

THE UNIVERSITY OF CHICAGO

THE BLACK HOLE WINDOW ON COSMIC INFLATION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ASTRONOMY AND ASTROPHYSICS

BY
SAMUEL PASSAGLIA

CHICAGO, ILLINOIS

DECEMBER 2020

Committee in charge

Professor Wayne Hu, Chair

Professor Daniel Hooper

Professor Edward W. Kolb

Professor Thomas Crawford

To my grandparents
Abraham and Diana Friedman,
Martin and Alfreda Passaglia.

When the mariner, sailing over tropic seas, looks for relief from his weary watch, he turns his eyes toward the southern cross, burning luridly above the tempest-vexed ocean. As the midnight approaches, the southern cross begins to bend, the whirling worlds change their places, and with starry finger-points the Almighty marks the passage of time upon the dial of the universe, and though no bell may beat the glad tidings, the lookout knows that the midnight is passing and that relief and rest are close at hand.

Let the people everywhere take heart of hope, for the cross is bending, the midnight is passing, and joy cometh with the morning.

— E. V. Debs, Statement to the Court

CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	viii
ACKNOWLEDGMENTS	ix
PREFACE	xi
1 INTRODUCTION	1
1.1 The Cosmic Microwave Puzzle	1
1.2 Cosmic Inflation	3
1.2.1 Canonical single-field slow roll	4
1.2.2 Inflationary perturbations	7
1.3 Primordial Black Hole Dark Matter	9
2 SINGLE-CLOCK INFLATION	12
2.1 Effective Field Theory	12
2.1.1 Unified Lagrangian	13
2.1.2 Scalar Perturbations	16
2.1.3 Cubic Action and Consistency Relation	20
2.1.4 Horndeski and GLPV Subclasses	23
2.2 Power Spectrum and Bispectrum in Generalized Slow Roll	26
2.2.1 GSR and In-In Formalisms	27
2.2.2 Consistency Relation	31
2.3 Worked Example: Transient G-Inflation	34
2.3.1 Scalar Power Spectra	34
2.3.2 GSR Bispectrum	37
2.4 Discussion	43
3 SINGLE-FIELD BEYOND SINGLE-CLOCK	45
3.1 Primordial Black Holes and Ultra-Slow Roll	46
3.1.1 No Go for Single Field Slow Roll	46
3.1.2 Ultra-Slow-Roll Inflation	52
3.2 Transient Ultra-Slow Roll and non-Gaussianity	56
3.2.1 Slow–Small Transition: Inflection-Point Inflation	57
3.2.2 Infinitely Fast Transitions	67
3.2.3 Fast/Slow–Large/Small Transitions	68
3.3 Discussion	83

4	THE HIGGS FIELD DURING INFLATION	85
4.1	The Higgs Instability Mechanism for Primordial Black Holes	85
4.2	Higgs During Inflation	94
4.2.1	Classical Roll vs Stochastic Kicks	96
4.2.2	Background and Linearization	97
4.2.3	Higgs Power Spectrum	101
4.3	Higgs and Reheating	107
4.3.1	Instantaneous Reheating	108
4.3.2	Nonlinear Curvature Evolution	111
4.3.3	Linear Conversion	114
4.3.4	Nonlinear Conversion	120
4.4	Discussion	131
5	CONCLUSION	133
	REFERENCES	134

LIST OF FIGURES

1.1	The cosmic microwave background	2
1.2	Cartoon inflationary potential	6
2.1	Tensor-to-scalar ratio and spectral tilt: theory and experiment	36
2.2	GSR power spectra	37
2.3	Squeezed bispectrum in small-step G-inflation	38
2.4	Squeezed bispectrum in large-step G-inflation	40
2.5	Equilateral bispectrum in small-step G-inflation	41
2.6	Equilateral bispectrum in large-step G-inflation	42
3.1	Phase spaces of slow-roll and ultra-slow roll	48
3.2	Inflection-point background	59
3.3	Inflection-point power spectrum and non-Gaussianity	61
3.4	Ultra-slow-roll transition models	72
3.5	Power spectrum and non-Gaussianity of <i>Large</i> transitions	74
3.6	Non-Gaussianity for <i>Large</i> and <i>Small</i> transitions	76
3.7	Phase space of ultra-slow-roll transitions	77
3.8	Non-Gaussianity from δN	81
4.1	Higgs instability	88
4.2	Position of the Higgs during inflation	90
4.3	Linearity of Higgs perturbations	99
4.4	Inflationary Higgs power spectrum	105
4.5	Higgs energy density around reheating	108
4.6	Induced radiation perturbations	117
4.7	Linear theory curvature perturbation after inflation	119
4.8	Non-linear curvature perturbation mapping	123
4.9	Curvature evolution after reheating	127
4.10	Non-linear transformation of real-space curvature field	129

LIST OF TABLES

2.1	GSR bispectrum operators and sources	32
-----	--	----

ACKNOWLEDGMENTS

This thesis is the product of years of sweat. If at any point the logic blurs and the arguments become confused, charge it only to me and to my faults; but if anything of value is found herein, render the credit to my advisor Wayne Hu. I thank him for his kindness, which I will remember always.

Without my steadfast collaborator Hayato Motohashi, I would have collapsed long ago. Scott Dodelson brought me to Chicago and equipped me for the climb, and I was carried to the summit by my committeemen Dan Hooper, Rocky Kolb, and Tom Crawford. Laticia Rebeles ensured I arrived on time, or nearly so.

The students and postdocs of Wayne's group over the years, that is to say Andrew Long, Austin Joyce, Chen Heinrich, David Zegeye, Giampaolo Benevento, Jose Ezquiaga, Macarena Lagos, Marco Raveri, Meng-Xiang Lin, and especially Pavel Motloch, all received patiently my weekly progress reports and each offered in return their insight and their friendship.

My officemates of the later period, Dimitrios Tanoglidis and Georgios Zacharegkas, supported me hour-by-hour through every trial. My officemates in the earlier period, Alessandro Manzotti and Ross Cawthon, refuse to fade into memory and I miss them deeply.

Héctor Ramírez contributed directly to this work and through shared struggle revealed himself a true friend. Chihway Chang showed unparalleled loyalty to this distant straggler. Kimmy Wu and Kassa Betre treated me gently all along the way.

This journey began with the education I received at the University of Pennsylvania, foremost through the grace of my undergraduate advisor Adam Lidz. The legion of professors there who endeavored to train me in this profession were led by Justin Khoury and Brig Williams. I thank also the City of Philadelphia, which though known for its fraternal love, adopted me as a son.

In the now distant past, Ron Revere instilled in me enough enthusiasm for the natural world to last a lifetime. Before then, the people of France, Japan, and Quebec took in a

stranger and welcomed me as their guest. And throughout my life, the enduring friendship of Lucas Leblanc has kept me steady.

These years in Chicago would have been hollow were it not for John-Henry Pezzuto and Madison Inman; Ruben Waldman, Elizabeth Ashley, and Ashley Guo; Judit Prat and Eric Oberla. Giulia Longhi treated me like family and made the city feel like home. Mr. Brightside and Runa, cats, allowed me to pick them up. Cosmic Microwave Background, though just a starving waif in Jackson Park, managed to do the same for me.

Ogura Satoko worked selflessly and tirelessly every single day for two years to help me achieve a dream, and Yiqi Yan and Shimeng Xu were the best companions in that quest that anyone could ask for.

India Weston supported me and cared for me boundlessly, though I could offer to her nothing in return.

Eric Passaglia and Abigail Friedman, and my siblings Abraham and Marta, know me well enough that no more need be written here. Martin and Alfreda Passaglia have made sure that some corner of this world is always kept warm for me. I walk in Abraham and Diana Friedman's footsteps, and I think of them every day.

Finally I thank you the reader, for your sympathetic regard.

PREFACE

I develop the principles governing the production of our universe's primordial inhomogeneities during its early phase of inflation. As a guiding thread I ask what physics during inflation can lead to perturbations so large that they form black holes in sufficient abundance to be the dark matter. Chapter 1 presents the simplest paradigm for inflation, a single canonical field which slowly rolls, and shows that it cannot produce primordial black hole dark matter. This thesis then proceeds by gradually relieving the assumptions in that simple model.

In Chapter 2, abridged from Refs. [1, 2], I use an effective field theory approach to generalize the canonical single-field model to single-clock inflation, inflation with only one dynamical degree of freedom. I present simple expressions for the power spectrum and bispectrum of the perturbations when slow roll is only transiently violated. I show through an example that they can be applied to the immense variety of models encompassed by the effective field theory. Despite this diversity, forming primordial black holes in single-clock inflation still requires violating slow roll.

In Chapter 3, abridged from Ref. [3], I present the prototypical single-field model which does produce primordial black holes by violating slow-roll, a model known as ultra-slow roll. Ultra-slow roll is an extreme scenario that also violates the assumption of an inflationary attractor solution which underlies single-clock inflation. Because of this, its non-Gaussianity violates a consistency relation and can amplify primordial black hole abundances, an effect I compute for both ultra-slow roll and its realistic implementations.

Finally, I establish in Chapter 4, abridged from Ref. [4], the circumstances under which primordial black holes can be produced from the fluctuations of an extra spectator field during inflation. I focus on the Standard Model Higgs field and study its evolution from its early stochastic phase, through its roll down its large-field instability, all the way to its highly non-linear behavior during reheating.

Chapter 5 provides some brief concluding remarks.

CHAPTER 1

INTRODUCTION

1.1 The Cosmic Microwave Puzzle

In 1964, Penzias and Wilson made the soothing discovery that we are bathed in a faint glow of microwave light [5]. This radiation is made up of relic thermal photons, emitted everywhere in space when our universe was hotter and denser than it is today. Once our universe cooled enough for electrons and protons to form hydrogen, this light could travel largely unimpeded all the way until it reached us today, stretched and diluted by the expansion of the universe to a comfortable 2.7°K. The discovery of the Cosmic Microwave Background (CMB) confirmed the Big Bang model and kickstarted cosmology as a physical science.

On further review though, the CMB exhibits the disconcerting feature that it is nearly the same temperature in all directions. Isotropy is uncomfortable because the light we receive from different directions was emitted from many different regions of space, and when we extrapolate those regions backwards in time, we find they have never been in causal contact with each other. Somehow our universe knows to be at nearly the same temperature everywhere.

Even more unsettling is the *nearly*. The COBE satellite in 1992 measured 1 part in 10^5 anisotropies in the temperature of the CMB [6]. On the one hand, we should be grateful for these small inhomogeneities in the universe at early times, because they seed the gravitational instabilities which lead to the formation of galaxies and cosmologists at late times. However, COBE found that the anisotropies are correlated across the many different causal horizons from which the CMB radiation was emitted – regions of space which have seemingly never talked to each other somehow know to have precisely correlated perturbations.

The WMAP and Planck satellites (see Fig. 1.1), along with many ground- and ballon-based experiments, have confirmed and extended COBE's discoveries, measuring the anisotropies

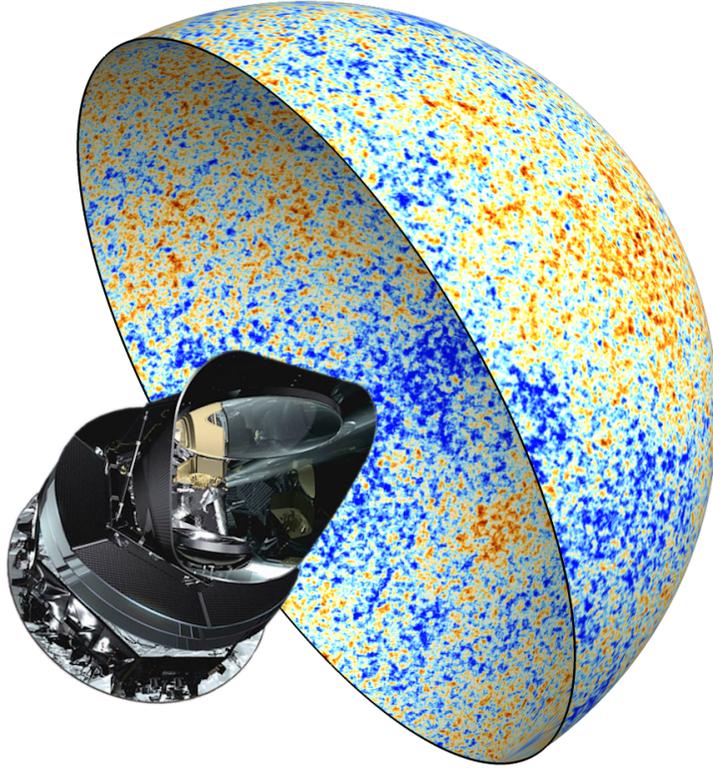


Figure 1.1: The *Planck* spacecraft is the most recent to have observed the microwave background, shown here over half the sky. It has detected superhorizon fluctuations in that map, sourced by the exponential expansion of our universe in its first instants, inflation. *Montage of European Space Agency images.*

of the CMB temperature and polarization to exquisite precision and accuracy over a wide range of scales. In Fourier space, the spectrum of the primordial curvature perturbations ζ as a function of comoving scale k has been measured to be

$$\Delta_{\zeta}^2 = A_s \left(\frac{k}{k_*} \right)^{n_s - 1}, \quad (1.1)$$

with the amplitude $A_s \simeq (2.11 \pm .03) \times 10^{-9}$ and the tilt $n_s \simeq .966 \pm .004$ [7]. $k_* = 0.05 \text{ Mpc}^{-1}$ is just a reference scale well measured by CMB experiments. A scale-invariant power spectrum would have $n_s = 1$, so beyond just the existence of superhorizon correlations we now have strong evidence that they are not exactly scale invariant.

The study of how this primordial curvature spectrum is transformed into the observed spectrum of temperature anisotropies has kept many people gainfully employed for decades now, and has enabled cosmologists to use the CMB as a powerful probe of our universe's contents and history. For example, study of the anisotropies has provided incontrovertible evidence for the existence of dark matter in our universe, to which we will return in §1.3.

This thesis will focus on the physics in the very early universe which explains the puzzling superhorizon correlations. The answer falls under the paradigm of *cosmic inflation*, which explains how seemingly acausal regions communicated, and what they communicated about. Inflation posits new physics at the highest energy scales to explain the appearance of our universe on the largest length scales, and therefore represents the ultimate unification of particle physics and cosmology.

1.2 Cosmic Inflation

In an expanding universe the maximum comoving distance traveled by light since the beginning of time is the causal horizon

$$\eta = \int_0^a \frac{d \ln \tilde{a}}{\tilde{a} H}, \quad (1.2)$$

as a function of the scale factor a . H is the Hubble rate, and with the Friedmann equations we can write the comoving Hubble radius $(aH)^{-1}$ as

$$(aH)^{-1} \propto a^{\frac{1+3w}{2}}, \quad (1.3)$$

where w is the equation of state of the universe. For a universe filled with radiation ($w = 1/3$) or matter ($w = 0$), the comoving Hubble radius is always increasing and the causal horizon is always dominated by the contribution from the last period of expansion.

