same way as the decaying mode of $\varphi_p$ does outside the horizon, as it should. The conditions that guarantee that $\varphi_p(t)$ is constant outside the horizon also ensure that $\mathcal{I}(t|k,p)$ freezes in the same manner.

# B    Failure of the Gram-Charlier series

In this appendix we review the Gram-Charlier (or Edgeworth) series and show that it is, in general, not a good approach to calculate the tail of a distribution if the latter deviates significantly from a Gaussian.

The Gram-Charlier expansion [54] consists in approximating a non-Gaussian distribution $P_{\mathrm{NG}}(x)$ as a series of corrections to a Gaussian distribution $P_{\mathrm{G}}(x)$ of variance $\sigma$,

$$P_{\mathrm{NG}}(x) = P_{\mathrm{G}}(x)\left[1 + \sum_{n=3}^{\infty} a_n H_n\left(x/\sqrt{2}\sigma\right)\right], \tag{80}$$

where $H_n$ denotes the $n$-th order Hermite polynomial

$$H_n(x) = (-1)^n e^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2}. \tag{81}$$

Using the orthogonality of the Hermite polynomials with respect to a Gaussian weight,

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2}\,\mathrm{d}x = \delta_{mn}2^n n!\sqrt{\pi}, \tag{82}$$

one can obtain the coefficients $a_n$ of the Gram-Charlier expansion:

$$a_n = \frac{1}{2^n n!} \int_{-\infty}^{\infty} H_n\left(x/\sqrt{2}\sigma\right) P_{\mathrm{NG}}(x)\,\mathrm{d}x. \tag{83}$$

It is important to remark that this expansion suffers from a series of shortcomings. One is that the PDF resulting from truncating the series at some given order is in general not positive definite. Another is that convergence is in general not guaranteed for an arbitrary PDF. A well-known result [54] is that a sufficient condition for the series to converge is that

$$\int_{-\infty}^{\infty} e^{x^2/4} P_{\mathrm{NG}}(x)\,\mathrm{d}x < \infty, \tag{84}$$

but this is a very strong condition that is not satisfied for a generic PDF.

As a simple example of a situation in which the above series does not converge, we can use the PDF in Eq. (23). We choose $\sigma = 0.2$ and $\eta_{\mathrm{CR}} = -2$. We find the coefficients via Eq. (83) and show the resulting PDF truncated order by order in Figure 7. It is apparent that the convergence of the series becomes progressively worse as more terms are added. Indeed, the Gram-Charlier series is, in general, not an appropriate way to reconstruct a PDF unless the latter is close enough to a Gaussian. Notice that the PDF from the Gram-Charlier series even becomes locally negative at certain perturbative orders, something that by construction does not happen with the procedure we have put forward in this work.
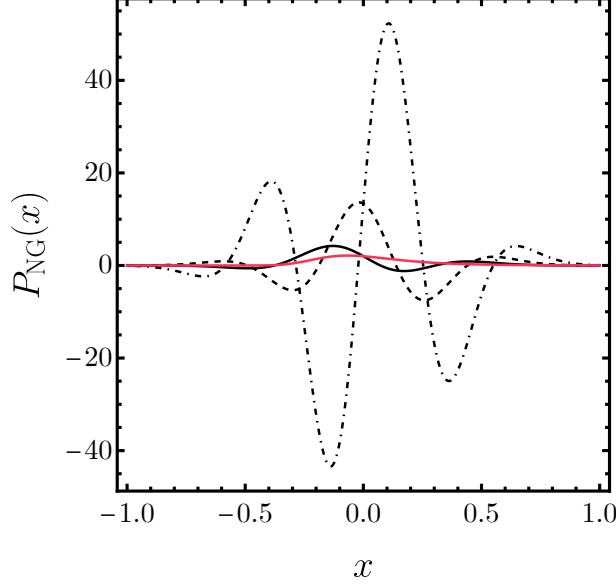
Figure 7: *Non-convergence of the Gram-Charlier series for the PDF in Eq. (23). The PDF is shown in red, and the Gram-Charlier series in Eq. (80) up to third, fourth and fifth order is depicted by the solid black, dashed black, and dot-dashed black lines, respectively.*

# References

[1] W. H. Kinney, "Horizon crossing and inflation with large eta," *Phys. Rev. D* **72** (2005) 023515, `arXiv:gr-qc/0503017`.

[2] J. Garcia-Bellido and E. Ruiz Morales, "Primordial black holes from single field models of inflation," *Phys. Dark Univ.* **18** (2017) 47–54, `arXiv:1702.03901 [astro-ph.CO]`.

[3] G. Ballesteros and M. Taoso, "Primordial black hole dark matter from single field inflation," *Phys. Rev. D* **97** no. 2, (2018) 023501, `arXiv:1709.05565 [hep-ph]`.