We see correlations in the CMB emitted by two regions separated by a comoving distance k^{-1} just today entering the comoving Hubble radius, and therefore $k\eta$ becoming larger than

1 only recently. At decoupling they must have $k\eta \ll 1$ and were therefore not causally connected. This is the root of the *horizon problem* introduced in the previous section.

Cosmic inflation posits that prior to the radiation dominated epoch in our universe there was a phase of exponential expansion: $H \sim \text{constant}$, $w \sim -1$, $a \propto e^{Ht}$. During such a phase the comoving Hubble radius (1.3) shrinks and the causal horizon (1.2) is dominated by early time contributions. Regions which we would today calculate, neglecting inflation, to have never been in contact, could in fact have been contained within the same comoving Hubble radius early on. It is then not surprising that the radiation emitted by those regions has a correlated temperature.

We shall soon see that an exponential expansion phase not only allows for distant regions to have once been in contact, but also provides a mechanism to source the perturbations themselves. This is the remarkable success of inflation, and the reason it has been the favored paradigm since its development in the 1980s.

In this introduction we cover the most basic model of inflation, canonical single-field slow roll. The rest of this thesis will then gradually relieve various assumptions of this model and search for the general principles that govern inflation and how these general principles translate to observations.

1.2.1 Canonical single-field slow roll

Scalar fields can easily accommodate the negative-pressure solutions needed to achieve $w \sim -1$ and exponential expansion. The only scalar field in the Standard Model is the Higgs, and we know enough about the Higgs to say that it is probably not the field driving inflation. We will therefore extend the Standard Model with a new scalar field ϕ , minimally coupled to gravity and with a canonical kinetic term, so that our action now has a component

$$S_\phi = \int d^4x \sqrt{-g} \left[\frac{1}{2} g^{\mu\nu} \phi_{,\mu} \phi_{,\nu} - V(\phi) \right]. \quad (1.4)$$

The field $\phi(\vec{x}, t)$ satisfies the Klein-Gordon equation,

$$\square\phi = \frac{dV}{d\phi}, \quad (1.5)$$

which can be linearized into a background piece $\phi(t)$ and a perturbation $\delta\phi(\vec{x}, t)$.

The evolution of the background is then determined by the equation of motion

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0, \quad (1.6)$$

and the Friedmann equation

$$H^2 = \frac{1}{3} \left[\frac{1}{2}\dot{\phi}^2 + V(\phi) \right]. \quad (1.7)$$

Overdots in this chapter denote coordinate time derivatives, and we set $M_{\text{Pl}} = 1$ throughout. We see that so long as the inflaton's energy is dominated by a slowly varying potential, we can have an inflationary solution. We define a slow roll parameter

$$\epsilon_H \equiv -\frac{d \ln H}{dN} = \frac{1}{2} \frac{\dot{\phi}^2}{H^2}, \quad (1.8)$$

which should therefore be much less than unity. We will often use the e -folds of expansion $N \equiv \ln a$ as a time variable during inflation.

We need inflation to last many e -folds (at least ~ 60 based on the horizon size today), and therefore we define a second slow-roll parameter,

$$\eta_H \equiv \frac{d \ln \epsilon_H}{dN}, \quad (1.9)$$

which should also be much less than 1.

When these assumptions hold, the second-order equation of motion for ϕ can be approx-

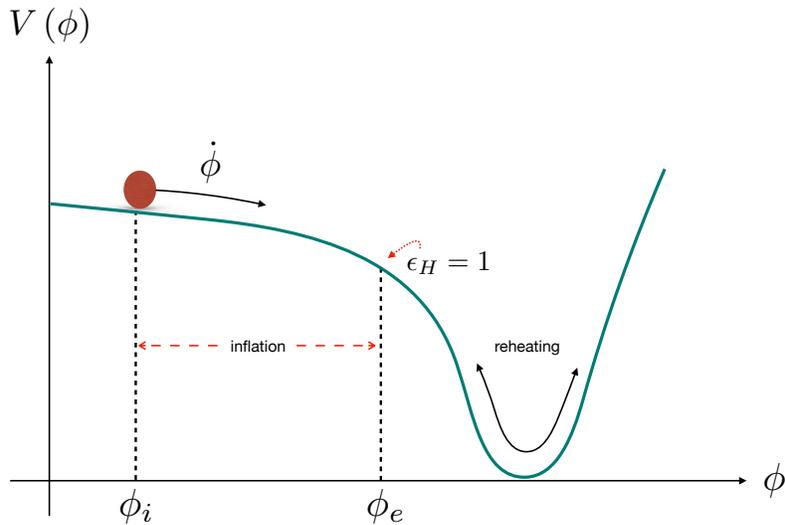


Figure 1.2: A typical inflationary potential. Observable scales cross the horizon early during inflation at ϕ_i , and inflation proceeds until $\epsilon_H \sim 1$ at ϕ_e . Subsequently reheating occurs and the inflaton transfers its energy to Standard Model particles. *Courtesy of Héctor Ramírez.*

imated by the first-order equation

$$\dot{\phi} \simeq -\frac{1}{3H} \frac{dV}{d\phi}, \quad (1.10)$$

with the velocity of the field a function only of the field's position along its potential. This is known as the slow-roll attractor solution.

In Fig. 1.2, we show a cartoon depiction of the inflationary potential convenient for visualization. Inflation occurs on a relatively flat region which supports slow-roll expansion. Observable scales might cross the shrinking comoving Hubble radius and leave causal contact at ϕ_i , while inflation proceeds until $\epsilon_H \sim 1$ at ϕ_e .

We then enter a new phase, reheating, which results in the transfer of the inflaton's energy into a Standard Model plasma to begin our universe's radiation dominated phase. Reheating is the most poorly understood aspect of inflationary theory, but it can usually be

glossed over because cosmologically relevant perturbations are much larger than the comoving Hubble radius at this time and thus reheating can only affect them in constrained ways.

So far we have managed to use single-field slow roll to explain why the CMB is isotropic: everything we see today was once in causal contact early in inflation. Now we will show that inflation can source the small and correlated anisotropies that we see in the sky.

1.2.2 Inflationary perturbations

Linearizing the Klein-Gordon equation (1.5) yields the evolution equation for the Fourier mode $\delta\phi_k$ of the linear perturbation,

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \frac{k^2}{a^2}\delta\phi_k + \frac{d^2V}{d\phi^2}\delta\phi_k = 0. \quad (1.11)$$

Here we have dropped perturbations to the metric, which are not important in slow roll in the spatially flat gauge.

We can scale out the expansion by using the Mukhanov variable $u_k \equiv a\delta\phi_k$ and, to lowest order in slow-roll, the equation of motion simplifies

$$\frac{d^2u_k}{d\tilde{\eta}^2} + \left(k^2 - \frac{2}{\tilde{\eta}^2}\right)u_k = 0, \quad (1.12)$$

where $\tilde{\eta}$ is the positive decreasing conformal time remaining until the end of inflation,

$$\tilde{\eta} \equiv \eta_{end} - \eta \sim \frac{1}{aH}. \quad (1.13)$$

In the subhorizon regime $k\tilde{\eta} \gg 1$, this equation looks like a harmonic oscillator which can be quantized by identifying u_k as the wavefunction a quantum operator \hat{u}_k

$$\hat{u}_k = u_k\hat{a}_k + u_k^*\hat{a}_k^\dagger, \quad (1.14)$$

where \hat{a}_k and \hat{a}_k^\dagger are lowering and raising operators satisfying the usual commutation relation $[\hat{a}_k, \hat{a}_k^\dagger] = 1$. The wavefunction is normalized by imposing the canonical commutation relation $[\hat{u}_k, d\hat{u}_k/d\eta] = i$, yielding

$$u_k(k\tilde{\eta} \rightarrow \infty) = \frac{1}{\sqrt{2k}} e^{-ik\tilde{\eta}} \quad (1.15)$$

which is known as the Bunch-Davies initial condition. Eq. (1.12) classically evolves these subhorizon quantum zero-point modes through horizon crossing, and in slow-roll this has an analytic solution

$$u_k(\tilde{\eta}) = \frac{1}{\sqrt{2k}} \left(1 - \frac{i}{k\tilde{\eta}}\right) e^{-ik\tilde{\eta}}, \quad (1.16)$$

leading to the late time ($k\tilde{\eta} \rightarrow 0$) field fluctuation

$$\delta\phi_k \simeq \frac{iH}{\sqrt{2k^3}}. \quad (1.17)$$

This field fluctuation keeps evolving at late times as H evolves, and therefore it is convenient to switch to a variable which is constant on superhorizon scales, like the curvature perturbation on uniform density slices ζ^1 ,

$$\zeta = -\frac{\delta\rho}{\rho'} \simeq -\frac{\delta\phi}{\phi'}, \quad (1.18)$$

leading to a curvature power spectrum

$$\Delta_\zeta^2(k) = \frac{k^3 |\zeta_k|^2}{2\pi^2} \simeq \frac{H^2}{8\pi^2 \epsilon_H} \Big|_{k\tilde{\eta}=1}, \quad (1.19)$$

which should be evaluated at horizon crossing and is then constant until the mode reenters the horizon in the late universe.

1. Chapters 3 and 4 will cover in great detail scenarios in which ζ is not constant.

This curvature power spectrum is precisely what is measured by the CMB in Eq. 1.1. We see that the near scale-invariance of the measured power spectrum is a reflection of near-constancy of H during inflation. Moreover, the small red-tilt of measured power spectrum is a reflection of the fact that the slow-roll parameters are not exactly zero,

$$n_s - 1 = \frac{d \ln \Delta_\zeta^2}{d \ln k} \sim -4\epsilon_H - 2\eta_H. \quad (1.20)$$

Now that we have identified the observed curvature power spectrum with the two-point vacuum fluctuations of the scalar field, we expect that higher-point curvature correlators encode information about the interactions of that field. In the canonical slow-roll model we have constructed here, however, this is a purely second-order effect, suppressed by additional factors of $\zeta \sim 10^{-5}$, and therefore unlikely to ever be observable.

However, in Chapter 2 we will generalize the action (1.4) by viewing ϕ as the Goldstone boson of broken time-translation symmetries during inflation, and we will drop the assumption of slow-roll. In this case, we will detail how to compute the three-point function and show that it is a powerful tool to distinguish the types of interactions felt by the inflaton. In Chapter 3, we will focus on a particular configuration, the squeezed-limit three-point function, which encodes the effect of a spatial dilation symmetry during inflation and which has important observational consequences.

1.3 Primordial Black Hole Dark Matter

The superhorizon curvature perturbations which we have just computed eventually reenter the horizon, and if they are large enough they can then collapse to primordial black holes (PBHs). PBHs are interesting objects for many reasons, but in this thesis we focus on the fact that in some mass ranges they can be the dark matter (DM).

The fractional energy-density in PBHs at formation time is related to the probability of

the curvature lying above some threshold ζ_c ,

$$\begin{aligned}\beta \equiv \frac{\rho_{\text{PBH}}}{\rho_{\text{tot}}} &= 2 \int_{\zeta_c}^{\infty} d\zeta \frac{1}{\sqrt{2\pi}\Delta_\zeta} e^{-\zeta^2/(2\Delta_\zeta^2)} \\ &\simeq \frac{\sqrt{2}}{\pi} \frac{\Delta_\zeta}{\zeta_c} e^{-\zeta_c^2/(2\Delta_\zeta^2)},\end{aligned}\tag{1.21}$$

where in the last equality we assume PBH formation is a rare event and thus $\zeta_c \gg \Delta_\zeta$. The threshold can be estimated from simulations to be $\mathcal{O}(1)$, which means achieving a sizable abundance requires the power spectrum Δ_ζ^2 to grow relative to its small value at CMB scales (1.1). We therefore usually want to produce PBH DM on very small physical scales, so that there is more time during inflation for this to happen.

However, the mass of a PBH formed by spherical collapse is equal to the mass contained in the horizon volume at collapse time,

$$M_H = \frac{4\pi\rho}{3H^3} = \frac{4\pi}{H},\tag{1.22}$$

and thus smaller scale perturbations lead to smaller mass PBHs. If the PBH mass is too small, the PBH DM will evaporate completely on cosmologically relevant timescales. The lightest PBHs which do not evaporate before matter-radiation equality have mass [8]

$$M \sim 10^{-21} M_\odot,\tag{1.23}$$

which under some light assumptions can be shown to corresponds to physical scales which crossed the horizon

$$N_{\text{PBH}} - N_{\text{CMB}} \sim 40\tag{1.24}$$

e -folds after the CMB scales. To get the right DM abundance, the power spectrum at this

scale has to be

$$\Delta_\zeta(k_{\text{PBH}}) \sim 10^{-2}. \quad (1.25)$$

In order to successfully produce PBH DM using inflationary perturbations, we therefore require a $\sim 10^7$ enhancement of the power spectrum relative to its value on CMB scales (1.1) within ~ 40 e -folds.

From the slow-roll expression for the power spectrum (1.19), this implies that the fractional variation per e -fold of the first slow-roll ϵ_H

$$\left| \frac{\Delta \ln \epsilon_H}{\Delta N} \right| \sim 0.4, \quad (1.26)$$

meaning that there is necessarily an $\mathcal{O}(1)$ violation of the slow-roll condition (1.9) if PBHs are to make up the dark matter.

In fact the estimate here is a very conservative one, since such light PBHs would have observable levels of gamma ray Hawking radiation emission. PBH DM models therefore usually aim to produce PBHs in the ‘asteroid-mass’ window around $10^{-17} \sim 10^{-12} M_\odot$ [9], and require a correspondingly larger slow-roll violation.

Therefore we now have a no-go theorem (originally due to Ref. [8]) for primordial black hole production in canonical single-field slow-roll inflation. Any canonical single-field model which produces primordial black holes in sufficient abundance to be the dark matter must violate slow roll.

The course of this thesis will now be to use primordial black holes as a guide through inflationary theory by gradually generalizing the assumptions made here.

CHAPTER 2

SINGLE-CLOCK INFLATION

In this chapter, based on Refs. [1, 2], we develop the unified effective field theory (EFT) of inflation, the most general form for the inflationary action which is consistent with unbroken spatial diffeomorphisms and a preferred temporal coordinate that represents the inflationary “clock.”

This framework encompasses a wide variety of inflationary models, and we develop a complete formulation of the scalar power spectrum and bispectrum for the EFT in terms of a set of simple one-dimensional integrals which remain valid even when slow-roll is transiently violated. We show analytically that our expressions explicitly preserve the consistency relation between the power spectrum and the squeezed-limit bispectrum so long as the curvature perturbation is conserved outside the horizon.

As an example application of our results, we compute the scalar power spectrum and bispectrum in a model in which potential-driven G-inflation at early times transitions to chaotic inflation at late times, showing that our expressions accurately track the power spectrum and bispectrum when conventional slow-roll approximations fail.

Despite the freedom inherent in the single-clock EFT, producing primordial black hole dark matter still generally requires violations of slow roll.