[4] 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 [astro-ph.CO]`.

[5] A. M. Green and B. J. Kavanagh, "Primordial Black Holes as a dark matter candidate," *J. Phys. G* **48** no. 4, (2021) 043001, `arXiv:2007.10722 [astro-ph.CO]`.

[6] 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 [astro-ph.CO]`.

[7] D. G. Figueroa, S. Raatikainen, S. Rasanen, and E. Tomberg, "Non-Gaussian Tail of the Curvature Perturbation in Stochastic Ultraslow-Roll Inflation: Implications for Primordial Black Hole Production," *Phys. Rev. Lett.* **127** no. 10, (2021) 101302, `arXiv:2012.06551 [astro-ph.CO]`.

[8] M. Celoria, P. Creminelli, G. Tambalo, and V. Yingcharoenrat, "Beyond perturbation theory in inflation," *JCAP* **06** (2021) 051, arXiv:2103.09244 [hep-th].

[9] A. Caravano, G. Franciolini, and S. Renaux-Petel, "Ultra-Slow-Roll Inflation on the Lattice: Backreaction and Nonlinear Effects," arXiv:2410.23942 [astro-ph.CO].

[10] 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 [astro-ph].

[11] M. Taoso and A. Urbano, "Non-gaussianities for primordial black hole formation," *JCAP* **08** (2021) 016, arXiv:2102.03610 [astro-ph.CO].

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

[13] V. Atal and C. Germani, "The role of non-gaussianities in Primordial Black Hole formation," *Phys. Dark Univ.* **24** (2019) 100275, arXiv:1811.07857 [astro-ph.CO].

[14] 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 [astro-ph.CO].

[15] A. Kalaja, N. Bellomo, N. Bartolo, D. Bertacca, S. Matarrese, I. Musco, A. Raccanelli, and L. Verde, "From Primordial Black Holes Abundance to Primordial Curvature Power Spectrum (and back)," *JCAP* **10** (2019) 031, arXiv:1908.03596 [astro-ph.CO].

[16] D. H. Lyth, K. A. Malik, and M. Sasaki, "A General proof of the conservation of the curvature perturbation," *JCAP* **05** (2005) 004, arXiv:astro-ph/0411220.

[17] N. S. Sugiyama, E. Komatsu, and T. Futamase, "$\delta N$ formalism," *Phys. Rev. D* **87** no. 2, (2013) 023530, arXiv:1208.1073 [gr-qc].

[18] G. Ballesteros and J. G. Egea, "One-loop power spectrum in ultra slow-roll inflation and implications for primordial black hole dark matter," arXiv:2404.07196 [astro-ph.CO].

[19] E. Tomberg, "Stochastic constant-roll inflation and primordial black holes," *Phys. Rev. D* **108** no. 4, (2023) 043502, arXiv:2304.10903 [astro-ph.CO].

[20] G. Ballesteros, T. Konstandin, A. Pérez Rodríguez, M. Pierre, and J. Rey, "Non-Gaussian tails without stochastic inflation," *JCAP* **11** (2024) 013, arXiv:2406.02417 [astro-ph.CO].

[21] D. Cruces, C. Germani, A. Nassiri-Rad, and M. Yamaguchi, "Small noise expansion of stochastic inflation," arXiv:2410.17987 [astro-ph.CO].

[22] J. H. P. Jackson, H. Assadullahi, A. D. Gow, K. Koyama, V. Vennin, and D. Wands, "Stochastic inflation beyond slow roll: noise modelling and importance sampling," arXiv:2410.13683 [astro-ph.CO].

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

[24] V. Atal, J. Cid, A. Escrivà, and J. Garriga, "PBH in single field inflation: the effect of shape dispersion and non-Gaussianities," *JCAP* **05** (2020) 022, `arXiv:1908.11357 [astro-ph.CO]`.

[25] M. Biagetti, V. De Luca, G. Franciolini, A. Kehagias, and A. Riotto, "The formation probability of primordial black holes," *Phys. Lett. B* **820** (2021) 136602, `arXiv:2105.07810 [astro-ph.CO]`.

[26] S. Hooshangi, M. H. Namjoo, and M. Noorbala, "Rare events are nonperturbative: Primordial black holes from heavy-tailed distributions," *Phys. Lett. B* **834** (2022) 137400, `arXiv:2112.04520 [astro-ph.CO]`.

[27] G. Ferrante, G. Franciolini, A. Iovino, Junior., and A. Urbano, "Primordial non-Gaussianity up to all orders: Theoretical aspects and implications for primordial black hole models," *Phys. Rev. D* **107** no. 4, (2023) 043520, `arXiv:2211.01728 [astro-ph.CO]`.

[28] S. Pi and M. Sasaki, "Logarithmic Duality of the Curvature Perturbation," *Phys. Rev. Lett.* **131** no. 1, (2023) 011002, `arXiv:2211.13932 [astro-ph.CO]`.

[29] S. Hooshangi, M. H. Namjoo, and M. Noorbala, "Tail diversity from inflation," *JCAP* **09** (2023) 023, `arXiv:2305.19257 [astro-ph.CO]`.

[30] P. Creminelli, "On non-Gaussianities in single-field inflation," *JCAP* **10** (2003) 003, `arXiv:astro-ph/0306122`.

[31] 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 [astro-ph.CO]`.

[32] 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 [astro-ph.CO]`.

[33] R. L. Arnowitt, S. Deser, and C. W. Misner, "The Dynamics of general relativity," *Gen. Rel. Grav.* **40** (2008) 1997–2027, `arXiv:gr-qc/0405109`.

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

[35] X. Chen, "Primordial Non-Gaussianities from Inflation Models," *Adv. Astron.* **2010** (2010) 638979, `arXiv:1002.1416 [astro-ph.CO]`.

[36] Y. Wang, "Inflation, Cosmic Perturbations and Non-Gaussianities," *Commun. Theor. Phys.* **62** (2014) 109–166, `arXiv:1303.1523 [hep-th]`.

[37] G. Ballesteros, J. Gambín Egea, and F. Riccardi, "Finite parts of inflationary loops," `arXiv:2411.19674 [hep-th]`.

[38] S. Weinberg, "Adiabatic modes in cosmology," *Phys. Rev. D* **67** (2003) 123504, `arXiv:astro-ph/0302326`.

[39] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," *JHEP* **05** (2003) 013, `arXiv:astro-ph/0210603`.

[40] S. Young, "Computing the abundance of primordial black holes," `arXiv:2405.13259 [astro-ph.CO]`.

[41] P. Creminelli and M. Zaldarriaga, "Single field consistency relation for the 3-point function," *JCAP* **10** (2004) 006, arXiv:astro-ph/0407059.

[42] J. Kristiano and J. Yokoyama, "Note on the bispectrum and one-loop corrections in single-field inflation with primordial black hole formation," *Phys. Rev. D* **109** no. 10, (2024) 103541, arXiv:2303.00341 [hep-th].

[43] H. Motohashi and Y. Tada, "Squeezed bispectrum and one-loop corrections in transient constant-roll inflation," *JCAP* **08** (2023) 069, arXiv:2303.16035 [astro-ph.CO].

[44] M. H. Namjoo, H. Firouzjahi, and M. Sasaki, "Violation of non-Gaussianity consistency relation in a single field inflationary model," *EPL* **101** no. 3, (2013) 39001, arXiv:1210.3692 [astro-ph.CO].

[45] J. Martin, H. Motohashi, and T. Suyama, "Ultra Slow-Roll Inflation and the non-Gaussianity Consistency Relation," *Phys. Rev. D* **87** no. 2, (2013) 023514, arXiv:1211.0083 [astro-ph.CO].

[46] R. Bravo, S. Mooij, G. A. Palma, and B. Pradenas, "A generalized non-Gaussian consistency relation for single field inflation," *JCAP* **05** (2018) 024, arXiv:1711.02680 [astro-ph.CO].

[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 [astro-ph.CO].

[48] S. Passaglia, W. Hu, and H. Motohashi, "Primordial black holes and local non-Gaussianity in canonical inflation," *Phys. Rev. D* **99** no. 4, (2019) 043536, arXiv:1812.08243 [astro-ph.CO].

[49] D. Wands, "Duality invariance of cosmological perturbation spectra," *Phys. Rev. D* **60** (1999) 023507, arXiv:gr-qc/9809062.

[50] G. L. Pimentel, L. Senatore, and M. Zaldarriaga, "On Loops in Inflation III: Time Independence of zeta in Single Clock Inflation," *JHEP* **07** (2012) 166, arXiv:1203.6651 [hep-th].

[51] V. Assassi, D. Baumann, and D. Green, "Symmetries and Loops in Inflation," *JHEP* **02** (2013) 151, arXiv:1210.7792 [hep-th].

[52] L. Senatore and M. Zaldarriaga, "The constancy of $\zeta$ in single-clock Inflation at all loops," *JHEP* **09** (2013) 148, arXiv:1210.6048 [hep-th].

[53] D. Green and K. Gupta, "Soft Metric Fluctuations During Inflation," arXiv:2410.11973 [hep-th].

[54] H. Cramér, *Mathematical Methods of Statistics*. Goldstine Printed Materials. Princeton University Press, 1946.