2.1 Effective Field Theory

In this section we derive the action for scalar metric perturbations up to cubic order in the unified EFT of inflation. We begin in §2.1.1 by reviewing and generalizing the construction of the Lagrangian of the EFT of inflation, which we then expand to cubic order in scalar metric perturbations in §2.1.2. We rewrite this action to make the squeezed-limit consistency relation manifest in §2.1.3. Finally, we study the structure of the EFT in the limits where

it reduces to the Horndeski and beyond-Horndeski GLPV subclasses in §2.1.4.

In general, we find that the cubic action for scalar perturbations can be written in terms of ten operators and manifestly leads to the squeezed-limit consistency relation during slow-roll. In the Horndeski and GLPV limits, six of the ten operators are present.

2.1.1 *Unified Lagrangian*

The unified EFT of inflation was presented in Ref. [10] with the complete set of quadratic operators that contribute to theories where the metric perturbations obey a second-order equation in both time and space and temporal components of the metric remain non-dynamical. These restrictions ensure that the power spectra of scalar and tensor metric fluctuations obey their usual form. We summarize here some of the essential features of that construction while extending it to include the complete set of cubic operators that contribute to the bispectrum.

In the EFT construction, we seek the most general form for the action that is consistent with unbroken spatial diffeomorphisms and a preferred temporal coordinate that represents the inflationary clock. Using this preferred slicing, we decompose the metric into its 3 + 1 Arnowitt-Deser-Misner () form

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \quad (2.1)$$

with the lapse N , the shift N^i , and the spatial metric h_{ij} .

This metric and a unit timelike vector n_μ orthogonal to constant t surfaces define the spatial tensors that compose the EFT action. We construct an action invariant under spatial diffeomorphisms out of a general scalar function of these quantities

$$S = \int d^4x N \sqrt{h} L(N, K^i_j, R^i_j, t), \quad (2.2)$$

in which $K_{\mu\nu} = n_{\mu;\nu} + n_\nu n_{\mu;\beta} n^\beta$ is the extrinsic curvature, R_{ij} is the three-dimensional Ricci tensor with trace $R = R^i{}_i$, and h is the determinant of the three-dimensional metric h_{ij} . Semicolons here and throughout denote covariant derivatives with respect to the metric $g_{\mu\nu}$. Latin indices denote spatial coordinates, which are raised and lowered using h_{ij} . We use the shorthand summation convention

$$S_{i\dots j} T_{i\dots j} \equiv \delta^{i i'} \dots \delta^{j j'} S_{i\dots j} T_{i' \dots j'}, \quad (2.3)$$

for any two spatial tensors S and T .

We have not allowed additional spatial derivatives in Eq. (2.2) since they lead to equations of motion that are beyond second-order in spatial derivatives. Thus we do not encompass the spatially covariant gravity [11, 12] or the Hořava-Lifshitz theories [13, 14, 15]. We have also not allowed the lapse or shift to be dynamical, and thus we do not encompass the full set of degenerate higher order scalar tensor (DHOST)[16, 17] theories. For an extension of the EFT to such models, see Ref. [18].

Next we perturb the action (2.2) around a spatially flat FLRW background,

$$[N] = 1, \quad [N^i] = 0, \quad [h_{ij}] = a^2 \delta_{ij}, \quad (2.4)$$

on which the extrinsic and intrinsic curvature are

$$[K^i{}_j] = H \delta^i{}_j, \quad [R^i{}_j] = 0, \quad (2.5)$$

with $H \equiv d \ln a / dt$. Here and below the notation $[\dots]$ denotes evaluation on the background.

In order to keep all terms that are at most cubic in metric perturbations, we expand the Lagrangian to cubic order in the ADM variables around the background. We define the

Taylor coefficients

$$\begin{aligned}
[L] &= \mathcal{C}, \\
\left[\frac{\partial L}{\partial X_j^i} \right] &= \mathcal{C}_X \delta_i^j, \\
\left[\frac{\partial^2 L}{\partial X_j^i \partial Y_l^k} \right] &= \mathcal{C}_{XY} \delta_i^j \delta_k^l + \frac{\mathcal{C}_{\bar{X}\bar{Y}}}{2} (\delta_i^l \delta_k^j + \delta_{ik} \delta^{jl}), \\
\left[\frac{\partial^3 L}{\partial X_j^i \partial Y_l^k \partial Z_n^m} \right] &= \mathcal{C}_{XYZ} \delta_i^j \delta_k^l \delta_m^n \\
&\quad + \frac{\mathcal{C}_{\bar{X}\bar{Y}\bar{Z}}}{2} \delta_m^n (\delta_i^l \delta_k^j + \delta_{ik} \delta^{jl}) \\
&\quad + \frac{\mathcal{C}_{X\bar{Y}\bar{Z}}}{2} \delta_i^j (\delta_m^l \delta_k^n + \delta_{mk} \delta^{nl}) \\
&\quad + \frac{\mathcal{C}_{\bar{X}Y\bar{Z}}}{2} \delta_k^l (\delta_i^n \delta_m^j + \delta_{im} \delta^{jn}) \\
&\quad + \frac{\mathcal{C}_{\bar{X}\bar{Y}\bar{Z}}}{8} (\delta_k^j \delta_m^l \delta_i^n + \delta_k^j \delta^{ln} \delta_{mi} \\
&\quad + \delta^{jl} \delta_k^n \delta_{mi} + \delta^{jl} \delta_{km} \delta_i^n + \delta_{km} \delta_i^l \delta^{nj} \\
&\quad + \delta_k^n \delta_m^j \delta_i^l + \delta_{ik} \delta^{nj} \delta_m^l + \delta_{ik} \delta_m^j \delta^{nl}),
\end{aligned} \tag{2.6}$$

where $X, Y, Z \in \{N, K, R\}$ and the index structure is determined by the symmetry of the background. We treat scalars and traces with the same notation, so that the tensor $N^i_j = (N/3)\delta_j^i$. Thus $\mathcal{C}_{\bar{N}\bar{X}Y} = \mathcal{C}_{\bar{N}\bar{X}\bar{Y}} = 0$ for any X, Y . Otherwise, these coefficients are arbitrary functions of time which are invariant under subscript permutation in the EFT; they take different concrete forms in different specific inflationary models. Notationally, our $\mathcal{C}_{\bar{X}\bar{Y}}$ is

equal to the $\tilde{\mathcal{C}}_{XY}$ of Ref. [10]. Up to cubic order we can write

$$\begin{aligned}
L = & \frac{1}{3!} \sum_{X,Y,Z} \left(\mathcal{C}_{XYZ} \delta X \delta Y \delta Z + \mathcal{C}_{X\bar{Y}\bar{Z}} \delta X \delta Y^i_j \delta Z^j_i \right. \\
& + \mathcal{C}_{\bar{X}\bar{Y}\bar{Z}} \delta X^i_j \delta Y \delta Z^j_i + \mathcal{C}_{\bar{X}\bar{Y}\bar{Z}} \delta X^i_j \delta Y^j_i \delta Z \\
& \left. + \mathcal{C}_{\bar{X}\bar{Y}\bar{Z}} \delta X^i_j \delta Y^j_k \delta Z^k_i \right) \\
& + \frac{1}{2} \sum_{Y,Z} \left(\mathcal{C}_{YZ} \delta Y \delta Z + \mathcal{C}_{\bar{Y}\bar{Z}} \delta Y^i_j \delta Z^j_i \right) \\
& + \mathcal{C}_N \delta N + \mathcal{C}_R \delta R + \frac{\mathcal{C}_N}{N} - \mathcal{C}_N, \tag{2.7}
\end{aligned}$$

with the sums running through all variable permutations with replacement. We have followed Ref. [10] in using integration by parts to eliminate the linear δK term up to a total derivative term as well as in using the background equation of motion to simplify some of the terms which are constant or linear in geometric quantities.

Finally, to ensure only second-order spatial derivatives in the equation of motion of perturbations we impose

$$\mathcal{C}_{\bar{K}\bar{K}} = -\mathcal{C}_{KK}, \quad \mathcal{C}_{\bar{K}\bar{R}} = -2\mathcal{C}_{KR}, \quad \mathcal{C}_{\bar{R}\bar{R}} = -\frac{8}{3}\mathcal{C}_{RR}. \tag{2.8}$$

This includes the Horndeski and GLPV classes.

2.1.2 Scalar Perturbations

We now restrict our attention to scalar metric perturbations and derive the quadratic and cubic actions for their dynamical field, the curvature perturbation. For scalar perturbations the ADM metric (2.1) takes the form

$$N = 1 + \delta N, \quad N_i = \partial_i \psi, \quad h_{ij} = a^2 e^{2\zeta} \delta_{ij}, \tag{2.9}$$

where we have fixed the residual gauge freedom associated with spatial diffeomorphism invariance by taking a diagonal form for h_{ij} [19]. We call this choice unitary gauge.

In unitary gauge, the perturbed geometric quantities are

$$\begin{aligned}\delta K^i_j &= \frac{1}{1 + \delta N} \left[\delta^i_j \left(\dot{\zeta} - H\delta N \right) + a^{-2} e^{-2\zeta} \left(\delta^{ik} \partial_k \zeta \partial_j \psi \right. \right. \\ &\quad \left. \left. + \delta^{ik} \partial_j \zeta \partial_k \psi - \delta^{ik} \partial_k \partial_j \psi - \delta^i_j \delta^{ab} \partial_a \zeta \partial_b \psi \right) \right], \\ \delta R^i_j &= -a^{-2} e^{-2\zeta} \left[\delta^{ik} \partial_k \partial_j \zeta + \delta^i_j \partial^2 \zeta + \delta^i_j (\partial \zeta)^2 - \delta^{ik} \partial_k \zeta \partial_j \zeta \right].\end{aligned}\quad (2.10)$$

Overdots in this chapter denote coordinate time derivatives, and here and throughout $(\partial \zeta)^2 \equiv \delta^{ab} \partial_a \zeta \partial_b \zeta$ and $\partial^2 \zeta \equiv \delta^{ab} \partial_a \partial_b \zeta$. Variation of the quadratic action with respect to the lapse and shift yields the Hamiltonian and momentum constraints

$$\delta N = D_1 \dot{\zeta}, \quad \psi = D_2 \zeta + a^2 D_3 \chi, \quad (2.11)$$

where χ is an auxiliary variable satisfying $\partial^2 \chi = \dot{\zeta}$ and the parameters D_1 , D_2 , and D_3 are

$$\begin{aligned}D_1 &= \frac{2\mathcal{C}_{KK}}{2H\mathcal{C}_{KK} - \mathcal{C}_{NK}}, \\ D_2 &= \frac{4(\mathcal{C}_{NR} + \mathcal{C}_R - H\mathcal{C}_{KR})}{2H\mathcal{C}_{KK} - \mathcal{C}_{NK}}, \\ D_3 &= \frac{3\mathcal{C}_{NK}^2 - 2\mathcal{C}_{KK}(2\mathcal{C}_N + \mathcal{C}_{NN})}{(2H\mathcal{C}_{KK} - \mathcal{C}_{NK})^2}.\end{aligned}\quad (2.12)$$

Since we are interested in the action to cubic order in perturbations, the lapse and shift should a priori be expanded beyond linear order. However, direct computation shows that the $\mathcal{O}(\zeta^2)$ lapse and shift parameters do not contribute to the cubic action. This is an example of the general result that the $\mathcal{O}(\zeta^2)$ lapse and shift parameters multiply the order $\mathcal{O}(\zeta)$ constraint equations and therefore do not contribute to the cubic action [20, 21, 22].

After eliminating the lapse and shift, the quadratic action for the curvature ζ becomes

$$S_2 = \int d^4x a^3 Q \left[\dot{\zeta}^2 - \frac{c_s^2}{a^2} (\partial\zeta)^2 \right], \quad (2.13)$$

in which Q and c_s^2 are

$$Q = \frac{\mathcal{C}_{KK} (2\mathcal{C}_{KK} (2\mathcal{C}_N + \mathcal{C}_{NN}) - 3\mathcal{C}_{NK}^2)}{(2H\mathcal{C}_{KK} - \mathcal{C}_{NK})^2}, \quad (2.14)$$

$$c_s^2 = \frac{2}{aQ} \left[\frac{d}{dt} \left(a \frac{2\mathcal{C}_{KK}(\mathcal{C}_{NR} + \mathcal{C}_R) - \mathcal{C}_{KR}\mathcal{C}_{NK}}{2H\mathcal{C}_{KK} - \mathcal{C}_{NK}} \right) - a\mathcal{C}_R \right].$$

In terms of the b_s parameter defined in Ref. [10], $Q \equiv \epsilon_H b_s / c_s^2$. The quadratic action provides the linearized equation of motion

$$\partial^2 \zeta = \frac{1}{aQc_s^2} \frac{d}{dt} (a^3 Q \dot{\zeta}). \quad (2.15)$$

We now plug in the perturbed geometric quantities (2.10) into the action (2.2) with the Lagrangian (2.7), eliminating the lapse and shift using the constraint equations (2.11) and retaining terms up to cubic order in ζ . We can also simplify the resulting action using integration by parts. Spatial boundary terms will not contribute to the in-in bispectrum, by momentum conservation, and will be omitted. Temporal boundary terms can contribute significantly and therefore must be retained [23, 24].

Finally, we can also use the linear equation of motion (2.15) to eliminate $\ddot{\zeta}$ -type terms

[25, 26, 27]. The resulting cubic action is

$$\begin{aligned}
S_3 = S_3^{\text{Boundary}} + \int d^3x dt & \left[a^3 F_1 \dot{\zeta}^2 + a F_2 \zeta (\partial\zeta)^2 \right. \\
& + a^3 \frac{F_3}{H} \dot{\zeta}^3 + a^3 F_4 \dot{\zeta} \partial_a \zeta \partial_a \chi + a^3 F_5 \partial^2 \zeta (\partial\chi)^2 \\
& + \frac{F_6}{H^3 a} \dot{\zeta} \partial^2 \zeta \partial^2 \zeta + \frac{F_7}{H^4 a^3} (\partial_a \partial_b \zeta)^2 \partial^2 \zeta \\
& \left. + \frac{F_8}{H^4 a^3} \partial^2 \zeta \partial^2 \zeta \partial^2 \zeta + \frac{F_9}{H^3 a} \partial^2 \zeta (\partial_a \partial_b \zeta) (\partial_a \partial_b \chi) \right], \tag{2.16}
\end{aligned}$$

in which F_1 through F_9 are dimensionless time-dependent functions presented in Ref. [1].

The temporal boundary terms are

$$\begin{aligned}
S_3^{\text{Boundary}} = \int d^3x dt \frac{d}{dt} & \left[a^3 G_1 \dot{\zeta}^3 + a^3 G_2 \zeta \dot{\zeta}^2 \right. \\
& + a G_3 \zeta (\partial\zeta)^2 + a^3 G_4 \dot{\zeta} \partial_a \zeta \partial_a \chi \\
& + a G_5 \dot{\zeta} (\partial\zeta)^2 + \frac{G_6}{a} (\partial\zeta)^2 \partial^2 \zeta \\
& + a^3 G_7 \partial^2 \zeta (\partial\chi)^2 + a G_8 \partial_a \zeta \partial_b \zeta \partial_a \partial_b \chi \\
& \left. + a^3 G_9 \dot{\zeta} (\partial_a \partial_b \chi)^2 \right], \tag{2.17}
\end{aligned}$$

in which G_1 through G_9 are time-dependent functions.

The G_3 and G_6 terms contain no time-derivatives of the fields and therefore do not contribute to bispectrum in the in-in formalism regardless of the behavior of their coefficients [25, 28].

The remaining terms are suppressed relative to the usual $a^3 \zeta^2 \dot{\zeta}$ boundary operator, which shall appear later in our construction, by the presence either of spatial derivatives, which yield relative factors of $k/aH \ll 1$, or by the presence of additional factors of $\dot{\zeta}$, which is suppressed outside the horizon. Therefore none of these terms contribute unless G_n grows sufficiently quickly, so long as the boundary is taken when all modes are outside the horizon.

We restrict our attention to scenarios which satisfy these mild conditions on the EFT

parameters and therefore we hereafter discard S_3^{Boundary} entirely.

2.1.3 Cubic Action and Consistency Relation

We can use to our advantage our ability to reorganize the cubic action using integration by parts and the equation of motion for ζ derived from the quadratic action. In particular, it is well known that in inflation with a single dynamical degree of freedom and a curvature perturbation which remains constant outside the horizon, the bispectrum in the squeezed limit should satisfy the consistency relation [29, 30, 20, 31]

$$\lim_{k_S \rightarrow 0} \frac{B_\zeta(k_S, k_L, k_L)}{P_\zeta(k_S)P_\zeta(k_L)} = -\frac{d \ln \Delta_\zeta^2(k_L)}{d \ln k_L}, \quad (2.18)$$

where B_ζ denotes the curvature bispectrum (see §2.2 for notation). Here the power spectrum P_ζ is related to the dimensionless power spectrum Δ_ζ^2 by

$$\frac{k^3}{2\pi^2} P_\zeta \equiv \Delta_\zeta^2 \simeq \frac{H^2}{8\pi^2 Q c_s^3}, \quad (2.19)$$

where here and below \simeq denotes a slow-roll relation. In the slow-roll approximation, the local slope of the power spectrum is nearly constant and is called the tilt

$$\frac{d \ln \Delta_\zeta^2}{d \ln k} \simeq n_s - 1 = (-2\epsilon_H - q - 3\sigma), \quad (2.20)$$

where $q \equiv \dot{Q}/(HQ)$, $\sigma \equiv \dot{c}_s/(Hc_s)$.

We expect the consistency relation to hold here, but at first glance – or, in the language of §2.2, when plugging in zeroth-order modefunctions – the squeezed-contributing interactions $\zeta \dot{\zeta}^2$ and $\zeta(\partial\zeta)^2$ with their sources F_1 and F_2 are not obviously related to the tilt (2.20). We can rewrite these terms in such a way as to make the consistency relation manifest by generalizing the procedure in Refs. [32, 31].

We first rewrite the squeezed-contributing action in terms of the quadratic Hamiltonian density

$$\mathcal{H}_2 = a^3 Q \left[\dot{\zeta}^2 + \frac{c_s^2}{a^2} (\partial\zeta)^2 \right], \quad (2.21)$$

and the quadratic Lagrangian density

$$\mathcal{L}_2 = a^3 Q \left[\dot{\zeta}^2 - \frac{c_s^2}{a^2} (\partial\zeta)^2 \right], \quad (2.22)$$

such that

$$\begin{aligned} S_{\text{squeezed}} = \int d^3x dt \frac{\zeta}{2Q} & \left[(\mathcal{H}_2 + \mathcal{L}_2) F_1 \right. \\ & \left. + (\mathcal{H}_2 - \mathcal{L}_2) \frac{F_2}{c_s^2} \right]. \end{aligned} \quad (2.23)$$

Next we note that several terms can be grouped into a vanishing boundary term. For a general function of time F ,

$$\begin{aligned} \frac{1}{F} \frac{d}{dt} \left(\frac{F\zeta\mathcal{H}_2}{H} \right) &= \frac{\dot{\zeta}}{H} \mathcal{L}_2 - \zeta (\mathcal{H}_2 + 2\mathcal{L}_2) - (q + \sigma) \zeta \mathcal{L}_2 \\ &+ \left(\frac{\dot{F}}{HF} + \epsilon_H + \sigma \right) \zeta \mathcal{H}_2. \end{aligned} \quad (2.24)$$

Ref. [32] uses a similar relation with $F = 1/c_s^2$ to simplify the action in k -inflation. Here we generalize this grouping using

$$F = \frac{1}{2 + q + \sigma} \left(2 - \frac{F_1}{2Q} + \frac{F_2}{2c_s^2 Q} \right), \quad (2.25)$$

such that the total $\zeta\mathcal{L}_2$ term on the right-hand side of Eq. (2.24) corresponds to the $\zeta\mathcal{L}_2$ term in Eq. (2.23), plus an additional factor of $2\zeta\mathcal{L}_2$.

Making this substitution and using the specific functional forms of F_1 and F_2 , we find

a significant cancellation among the $\zeta\mathcal{H}_2$ terms which results in the squeezed action taking the form

$$S_{\text{squeezed}} = \int dt d^3x \left[\zeta(\mathcal{H}_2 + 2\mathcal{L}_2) - \frac{F}{H}\dot{\zeta}\mathcal{L}_2 + \frac{d}{dt} \left(\frac{F}{H}\zeta\mathcal{H}_2 \right) \right]. \quad (2.26)$$

The boundary term here does not contribute to the bispectrum (see Ref. [32]), and therefore we discard it. The $\dot{\zeta}\mathcal{L}_2$ term does not contribute to the squeezed limit. In order to make the consistency relation more manifest, we undo the grouping by using Eq. (2.24) with $F = 1$. We also use

$$2G\zeta\mathcal{L}_2 = \frac{d}{dt}(Ga^3Q\zeta^2\dot{\zeta}) - \dot{G}a^3Q\zeta^2\dot{\zeta}, \quad (2.27)$$

which holds for all functions of time G , and in particular we use it with $G = \epsilon_H + 3\sigma/2 + q/2$.

After these substitutions and including the terms in Eq. (2.16) that do not contribute to the squeezed limit, we obtain the cubic action for metric perturbations

$$\begin{aligned} S_3 = & \int d^3x dt \left[a^3Q \frac{d}{dt} \left(\epsilon_H + \frac{3}{2}\sigma + \frac{q}{2} \right) \zeta^2\dot{\zeta} \right. \\ & - \frac{d}{dt} \left[a^3Q \left(\epsilon_H + \frac{3}{2}\sigma + \frac{q}{2} \right) \zeta^2\dot{\zeta} \right] \\ & + (\sigma + \epsilon_H)\zeta(\mathcal{H}_2 + 2\mathcal{L}_2) + (1 - F)\frac{\dot{\zeta}\mathcal{L}_2}{H} \\ & + a^3\frac{F_3}{H}\dot{\zeta}^3 + a^3F_4\dot{\zeta}\partial_a\zeta\partial_a\chi \\ & + a^3F_5\partial^2\zeta(\partial\chi)^2 + \frac{F_6}{H^3a}\dot{\zeta}\partial^2\zeta\partial^2\zeta \\ & + \frac{F_7}{H^4a^3}\partial^2\zeta(\partial_a\partial_b\zeta)^2 + \frac{F_8}{H^4a^3}\partial^2\zeta\partial^2\zeta\partial^2\zeta \\ & \left. + \frac{F_9}{H^3a}\partial^2\zeta(\partial_a\partial_b\zeta)(\partial_a\partial_b\chi) \right]. \quad (2.28) \end{aligned}$$

Refs. [32, 31] show explicitly in the context of more restricted inflationary models that the boundary term yields the slow-roll squeezed-limit consistency relation, while the first

term on the first line contributes to the squeezed-limit at higher order in slow-roll, as does the first term on the third line (which can be seen by re-application of Eq. (2.24)). No other term contributes to the squeezed limit at lowest order in slow-roll, and therefore we can immediately see from Eq. (2.28) that the squeezed-limit consistency relation holds in the unified EFT of inflation during slow-roll. In §2.2, we will show that the consistency relation holds even beyond slow-roll.

While the cubic action (2.28) ensures the consistency relation holds in slow-roll, no assumption of slow-roll has been made in its derivation.

2.1.4 *Horndeski and GLPV Subclasses*

Though we write the EFT directly in terms of the metric, the EFT can also be viewed as a four-dimensional scalar-tensor theory by transforming out of unitary gauge using the Stuckelburg trick [11, 33]. In this way, the EFT of inflation presented in §2.1.1 encompasses a large space of fully covariant models. In this section, we study the structure of the cubic action (2.28) derived in §2.1.3 in the Horndeski and GLPV model classes.

The Horndeski and GLPV classes are constructed to avoid the Ostrogradsky instability [34, 35]. The Horndeski class [36] is the most general 4-dimensional scalar-tensor theory with second-order equations of motion for the scalar field ϕ . The Horndeski class can be broadened to include models which have higher than second-order equations of motion yet due to a degeneracy condition do not propagate an Ostrogradsky mode. This is the beyond-Horndeski GLPV class [37], of which the Horndeski class is a subset. The GLPV class is an example of a DHOST theory [16]. While the GLPV model can be represented with an action of the form (2.2), writing the other DHOST theories in our EFT would require generalizing Eq. (2.2) to include time derivatives of the lapse function [17].

The cubic action (2.28) and the resultant bispectrum takes on a restricted form in the Horndeski and GLPV classes. This restriction follows from the ADM representation of the

action for Horndeski and GLPV models [38],

$$\begin{aligned}
L = & A_2 + A_3 K + A_4 (K^2 - K^i_j K^j_i) + B_4 R \\
& + A_5 (K^3 - 3K K^i_j K^j_i + 2K^i_j K^j_k K^k_i) \\
& + B_5 (K^i_j R^j_i - \frac{1}{2} K R).
\end{aligned} \tag{2.29}$$

Here $A_n(X, \phi)$ and $B_n(X, \phi)$ are functions of the kinetic term $X = \nabla^\mu \phi \nabla_\mu \phi$ and field ϕ . In the unitary gauge of ADM, $\phi \rightarrow \phi(t)$ and thus $X = -\dot{\phi}^2/N^2$, so these quantities may also be considered as functions of N and t . In the GLPV class, these functions are completely general, while in the Horndeski class they satisfy

$$\begin{aligned}
A_4 &= 2X B_{4,X} - B_4, \\
A_5 &= -\frac{1}{3} X B_{5,X}.
\end{aligned} \tag{2.30}$$

We then take the appropriate partial derivatives in Eq. (2.7) to get the various \mathcal{C} variables in the Horndeski and GLPV theories. We find

$$\begin{aligned}
\mathcal{C}_N &= -2X (A_{2,X} + 3A_{3,X}H + 6A_{4,X}H^2 + 6A_{5,X}H^3), \\
\mathcal{C}_R &= B_4 - \frac{B_5 H}{2}, \\
\mathcal{C}_{NN} &= 6X \left(A_{2,X} + 3A_{3,X}H + 6A_{4,X}H^2 + 6A_{5,X}H^3 \right) \\
&\quad + 4X^2 \left(A_{2,XX} + 3A_{3,XX}H + 6A_{4,XX}H^2 \right. \\
&\quad \left. + 6A_{5,XX}H^3 \right), \\
\mathcal{C}_K &= A_3 + 4A_4 H + 6A_5 H^2, \\
\mathcal{C}_{KK} &= 2(A_4 + 3A_5 H), \\
\mathcal{C}_{NK} &= -2 \left(A_{3,X} + 4A_{4,X}H + 6A_{5,X}H^2 \right) X,
\end{aligned}$$

$$\begin{aligned}
\mathcal{C}_{NR} &= -2 \left(B_{4,X} - \frac{B_{5,XH}}{2} \right) X, \\
\mathcal{C}_{KR} &= -\frac{B_5}{2}, \\
\mathcal{C}_{NNN} &= -24X \left(A_{2,X} + 3A_{3,XH} + 6A_{4,X}H^2 \right. \\
&\quad \left. + 6A_{5,X}H^3 \right) \\
&\quad - 36X^2 \left(A_{2,XX} + 3A_{3,XXH} + 6A_{4,XX}H^2 \right. \\
&\quad \left. + 6A_{5,XX}H^3 \right) \\
&\quad - 8X^3 \left(A_{2,XXX} + 3A_{3,XXXH} + 6A_{4,XXX}H^2 \right. \\
&\quad \left. + 6A_{5,XXX}H^3 \right), \\
\mathcal{C}_{NNK} &= 6X \left(A_{3,X} + 4A_{4,X}H + 6A_{5,X}H^2 \right) \\
&\quad + 4X^2 \left(A_{3,XX} + 4A_{4,XXH} + 6A_{5,XX}H^2 \right), \\
\mathcal{C}_{NNR} &= 6X \left(B_{4,X} - \frac{B_{5,XH}}{2} \right) \\
&\quad + 4X^2 \left(B_{4,XX} - \frac{B_{5,XXH}}{2} \right), \\
\mathcal{C}_{NKK} &= -4X(A_{4,X} + 3A_{5,X}H), \\
\mathcal{C}_{N\bar{K}\bar{K}} &= 4X(A_{4,X} + 3A_{5,X}H), \\
\mathcal{C}_{NKR} &= XB_{5,X}, \\
\mathcal{C}_{N\bar{K}\bar{R}} &= -2XB_{5,X}, \\
\mathcal{C}_{KKK} &= 6A_5, \\
\mathcal{C}_{\bar{K}\bar{K}\bar{K}} &= -6A_5, \\
\mathcal{C}_{\bar{K}\bar{K}\bar{K}} &= 12A_5, \tag{2.31}
\end{aligned}$$

in which ${}_X \equiv d/dX$ and all other coefficients are either zero or determined by Eq. (2.8).

Using these coefficients, one can show that in the Horndeski and GLPV cases F_6 , F_7 , F_8 , and F_9 are identically zero using the expressions in Ref. [1]. Thus the cubic action reduces

to

$$\begin{aligned}
S_3^{\text{GLPV}} = & \int d^3x dt \left[a^3 Q \frac{d}{dt} \left(\epsilon_H + \frac{3}{2} \sigma + \frac{q}{2} \right) \zeta^2 \dot{\zeta} \right. \\
& - \frac{d}{dt} \left[a^3 Q \left(\epsilon_H + \frac{3}{2} \sigma + \frac{q}{2} \right) \zeta^2 \dot{\zeta} \right] \\
& + (\sigma + \epsilon_H) \zeta (\mathcal{H}_2 + 2\mathcal{L}_2) + (1 - F) \frac{\dot{\zeta} \mathcal{L}_2}{H} \\
& + a^3 \frac{F_3}{H} \dot{\zeta}^3 + a^3 F_4 \dot{\zeta} \partial_a \zeta \partial_a \chi \\
& \left. + a^3 F_5 \partial^2 \zeta (\partial \chi)^2 \right]. \tag{2.32}
\end{aligned}$$

It can also be shown that the F_4 and F_5 operators are suppressed by an additional factor of slow-roll parameters relative to the other operators. This result was shown for the Horndeski class in Ref. [39], and holds also for the GLPV class.

This is the same form of the action as shown in Refs. [26, 40, 41], after undoing our grouping of the F_1 and F_2 terms. Our novel squeezed-action grouping of F_1 and F_2 also confirms the result of Ref. [39] that the squeezed-limit consistency relation holds in slow-roll in Horndeski models and corroborates the result in Ref. [42] that GLPV leads to no new scalar bispectrum shapes relative to Horndeski. By writing it in this form we show that the squeezed-limit consistency relation holds in GLPV models in slow-roll.

2.2 Power Spectrum and Bispectrum in Generalized Slow Roll

From our results so far, it is already clear that it is not much easier to produce PBH DM in the single-clock EFT than in canonical single-field. The slow-roll power spectrum (2.19)

$$\Delta_\zeta^2 \simeq \frac{H^2}{8\pi^2 Q c_s^3}, \tag{2.33}$$

will still have to grow by $\sim 10^7$ in ~ 40 e -folds, as shown in Chapter 1. Instead of implying a large drop in the background ϵ_H as in canonical single field, this now implies a large change in Qc_s^3 ,

$$\left| \frac{d \ln Qc_s^3}{dN} \right| \sim \mathcal{O}(1). \quad (2.34)$$

Depending on how this change is achieved in terms of EFT parameters, it is possible that the background nonetheless proceeds in slow-roll. Regardless, slow-roll approximations for the perturbations must be violated.

We now present in §2.2.1 the generalized slow-roll and in-in formalisms, which we use to construct a complete integral formulation of the power spectrum and bispectrum resulting from the EFT action derived in §2.1, valid when slow-roll is transiently violated. In §2.2.2, we study the squeezed-limit of the bispectrum and show that the consistency relation holds beyond slow-roll. We relegate the explicit forms for the components of some expressions to the published version of this work, Ref. [1].

2.2.1 GSR and In-In Formalisms

The tree-level three-point correlation function in the in-in formalism is given by [32, 20, 43, 44]

$$\langle \hat{\zeta}_{\mathbf{k}_1}^I(t_*) \hat{\zeta}_{\mathbf{k}_2}^I(t_*) \hat{\zeta}_{\mathbf{k}_3}^I(t_*) \rangle = 2 \operatorname{Re} \left[-i \int_{-\infty(1+i\epsilon)}^{t_*} dt \langle \hat{\zeta}_{\mathbf{k}_1}^I(t_*) \hat{\zeta}_{\mathbf{k}_2}^I(t_*) \hat{\zeta}_{\mathbf{k}_3}^I(t_*) H_I(t) \rangle \right], \quad (2.35)$$

with $H_I \simeq -\int d^3x \mathcal{L}_3$ at cubic order [45].

The field operators $\hat{\zeta}^I$ are in the interaction picture, which means their corresponding modefunctions satisfy the free Hamiltonian's equation of motion (2.15). $\hat{\zeta}_{\mathbf{k}}^I$ is the Fourier transform of the operator. We define the corresponding modefunctions $\zeta_k(t)$ as

$$\hat{\zeta}_{\mathbf{k}}^I(t) = \zeta_k(t) \hat{a}(\mathbf{k}) + \zeta_k^* \hat{a}^\dagger(-\mathbf{k}), \quad (2.36)$$

where the annihilation and creation operators satisfy

$$[\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')] = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') \quad (2.37)$$

as usual. Using these relations the power spectrum can be evaluated from the modefunctions at a time t_* taken to be after all the relevant modes have left the horizon

$$\begin{aligned} \langle \hat{\zeta}_{\mathbf{k}}^I(t_*) \hat{\zeta}_{\mathbf{k}'}^I(t_*) \rangle &= (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{k}') |\zeta_k(t_*)|^2 \\ &\equiv (2\pi)^3 \delta^3(\mathbf{k} + \mathbf{k}') P_\zeta(k). \end{aligned} \quad (2.38)$$

Translational and rotational invariance requires that the three-point correlators be encapsulated in the bispectrum B_ζ as

$$\langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle = (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_\zeta(k_1, k_2, k_3), \quad (2.39)$$

in which we have suppressed the evaluation at t_* . The dimensionless parameter conventionally constrained by experiment is

$$f_{\text{NL}}(k_1, k_2, k_3) \equiv \frac{5}{6} \frac{B_\zeta(k_1, k_2, k_3)}{P_\zeta(k_1)P_\zeta(k_2) + \text{perm.}}. \quad (2.40)$$

Here and throughout ‘+perm.’ denotes the two additional cyclic permutations of indices.

In order to evaluate the in-in integral (2.35) and compute $B_\zeta(k_1, k_2, k_3)$, we need to solve the equation of motion (2.15) for the interaction picture modefunctions $\zeta_k(t)$. However, beyond the slow-roll approximation, there is no general analytic solution to the equation of motion. The generalized slow-roll approach is to solve the equation of motion iteratively [32, 46, 47, 48, 49, 50]. It is convenient to express the modefunction in dimensionless form

as

$$y \equiv \sqrt{\frac{k^3}{2\pi^2} \frac{f}{x}} \zeta_k, \quad (2.41)$$

where

$$f \equiv 2\pi a s \sqrt{2Qc_s}, \quad (2.42)$$

$x = ks$ and the sound horizon

$$s \equiv \int_a^{a_{\text{end}}} \frac{d\tilde{a}}{\tilde{a}} \frac{c_s}{\tilde{a}H}, \quad (2.43)$$

with a_{end} denoting the end of inflation.

The formal solution to Eq. (2.15) is

$$y(x) = y_0(x) - \int_x^\infty \frac{d\tilde{x}}{\tilde{x}^2} g(\ln \tilde{s}) y(\tilde{x}) \text{Im} [y_0^*(\tilde{x}) y_0(x)], \quad (2.44)$$

in which $\tilde{x} \equiv k\tilde{s}$, $g(\ln s) \equiv (f'' - 3f')/f$ and $' \equiv d/d \ln s$. The zeroth order solution with Bunch-Davies initial conditions is

$$y_0(x) = \left(1 + \frac{i}{x}\right) e^{ix}. \quad (2.45)$$

The first-order solution is obtained by plugging in the zeroth-order solution into the right hand side of Eq. (2.44).

Every order in the GSR hierarchy of solutions is suppressed relative to the previous order by the g factor, whose time integral is assumed to be small but whose value can evolve and become transiently large unlike in the slow-roll approximation – we call such a case “slow-roll suppressed”. When operators in the cubic action are also slow-roll suppressed, as is the case for the $\zeta^2 \dot{\zeta}$ and $\zeta(\mathcal{H}_2 + 2\mathcal{L}_2)$ terms, it suffices to use the zeroth-order solution for the modefunctions in computing the bispectrum to first-order in slow-roll parameters. Operators with general EFT coefficients, however, are not necessarily slow-roll suppressed and therefore the first-order modefunction solution must be used in order to maintain a

consistent first-order solution.

In the GSR formalism, the power spectrum to first-order in slow-roll parameters is [51, 10]

$$\ln \Delta_{\zeta}^2 = G(\ln s_*) + \int_{s_*}^{\infty} \frac{ds}{s} W(ks) G'(\ln s), \quad (2.46)$$

with the power spectrum window function

$$W(u) = \frac{3 \sin(2u)}{2u^3} - \frac{3 \cos(2u)}{u^2} - \frac{3 \sin(2u)}{2u}, \quad (2.47)$$

the power spectrum source

$$G = -2 \ln f + \frac{2}{3} (\ln f)', \quad (2.48)$$

and s_* the sound horizon corresponding to the superhorizon time t_* .

The first-order bispectrum result follows the same schematic form as the first-order power spectrum result: a windowed integral over a source. Each operator i in the cubic action contributes a set of sources and windows to the bispectrum which are indexed by j according to their asymptotic scalings at $x \ll 1$ and $x \gg 1$. Thus we denote these sources and windows as S_{ij} , W_{ij} .

At zeroth-order in GSR modefunctions, the bispectrum integrals depend only on the triangle perimeter $K \equiv k_1 + k_2 + k_3$, and all shape dependence is held outside the integrand by corresponding k -weights T_{ij} . The integrals take the form

$$I_{ij}(K) = S_{ij}(\ln s_*) W_{ij}(K s_*) + \int_{s_*}^{\infty} \frac{ds}{s} S'_{ij}(\ln s) W_{ij}(K s). \quad (2.49)$$

At first-order in the GSR modefunctions, each operator yields a shape-dependent boundary contribution resulting from the removal of certain nested integrals using integration by parts [32]. These contributions are of the form $[T_{iB} I_{iB}(2k_3) + \text{perm.}]$. Together, the perimeter-dependent and shape-dependent integrals enable computation of the complete

bispectrum of the effective field theory of inflation to first-order in slow-roll parameters,

$$B_\zeta(k_1, k_2, k_3) = \frac{(2\pi)^4 \Delta_\zeta(k_1)\Delta_\zeta(k_2)\Delta_\zeta(k_3)}{4 k_1^2 k_2^2 k_3^2} \quad (2.50)$$

$$\times \left\{ \sum_{ij} T_{ij} I_{ij}(K) + \sum_{i=2}^9 [T_{iB} I_{iB}(2k_3) + \text{perm.}] \right\}.$$

We provide the sources, windows, and k -weights that each operator in the cubic action contributes to this expression in the full published version of this work, Ref. [1]. In Tab. 2.1, we give summary information for each operator. The following section focuses on establishing the squeezed-limit consistency relation beyond slow-roll from these results.

2.2.2 Consistency Relation

In §2.1.3, we argued from the cubic action that the squeezed-limit consistency relation (2.18) holds during slow-roll inflation. Now that we have the complete integral forms of the bispectrum to first-order in slow-roll parameters, we can examine the squeezed-limit consistency relation in more detail, in particular focusing on its form beyond slow-roll.

We first confirm our expectation from §2.1.3 that only the $i = 0$ and $i = 1$ operators contribute in the squeezed-limit. In the published version of this work, Ref. [1], we show from the full expressions for GSR sources, windows, and weights that in the squeezed-limit $x_L/x_S \gg 1$, $x_S \ll 1$, we have

$$\sum_{i=2}^9 \left[\sum_j T_{ij} I_{ij} + [T_{iB} I_{iB}(2k_3) + \text{perm.}] \right] = 0, \quad (2.51)$$

and therefore the operators $i = 2$ to 9 have no net squeezed contribution.

As for the $i = 0$ and $i = 1$ operators, only I_{01} , I_{02} , I_{11} , and I_{12} contribute to squeezed triangles as k_L/k_S . We can then generalize a calculation from Ref. [32] to show that the consistency relation holds even beyond slow-roll. The GSR expression for the squeezed

i	Operator	Source	Squeezed	GLPV
0	$\zeta^2 \dot{\zeta}$	$\frac{2\epsilon_H + 3\sigma + q}{2f}$	yes	Supp.
1	$\zeta(\mathcal{H}_2 + 2\mathcal{L}_2)$	$\frac{\sigma + \epsilon_H}{f}$	yes	Supp.
2	$\dot{\zeta}\mathcal{L}_2$	$\frac{c_s}{aHs} \frac{F-1}{f}$	no	Free
3	$\dot{\zeta}^3$	$-\frac{1}{Q} \frac{c_s}{aHs} \frac{F_3}{f}$	no	Free
4	$\dot{\zeta}(\partial\zeta)\partial\chi$	$-\frac{1}{2Q} \frac{F_4}{f}$	no	Supp.
5	$\partial^2\zeta(\partial\chi)^2$	$\frac{1}{Q} \frac{F_5}{f}$	no	Supp.
6	$\dot{\zeta}\partial^2\zeta\partial^2\zeta$	$\frac{1}{Qc_s^4} \left(\frac{c_s}{aHs}\right)^3 \frac{F_6}{f}$	no	—
7	$(\partial_a\partial_b\zeta)^2\partial^2\zeta$	$\frac{1}{Qc_s^6} \left(\frac{c_s}{aHs}\right)^4 \frac{F_7}{f}$	no	—
8	$(\partial^2\zeta)(\partial^2\zeta)(\partial^2\zeta)$	$\frac{1}{Qc_s^6} \left(\frac{c_s}{aHs}\right)^4 \frac{F_8}{f}$	no	—
9	$(\partial^2\zeta)(\partial_a\partial_b\zeta)(\partial_a\partial_b\chi)$	$\frac{1}{Qc_s^4} \left(\frac{c_s}{aHs}\right)^3 \frac{F_9}{f}$	no	—

Table 2.1: GSR bispectrum operators, sources, whether they contribute to the squeezed-limit, and their status in the GLPV class and its subset the Horndeski class. “Supp.” denotes an operator which is slow-roll suppressed, while “Free” an operator which is not. The $i = 6$, $i = 7$, $i = 8$, and $i = 9$ operators are identically zero in the GLPV and Horndeski classes.

bispectrum is

$$\frac{12}{5} f_{\text{NL}}^{\text{squeezed}} = \lim_{k_S \rightarrow 0} \frac{1}{\Delta(k_S)} \left[-2I_{01}(2k_L) + 4I_{02}(2k_L) + 2I_{11}(2k_L) - 2I_{12}(2k_L) \right]. \quad (2.52)$$

To leading order we can substitute $1/\Delta(k_S) \rightarrow f_*$, resulting in

$$\begin{aligned} \frac{12}{5} f_{\text{NL}} \approx & -2 \frac{f'}{f} \Big|_{s_*} + f_* \int_{s_*}^{\infty} \frac{ds}{s} \left[\left(\frac{\epsilon_H}{f} \right)' W_\epsilon(k_L s) \right. \\ & \left. + \left(\frac{\sigma}{f} \right)' W_\sigma(k_L s) + \left(\frac{q}{f} \right)' W_q(k_L s) \right], \end{aligned} \quad (2.53)$$

where

$$\begin{aligned}
W_\epsilon(x) &= \frac{1}{x} \sin(2x), \\
W_\sigma(x) &= \frac{2}{x} \sin(2x) - \cos(2x), \\
W_q(x) &= \frac{1}{x} \sin(2x) - \cos(2x),
\end{aligned} \tag{2.54}$$

and we have evaluated the boundary term during slow-roll as

$$(2\epsilon_H + 3\sigma + q)|_{s_*} \simeq -2 \frac{f'}{f} \Big|_{s_*}. \tag{2.55}$$

We need to compare this GSR expression for the squeezed bispectrum to the GSR expression for the tilt of the power spectrum [52],

$$\begin{aligned}
\frac{d \ln \Delta_\zeta^2}{d \ln k} \Big|_{k_L} &= \int_{s_*}^{\infty} \frac{ds}{s} W'(k_L s) G'(\ln s) \\
&= 2 \frac{f'}{f} \Big|_{s_*} + \int_{s_*}^{\infty} \frac{ds}{s} \left(\frac{f'}{f} \right)' W_n(k_L s),
\end{aligned} \tag{2.56}$$

where

$$W_n(x) = -2 \cos(2x) + \frac{2}{x} \sin(2x). \tag{2.57}$$

We see immediately from comparing the boundary terms in Eqs. (2.53) and (2.56) that the squeezed limit consistency relation holds in slow-roll. The integral contributions become significant during slow-roll violations. For a sharp feature at $k_L s \gg 1$, the parameters with the highest numbers of derivatives dominate and

$$\left(\frac{f'}{f} \right)' \approx \frac{f''}{f} \approx -\frac{f_*}{2} \left(\frac{\sigma + q}{f} \right)', \tag{2.58}$$

which, when combined with the windows in the desired limit, establishes consistency beyond

slow-roll between Eqs. (2.53) and (2.56).

We have made two assumptions in deriving the consistency relation beyond slow-roll. First, we have assumed that the net change in the power spectrum between two different scales is slow-roll suppressed and thus that we can send $1/\Delta(k_S)$ to f_* . Implicitly this requires that any slow-roll violation is highly transient so that the integrated effect of transient violations remains small. Therefore second, we assume that the sources of slow-roll violation are sharp in their temporal structure using Eq. (2.58). The inflationary model we consider in the following section can violate these approximations by allowing large changes in the power spectrum outside the well observed regime. Nonetheless, we expect that the consistency relation when computed exactly holds in general as long as ζ freezes out after horizon crossing.

2.3 Worked Example: Transient G-Inflation

In this section, we illustrate the calculation of the scalar bispectrum in our general formalism for the unified EFT of inflation with a specific model with cubic galileon interactions in which slow-roll is transiently violated. This model is not constructed to form PBH DM, which we shall instead do in the following chapter, but rather to achieve a specific tensor-to-scalar ratio and resolve certain reheating instabilities. We briefly review this transient G-inflation model and its power spectrum in §2.3.1 and present its bispectrum in §2.3.2.

2.3.1 Scalar Power Spectra

The transient G-inflation model is presented in detail along with its scalar and tensor power spectra in Ref. [2]. We briefly review it here.

We assume that the Lagrangian density takes the form

$$\mathcal{L} = -X/2 - V(\phi) + f_3(\phi)\frac{X}{2}\square\phi + \frac{R}{2}, \quad (2.59)$$

with the chaotic inflation potential $V(\phi) = m^2\phi^2/2$. In Ref. [53], this model is considered with a constant $f_3 = -M^{-3}$. The constant f_3 model suffers from two problems: for the measured value of the scalar tilt n_s , it predicts too large a tensor-to-scalar ratio r ; and for some values of m and M the inflaton has a gradient instability $c_s^2 < 0$ during reheating whose resolution would lie beyond the scope of the perturbative EFT.

Transient G-inflation shuts off the G-inflation term before the end of inflation by using a tanh step-like feature in f_3

$$f_3(\phi) = -M^{-3} \left[1 + \tanh \left(\frac{\phi - \phi_r}{d} \right) \right]. \quad (2.60)$$

Prior to the step, the inflaton is in a G-inflation regime, while after the step, the inflaton follows the slow-roll attractor solution of chaotic inflation. Because the $f_3 X \square \phi$ term in the Lagrangian becomes negligible after the step, the gradient instability at the end of inflation is avoided. By having the transition start just as the CMB scale exits the horizon, the tilt n_s is decoupled from the tensor-to-scalar ratio r and therefore the model can be consistent with observations, as seen in Fig. 2.1.

We consider two parameter sets for the transient G-inflation model, a ‘large-step’ model and a ‘small-step’ model. The large-step model is the fiducial model of Ref. [2]. The inflaton mass scale $m = 2.58 \times 10^{-6}$ is chosen to satisfy the *Planck* 2015 TT+lowP power spectrum amplitude. The Galileon mass scale $M = 1.303 \times 10^{-4}$ suppresses the tensor amplitude relative to the scalar amplitude when the CMB mode $k_{\text{CMB}} = 0.05 \text{ Mpc}^{-1}$ exits the horizon 55 e -folds before the end of inflation. The remaining parameters $\phi_r = 13.87$ and $d = 0.086$ control the step and are chosen such that the tilt and running satisfy observational constraints.

The small-step model is chosen by the same procedure, save for the parameter M which is selected for a larger tensor amplitude. The other parameters are adjusted to keep the tilt and amplitude of the power spectrum fixed. The resultant parameter set is $\{m, M, \phi_r, d\} =$

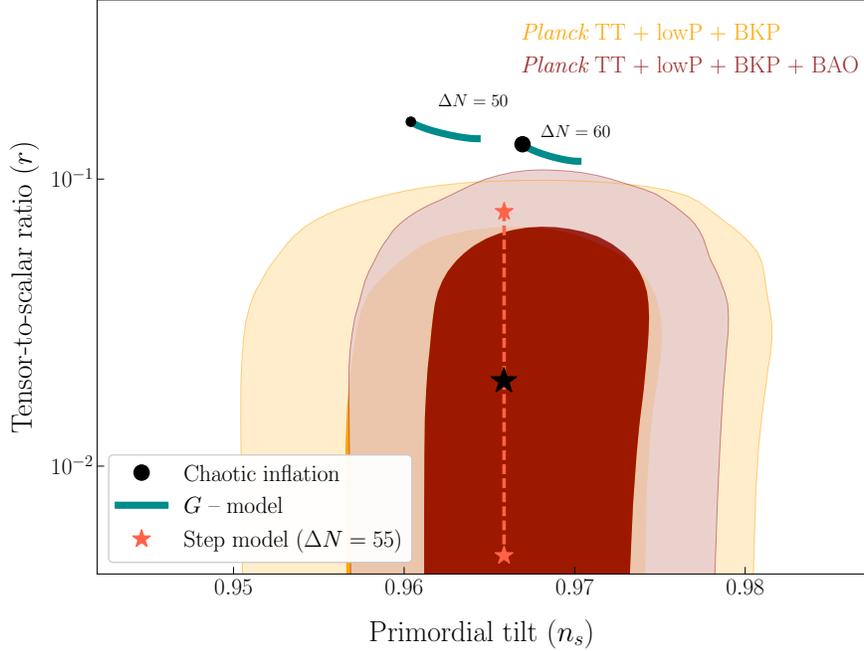


Figure 2.1: *Planck* and BICEP-Keck constraints in the (n_s, r) plane along with the range of predictions from canonical chaotic inflation (black points), G-inflation (green lines), and the transient G-inflation ‘step’ model (orange line and stars). The black star is the ‘large-step’ model of this work.

$\{6.50 \times 10^{-6}, 48.25 \times 10^{-4}, 14.67, 0.021\}$. In this model, the power spectrum evolution before and after the step is much smaller as inflation is never in a fully G-inflation dominated phase, and thus we are closer to the regime of validity of the argument in §2.2.2.

In both models, slow-roll is transiently violated as the inflaton traverses the step, and thus the GSR formalism should be used in place of the traditional slow-roll approach for power spectrum and bispectrum observables. We show the GSR power spectra for these models in Fig. 2.2. In the small-step model, the deviations from scale invariance are small in amplitude but rapidly varying in k (see inset). In the large-step model, they are large in amplitude but smoother in scale. We shall see next that these properties also apply to the bispectrum.

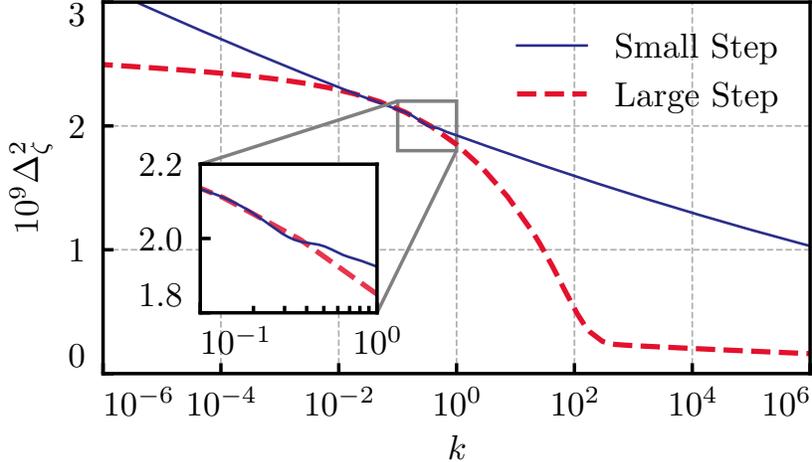


Figure 2.2: The GSR power spectra for the transient G-inflation models we consider. In the small step model, the transition has a small amplitude but rapid variation whereas in the large step model, it has a large amplitude and slow variation. For comparison to exact solutions of the Mukhanov-Sasaki equation, see Ref. [2].

2.3.2 GSR Bispectrum

We now compute the bispectrum for the transient G-inflation models of §2.3.1 using the GSR formulas from §2.2.

We begin by computing the squeezed bispectrum, where the consistency relation allows us to check our computations by comparing the bispectrum result in the squeezed limit to the slope of the GSR power spectrum using Eq. (2.18). We choose to fix the ratio $k_S/k_L = 10^{-2}$. From the analytic analysis in §2.2.2, we know that the only operators which contribute to the squeezed limit are the $i = 0$ and $i = 1$ operators, and their sources are manifestly related to the local slope of the power spectrum. Thus we expect these operators to enforce the consistency relation.

The accuracy of the GSR approximation in the squeezed-limit for the small-step case is shown in Fig 2.3. The GSR bispectrum result closely tracks the consistency relation result before, during, and after the step in the power spectrum. Slow-roll violations during the transition appear as sharp features in the sources which, when integrated against the

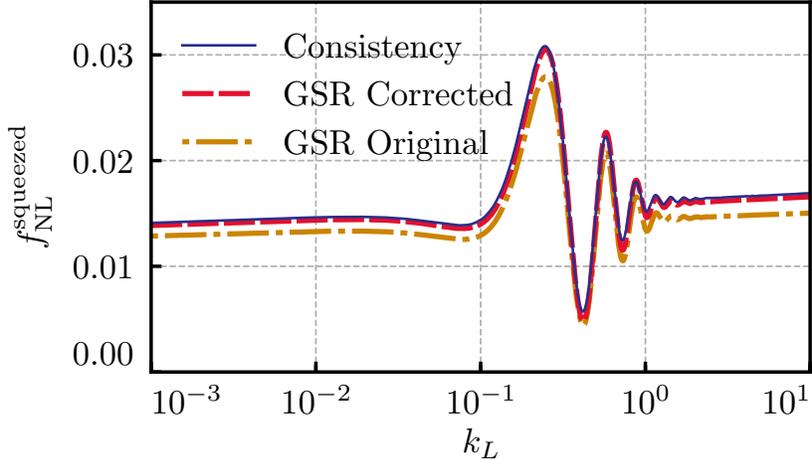


Figure 2.3: Squeezed bispectrum for small-step transient G-inflation. We see excellent agreement between the GSR bispectrum and the consistency relation curve, though with a slight amplitude error. By applying a simple correction to account for modefunction evolution outside the horizon, we can eliminate this error completely.

windows, induce oscillatory features in the squeezed bispectrum and in the tilt of the power spectrum.

While the GSR bispectrum calculation and the power spectrum based consistency relation expectation agree on the period and phase of these features, there is a small amplitude difference between the curves before, during, and after the transition. This error occurs because the bispectrum and power spectrum are calculated to first-order in slow-roll suppressed quantities. In particular the consistency relation check of §2.2.2 ignores corrections due to the evolution in f which would be picked up in the next order of the GSR iteration. Since there is some slow-roll suppressed evolution in f between the epochs when k_S and k_L freeze out, or equivalently in the power spectra at the two scales, a correspondingly small error is induced in the bispectrum.

In this case, where the change in the power spectrum between k_S and k_L is insignificant, this error is minor. Nonetheless, in the upcoming large-step example the power spectrum will significantly evolve across freeze-out epochs and this error will become large. In Refs. [32, 52],

it is shown that next-order terms in the GSR hierarchy provide a correction factor

$$R_0 = 1 + \frac{n_s - 1}{2} \ln \left(\frac{k_S}{k_L} \right), \quad (2.61)$$

assuming that the squeezed bispectrum integrals receive most of their contributions at horizon crossing for k_L .

This correction multiplies the zeroth-order bispectrum contributions from the $i = 0$ and $i = 1$ terms and corrects for the leading-order integrated evolution of f . Since in the following example the power spectrum evolution will be large, we generalize this correction to the non-leading integrated evolution of f by choosing.

$$R = \Delta(k_S)/\Delta(k_L). \quad (2.62)$$

We show in Fig. 2.3 that this correction eliminates the small amplitude error, improving the consistency between the squeezed bispectrum and the derivative of the power spectrum. This correction does not impact triangle shapes where all three modes are comparable in scale. For a formulation of GSR which avoids this type of error by maintaining order-by-order modefunction freeze-out, see Ref. [54].

We show in Fig. 2.4 the squeezed bispectrum for the large-step model. In the large-step model, the $i = 0$ and $i = 1$ sources are much wider than in the small-step case and thus the bispectrum appears as a single peak rather than an oscillatory function. In addition, for this choice of model parameters the power spectrum evolution is large and thus the GSR squeezed bispectrum makes a significant error across the step. Nonetheless, correcting the bispectrum for the integrated evolution f between k_S and k_L with Eq. (2.62) succeeds in explaining most of this discrepancy.

The residual errors in Fig. 2.4 at the peak of the squeezed bispectrum can be understood as a reflection of other iterative corrections in the GSR hierarchy, modes which converge only

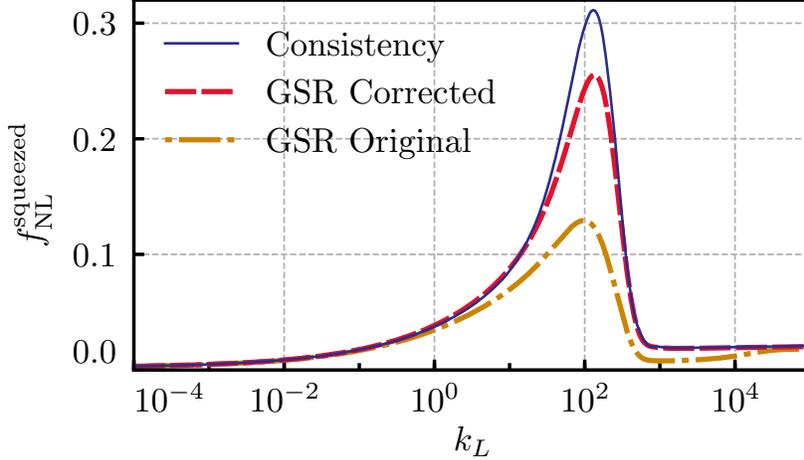


Figure 2.4: Squeezed bispectrum for large-step transient G-inflation. Despite the large evolution of the power spectrum in this model, the corrected GSR bispectrum tracks closely the consistency relation. The discrepancy between the corrected bispectrum result and the consistency relation at the peak indicates that the next-order term in the GSR hierarchy becomes important there.

slowly in this large-step case. These terms are associated with the dynamics of the k_L modes and similar corrections are required for the power spectrum as well. In fact, it is explicitly shown in Ref. [2] that the g terms in the power spectrum expansion reach order unity during the transition, which explains why higher order GSR contributions are necessary to ensure the consistency relation holds at the bispectrum peak.

We next turn to the equilateral bispectrum. Only the $i = 2$ and $i = 3$ operators yield contributions which are not slow-roll suppressed (see Tab. 2.1). In the slow-roll approximation one would take their sources to be constant in Eq. (2.50) and obtain

$$f_{\text{NL}}^{\text{lead; equil.}} \simeq \frac{35}{108}(1 - F)_{\text{SR}} + \frac{5}{81} \left(\frac{F_3}{Q} \right)_{\text{SR}}, \quad (2.63)$$

in which the “SR” subscript denotes that the functions should be expanded to zeroth order in slow-roll. This can be shown to agree analytically with the result for the leading-order equilateral bispectrum in the literature for Horndeski models, Eq. (97) of Ref. [39]. In the specific case of transient G-inflation, Eq. (2.63) takes the form

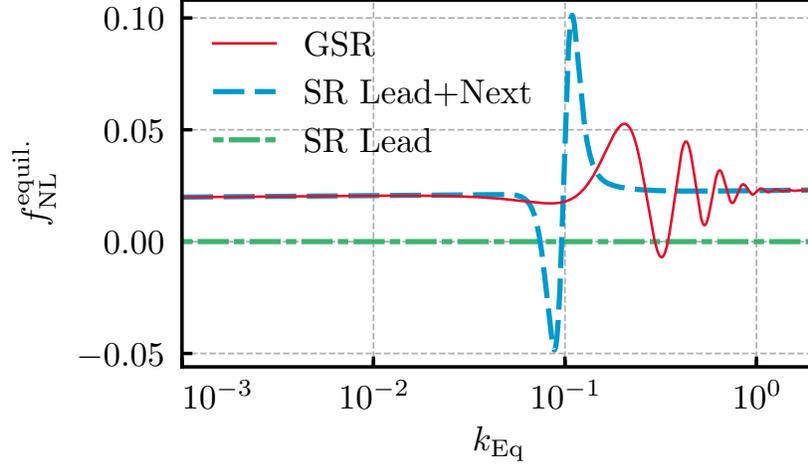


Figure 2.5: Equilateral bispectrum for small-step transient G-inflation. For this set of parameters, $|f_3| \ll 1$ and thus the leading-order slow-roll contribution Eq. (2.64) remains nearly 0. The next-to-leading order SR contribution, Eq. (100) of Ref. [39], dominates, and agrees with the GSR computation before and after the step. During the transition, the SR hierarchy is violated and the SR expression fails to accurately track the GSR bispectrum.

$$f_{\text{NL}}^{\text{SR, equil.}} \simeq \frac{5f_3^2 H^2 \dot{\phi}^2 \left(17 + 94f_3 H \dot{\phi} - 17f_{3,\phi} \dot{\phi}^2 \right)}{81 \left(1 + 4f_3 H \dot{\phi} - f_{3,\phi} \dot{\phi}^2 \right)^2 \left(1 + 6f_3 H \dot{\phi} - f_{3,\phi} \dot{\phi}^2 \right)}, \quad (2.64)$$

in which ${}_{,\phi} \equiv d/d\phi$.

When $|f_3|$ is large, as in pure G-inflation, the leading-order equilateral bispectrum dominates over slow-roll suppressed terms and leads to a larger bispectrum than in canonical inflation. However, when $|f_3|$ is small, as occurs in the small step model and after the transition in the wide step model, the leading order contribution to the equilateral bispectrum is subdominant to the slow-roll suppressed contributions from the $i = 0$ and $i = 1$ operators. For this case, Ref. [39] computes a next-to-leading order contribution to the bispectrum, which results from considering the contributions from slow-roll suppressed operators, the next-order in slow-roll contributions from the $i = 2$ and $i = 3$ operators, as well as SR corrections to the modefunctions.

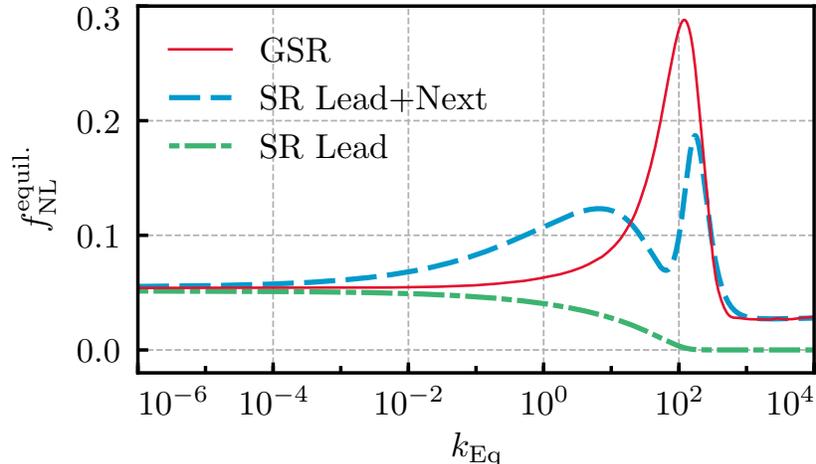


Figure 2.6: Equilateral bispectrum for large-step transient G-inflation. In this case the leading-order SR contribution Eq. (2.64) dominates prior to the transition where $|f_3| \gg 1$. After the transition, the leading-order SR contribution goes to zero while the next-to-leading order terms in Eq. (100) of Ref. [39] come to dominate. The GSR result again agrees with the SR results before and after the transition, while the SR result shows an erroneous double-peak feature during the transition. Despite the enhancement due to the slow-roll violation, $|f_{\text{NL}}^{\text{equil.}}| < 1$ at all times.

In Fig. 2.5 and Fig. 2.6, we compare the total equilateral bispectrum in GSR with the leading-order slow-roll expression (2.64) as well as Eqs. (97) and (100) of Ref. [39], formulas which include the next-to-leading order contributions.

In the small-step case, inflation before and after the transition is nearly canonical and thus the equilateral bispectrum is dominated by the $i = 0$ operator. At the transition the $i = 1$, $i = 2$, and $i = 3$ operators contribute, while the $i = 4$ and $i = 5$ operators remain subdominant throughout. As expected, the leading-order slow-roll bispectrum is subdominant throughout while the slow-roll formula including the next-to-leading order contributions agrees well with GSR before and after the transition. However, during the transition it displays radically different behavior from the GSR curve and fails to reproduce the oscillatory equilateral bispectrum resulting from the sharp sources.

In the large-step case, inflation before the step is in a G-inflation dominated phase. During this phase, the $i = 2$ and $i = 3$ operators dominate the equilateral bispectrum. In

the G-inflation dominated limit, $f_3 \rightarrow -\infty$, the leading-order contribution in slow-roll to the equilateral bispectrum (2.64) approaches $235/3888 \sim 0.06$. The slow-roll suppressed contribution only yields a small adjustment to this value. This is significantly smaller than might be expected from the k -inflation scaling, for example in DBI inflation $f_{\text{NL}}^{\text{equil.}} \simeq \frac{35}{108}(1 - 1/c_s^2)$, which with the G-inflation $c_s^2 \simeq 2/3$ yields $f_{\text{NL}}^{\text{equil.}} \simeq -0.16$. Note also the difference in sign.

After the transition, $f_3 \rightarrow 0$ and the leading-order contribution goes to zero while next-to-leading order contributions become important. Once more, while the leading-order and next-to-leading order SR formulas can accurately track the GSR bispectrum when the usual slow-roll hierarchy is maintained, they fail during the transition when this hierarchy is violated. In particular, the next-to-leading order SR formula predicts an erroneous double peak structure in the equilateral bispectrum.

2.4 Discussion

In this chapter, we developed an effective field theory approach for the study of the power spectrum and bispectrum in single-clock inflation beyond the usual slow-roll approximation. This approach begins with the most general action which breaks temporal diffeomorphisms but preserves spatial diffeomorphisms. In addition we require that the scalar degree of freedom obeys a standard dispersion relation at leading order so that power spectra behave in the usual way.

Our approach of studying the action directly in unitary gauge yields a wider set of terms in the action than explicitly considered in previous work [33, 55, 56, 57], and in particular our action encompasses the Horndeski [36] and GLPV [37] classes.

From this starting point we derive the quadratic and cubic actions for scalar curvature perturbations, making use of integration by parts and the equation of motion while discarding boundary terms which are suppressed outside the horizon. By appropriately grouping

the operators, we isolate the ones that contribute in the squeezed limit and highlight the consistency relation between the power spectrum and the squeezed bispectrum. The resultant cubic action contains ten operators, of which six are present in the Horndeski and GLPV classes, and of these six operators four are slow-roll suppressed.

We then compute the power spectrum and the tree-level bispectrum contribution from each operator using the in-in and GSR formalisms which are valid beyond the slow-roll limit. Our GSR results enable computation of the power spectrum and any bispectrum configuration for all the operators in our action from a set of simple one-dimensional integrals.

In particular the GSR expressions confirm that the consistency relation holds not just in the slow-roll approximation but also in the case of rapidly varying sources. This result extends works which show that the consistency relation explicitly holds in slow-roll, for specific models, or for certain subclasses of EFT operators [29, 58, 39, 32].

As an explicit example, we compute the power spectrum and bispectrum for a specific inflationary model in the Horndeski class in which slow-roll is transiently violated, the transient G-inflation model [2]. For this model, our first-order GSR results for the equilateral bispectrum show qualitatively different behavior from the slow-roll results in the literature during the slow-roll violating phase. This model also highlights corrections for squeezed configurations from non-leading GSR terms which can be important in models in which the power spectrum deviates dramatically from scale-invariance between freeze-out epochs.

The large number of time-dependent coefficients in the EFT of inflation allows a rich range of behavior of perturbations beyond slow-roll. By condensing this large family of coefficients into a small number of integrals, we have provided the tools with which the bispectrum for a very general class of inflation models can be easily studied.

CHAPTER 3

SINGLE-FIELD BEYOND SINGLE-CLOCK

The squeezed-limit consistency relation we emphasized in Chapter 2 came about because the single-clock background is an attractor, so long-wavelength perturbations appear to short-wavelength modes as a simple rescaling of the background. In this chapter, based on Ref. [3], we will first show that the non-Gaussianity produced by this coordinate shift has no effect on PBH abundances as measured in local coordinates, and therefore does not alleviate the no-go theorem for PBH DM in single-field slow-roll that we showed in Chapter 1.

We will then present a single-field model which does produce PBH DM, by violating the slow-roll approximation in a phase known as ultra-slow-roll (USR) [59]. In USR, the field velocity is no longer uniquely determined by the field position and the background is no longer an attractor – USR is therefore a single field model which is not single-clock, and it violates the squeezed-limit consistency relation.

We will therefore study the effect on PBH abundances of USR’s enhanced squeezed non-Gaussianity, and show that it does not have a qualitatively important effect on PBH abundances.

Finally, implementations of USR inflation which are consistent with CMB measurements must have the USR phase be transient. We explore in depth the phenomenology of such transient USR models, such as inflection-point inflation, to show that their squeezed non-Gaussianity is highly suppressed unless they satisfy certain specific conditions which we detail.

3.1 Primordial Black Holes and Ultra-Slow Roll

3.1.1 No Go for Single Field Slow Roll

Following Ref. [8], we showed in Chapter 1 that in canonical single field inflation the comoving curvature power spectrum

$$\Delta_\zeta^2 \equiv \frac{k^3}{2\pi^2} P_\zeta \quad (3.1)$$

must reach at least

$$\Delta_\zeta^2 \sim 10^{-2} \quad (3.2)$$

within ~ 42 e -folds from the epoch when CMB scales exited the horizon, at which $\Delta_\zeta^2 \simeq 10^{-9}$, for the dark matter to be entirely composed of PBHs. In slow roll, the power spectrum satisfies

$$\Delta_\zeta^2 \simeq \frac{H^2}{8\pi^2 \epsilon_H}. \quad (3.3)$$

Therefore such an enhancement of Δ_ζ^2 requires a slow-roll violation of at least $\eta \equiv d \ln \epsilon_H / dN \sim 1$ after horizon exit of the CMB modes but well before the end of inflation.

In this section we update this slow-roll no-go theorem to include local non-Gaussianity which modulates short-wavelength power in a long-wavelength mode. In particular since the formation of a PBH depends on the density fluctuation averaged on the horizon scale at reentry of the perturbations, horizon scale power that is modulated by superhorizon wavelength fluctuations can in principle enhance formation. We study whether such a modulation can make it possible to produce a substantial fraction of the dark matter in PBHs with slow-roll inflation.

In the presence of a long-wavelength fluctuation ζ_L , low pass filtered for comoving wavenumbers $k \leq k_L$, the power spectrum at $k_S \gg k_L$ becomes position dependent

$$P_\zeta(k_S, x) = P_\zeta(k_S) \left[1 + \frac{d \ln P_\zeta(k_S)}{d \zeta_L} \zeta_L(x) \right]. \quad (3.4)$$

By multiplying by and averaging over the long-wavelength mode,

$$\langle \zeta_L(x) P_\zeta(k_S, x) \rangle_{\zeta_L} \simeq \int \frac{d^3 k_L}{(2\pi)^3} B_\zeta(k_L, k_S, k_S), \quad (3.5)$$

we can relate the power spectrum response to the curvature bispectrum B_ζ ,

$$\frac{d \ln P_\zeta(k_S)}{d \zeta_L} \simeq \frac{B_\zeta(k_L, k_S, k_S)}{P_\zeta(k_S) P_\zeta(k_L)} \simeq \frac{12}{5} f_{\text{NL}}(k_L, k_S, k_S). \quad (3.6)$$

Here f_{NL} is the standard dimensionless non-Gaussianity parameter

$$f_{\text{NL}}(k_1, k_2, k_3) \equiv \frac{5}{6} \frac{B_\zeta(k_1, k_2, k_3)}{P_\zeta(k_1) P_\zeta(k_2) + \text{perm.}}, \quad (3.7)$$

in which ‘+ perm.’ denotes the two additional cyclic permutations of indices and the approximation (3.6) assumes the squeezed limit $k_S \gg k_L$.

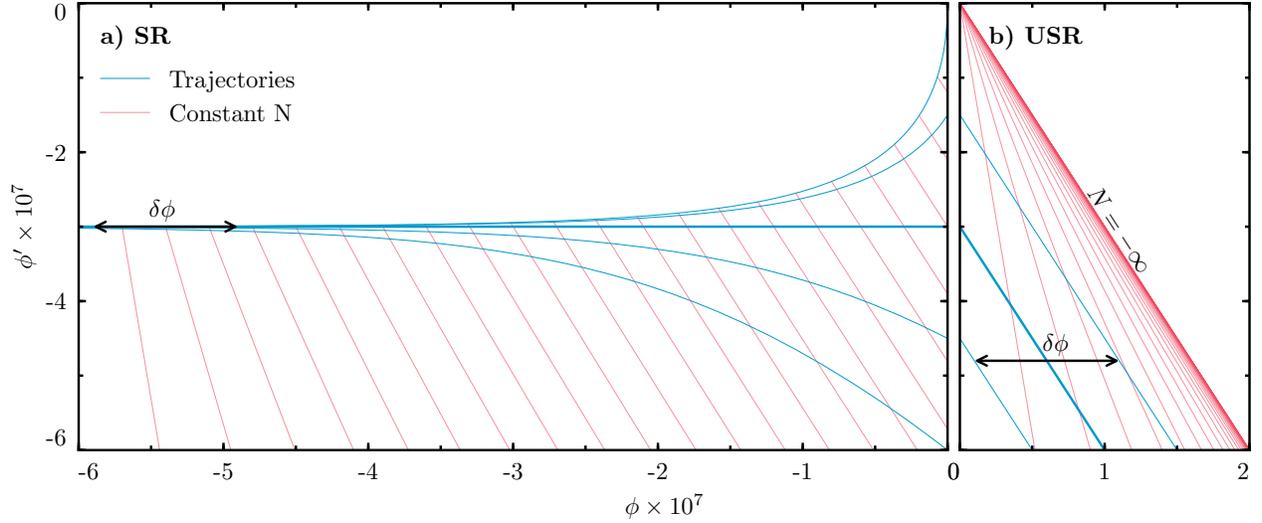


Figure 3.1: Phase space diagram for a) slow roll (SR) and b) ultra-slow roll (USR). Shown are background trajectories (blue lines), lines of constant e -folds (red lines) to the end of inflation (left edge of panels), and field fluctuations $\delta\phi$ (arrows). SR trajectories converge to the attractor for different initial kinetic energies at $\phi = 0$. SR field fluctuations $\delta\phi$ follow the attractor trajectory and can be absorbed into a change in e -folds leaving a change in the relationship between local and global coordinates, but no local imprint on observables once clocks are synchronized to a fixed field value at the end of inflation. USR field fluctuations can still be absorbed into a local background but no longer the background of the unperturbed universe (thick blue line). Since different USR trajectories experience different numbers of e -folds to the end, the power spectrum becomes position dependent, with f_{NL} reflecting the e -folding asymmetry between positive and negative $\delta\phi$ or $\partial^2 N/\partial\phi^2$.

In single-field inflation, $f_{\text{NL}}(k_L, k_S, k_S)$ has a constrained form when ζ is conserved outside the horizon. The curvature perturbation is equivalent to a field fluctuation in spatially flat gauge $\zeta = -\delta\phi/\phi'$, with primes denoting derivatives with respect to e -folds $' = d/dN$ here and throughout. Therefore for a constant ζ , the field fluctuation evolves according to

$$\delta\phi' = \frac{\phi''}{\phi'}\delta\phi, \quad (3.8)$$

and the phase-space trajectory of the long-wavelength field perturbation follows that of the background itself. Short-wavelength modes evolving in a long-wavelength perturbation then also follow the phase-space trajectory of the background, with the only difference being the local e -folds which determines the relationship between physical and comoving wavenumber (see Fig 3.1a).

Single-field inflation on the slow-roll attractor (3.8) therefore satisfies the consistency relation [20]

$$\lim_{k_L/k_S \rightarrow 0} \frac{12}{5} f_{\text{NL}}(k_L, k_S, k_S) = -\frac{d \ln \Delta_\zeta^2(k_S)}{d \ln k_S}. \quad (3.9)$$

This implies a modulation of the small-scale power spectrum due to the long-wavelength mode according to Eqs. (3.4) and (3.6) as

$$P_\zeta(k_S, x) = P_\zeta(k_S) \left[1 - \frac{d \ln \Delta_\zeta^2(k_S)}{d \ln k_S} \zeta_L(x) \right]. \quad (3.10)$$

This modulation is zero at the scale where the power spectrum peaks and corresponds to a dilation of scales rather than an amplitude enhancement.

In general the physical effect of a dilation of scales is to change the mass scale of PBHs rather than enhance their abundance. However in the slow-roll case, there is actually a change in neither abundance nor mass scale. Though the dilation (3.10) does occur in global comoving coordinates, in single-field inflation a freely-falling observer will not see this

dilation locally.

For a given perturbed metric, the standard Fermi normal coordinates (FNC) [60] can be constructed with respect to a central timelike geodesic of a comoving observer [61, 62], such that $g_{\mu\nu}^{\text{FNC}} \simeq \eta_{\mu\nu}$ up to tidal corrections. In order to absorb the effects of superhorizon perturbations out to the horizon scale of a local observer, as required for PBH calculations, we utilize conformal Fermi normal coordinates ($\overline{\text{FNC}}$) [63]. $\overline{\text{FNC}}$ are constructed such that $g_{\mu\nu}^{\overline{\text{FNC}}} \simeq a^2 \eta_{\mu\nu}$, i.e. a conformally flat, locally Friedmann-Lemaître-Robertson-Walker (FLRW) form where the global scale factor a of the background universe is evaluated at the proper time of the central observer.

As shown in Ref. [63], for single-field slow-roll inflation the bispectrum in $\overline{\text{FNC}}$ is related to the comoving-gauge bispectrum by an additional term proportional to the tilt of the power spectrum as

$$\begin{aligned} \lim_{k_L/k_S \rightarrow 0} B_{\bar{\zeta}}(k_L, k_S, k_S) &= P_{\bar{\zeta}}(k_L) P_{\bar{\zeta}}(k_S) \frac{d \ln \Delta_{\bar{\zeta}}^2(k_S)}{d \ln k_S} \\ &+ B_{\zeta}(k_L, k_S, k_S), \end{aligned} \quad (3.11)$$

where barred symbols denote quantities in the $\overline{\text{FNC}}$ frame. This additional term neatly cancels the comoving-gauge squeezed bispectrum from the consistency relation (3.9) and thus in single-field slow-roll inflation

$$\lim_{k_L/k_S \rightarrow 0} B_{\bar{\zeta}}(k_L, k_S, k_S) = 0. \quad (3.12)$$

There is therefore no modulation of the power spectrum in $\overline{\text{FNC}}$

$$P_{\bar{\zeta}}(k_S, x) = P_{\bar{\zeta}}(k_S), \quad (3.13)$$

and the small-scale power spectrum in $\overline{\text{FNC}}$ does not depend on the value of the long-

wavelength perturbation. All local observers therefore see the same small-scale power spectrum regardless of their position in the long-wavelength mode.

Physically, the cancellation in Eq. (3.12) occurs because the bispectrum from the consistency relation encodes the effect on small-wavelength modes of evolving in a separate universe with a background evolution defined by the long-wavelength mode. Once the long-wavelength mode is frozen, this effect is just to change coordinates in the separate universe relative to global coordinates. When making local observations, an observer knows nothing of the global coordinates and instead makes measurements in coordinates corresponding to the separate universe. The formation of PBHs is a local process and so their properties also do not depend on their position in the long-wavelength mode.

This lack of local modulation can also be understood from the phase-space diagram Fig. 3.1a. Relative to the end of inflation at a fixed field value, perturbed trajectories in slow roll are indistinguishable from the background trajectory and thus observers making measurements relative to the end of inflation cannot from any local measurement decide whether they inhabit different regions of a long-wavelength curvature perturbation.

This leads us to our first conclusion: squeezed non-Gaussianity cannot produce PBHs as a significant fraction of the dark matter in canonical single-field slow-roll inflation. For such PBHs to form in canonical single-field inflation, the slow-roll approximation must be violated, at least transiently, to either produce large Gaussian or non-Gaussian fluctuations. In this sense, the slow-roll no-go theorem shown in Ref. [8] is robust and does not change.

Models that evade this no-go result typically have a period when the inflaton rolls on a very flat potential where Hubble friction is insufficient to keep the inflation on the slow-roll attractor. The ultra-slow-roll model, where the inflaton potential is perfectly flat, provides the prototypical example for such studies as we shall see next.

3.1.2 Ultra-Slow-Roll Inflation

Ultra-slow roll [59] is a model of single-field inflation which greatly enhances the scalar power spectrum while also breaking the single-field consistency relation (3.9) for the squeezed bispectrum by violating the attractor condition (3.8) [64, 65]. It is therefore possible to spatially modulate the local power in small scale density fluctuations relevant for PBHs with long-wavelength modes. In this section we examine whether this non-Gaussian modulation can significantly enhance the PBH abundance in ultra-slow roll.

USR is characterized by a potential which is sufficiently flat before its end, which we denote with $\phi = 0$, that the Klein-Gordon equation takes the form

$$\ddot{\phi} \simeq -3H\dot{\phi}, \quad (3.14)$$

where here and throughout overdots denote derivatives with respect to the coordinate time t . If the potential energy dominates then $H \simeq \text{const.}$ and Eq. (3.14) then implies $\phi'' \simeq -3\phi'$ and hence $\phi' \simeq -3\phi + \text{const.}$, defining a family of trajectories in the phase-space diagram, as depicted by the blue trajectories in Fig. 3.1b. Therefore, the phase-space trajectory of the background evolution depends on the initial kinetic energy and does not exhibit attractor behavior.

For an exactly flat potential at $\phi > 0$, an inflaton with insufficient initial kinetic energy will not cross the plateau to reach $\phi = 0$, neglecting stochastic effects. In Fig. 3.1b we focus on classical trajectories that can reach $\phi = 0$ within finite e -folds, and hence the upper right triangle region is inaccessible.

The solution to Eq. (3.14) is $\dot{\phi} \propto a^{-3}$ and so $\epsilon_H \propto a^{-6}$ and $\eta = -6$. Since the analytic solution of the Mukhanov-Sasaki equation for ζ in the superhorizon limit is given by

$$\zeta \simeq c_1 + c_2 \int \frac{dt}{a^3 \epsilon_H}, \quad (3.15)$$

with integration constants c_1 and c_2 , it is dominated by the second mode which grows in USR since $(a^3\epsilon_H)^{-1} \propto a^3$ rather than decays as it does in slow roll. With $H \simeq \text{const.}$, Eq. (3.15) gives $\zeta \propto a^3$ and hence in the spatially flat gauge $\delta\phi = -\zeta\phi' = \text{const.}$, implying that $\delta\phi' = 0$, unlike the case of the slow-roll attractor (3.8).

The power spectrum in this model depends on the value of ϵ_H at the end of USR,

$$\Delta_\zeta^2 \simeq \frac{H^2}{8\pi^2\epsilon_{H,\text{end}}}, \quad (3.16)$$

and thus can be very large if $\epsilon_{H,\text{end}} \ll 1$.

One can employ a gauge transformation from spatially flat gauge to comoving gauge to show that the squeezed-limit non-Gaussianity takes the form (see App. B of Ref. [3])

$$\lim_{k_L/k_S \rightarrow 0} \frac{12}{5} f_{\text{NL}}(k_L, k_S, k_S) = 6. \quad (3.17)$$

Since the USR power spectrum is scale invariant, the large value of f_{NL} in USR violates the consistency relation.

The physical origin of this large value for f_{NL} can be seen from the phase-space diagram Fig. 3.1b. Due to the initial kinetic energy dependence of the background evolution, a USR perturbation cannot be mapped into a change in the background clock along the same phase-space trajectory. Instead, long-wavelength perturbations $\delta\phi$ carry no corresponding $\delta\phi'$ and so shift the USR trajectory to one with a different relationship between ϕ and ϕ' . On this shifted trajectory, the short-wavelength power spectrum attains a different value at the end of USR. More generally, if a local measurement is sensitive to ϕ' at the end of inflation, as in the case of $\Delta_\zeta^2(k_S)$, then different observers will produce different measurements depending on their position in the long-wavelength mode.

This graphical representation of f_{NL} can be turned into a computational method through the so-called δN formalism [66, 67, 68, 69]. When the expansion shear for a local observer

is negligible, as it is in USR above the horizon, the nonlinear evolution of the curvature fluctuation follows the evolution of local e -folds. On spatially flat hypersurfaces, the field fluctuation can be absorbed into a new conformally flat FLRW background on scales much shorter than the wavelength and so the local e -folds may be calculated from the Friedmann equation of a separate universe. The position-dependent power spectrum is therefore the second order change in e -folds due to a short-wavelength $\delta\phi_S$ on top of a long-wavelength $\delta\phi_L$. Since in USR these perturbations leave ϕ' unchanged, the non-Gaussianity parameter can be computed from the e -folds as a function of phase-space position of the background $N(\phi, \phi')$ as

$$\frac{12}{5}f_{\text{NL}} = 2 \frac{\partial^2 N}{\partial \phi^2} \bigg/ \left(\frac{\partial N}{\partial \phi} \right)^2, \quad (3.18)$$

at fixed ϕ' .

The consequence of this formula can be visualized through Fig. 3.1b as the effect of perturbations on phase-space trajectories. Around a chosen background trajectory, the long-wavelength perturbation is reabsorbed into a new background, a horizontal shift to a new trajectory. Short-wavelength perturbations living in this new background induce a second shift in the trajectory, hence the second derivative. Visually, the fact that for the same amplitude of field fluctuation $|\delta\phi|$, a positive fluctuation intersects more surfaces of constant N than a negative fluctuation indicates a large f_{NL} . Refs. [64, 70, 71, 72] follow this approach to analytically compute its value in complete agreement with the in-in approach or the gauge-transformation approach. We shall again exploit the δN formalism in §3.2.

Despite the violation in the consistency relation, the coordinate transformation for the bispectrum Eq. (3.11) still holds and the transformation from global comoving coordinates to $\overline{\text{FNC}}$ leads to the same additional tilt-dependent term in the bispectrum as in the canonical case so long as the transformation to $\overline{\text{FNC}}$ is performed when modes are frozen outside the horizon after the end of inflation.¹ After this time, the construction follows Ref. [63]

1. $\overline{\text{FNC}}$ can still be established during the USR phase but are more closely related to spatially flat gauge

exactly. This procedure of transforming coordinate systems after inflation is followed for slow-roll inflation in Ref. [73] to compute the next-to-leading order term in the bispectrum transformation. Practically, it corresponds to the clock-synchronization condition that all local observers make their measurements at fixed proper time after the end of inflation.

Given the scale invariance of the spectrum, the tilt-dependent transformation from comoving gauge to $\overline{\text{FNC}}$ leaves neither an enhancement of the local power in the long-wavelength mode nor a modulation of the mass of the PBHs. On the other hand, since the transformation term no longer cancels with the comoving-gauge f_{NL} itself, a large value of the latter can in principle enhance PBH formation locally.

If $f_{\text{NL}}(k_L, k_S, k_S)$ is described by the USR result Eq. (3.17), then the local power spectrum can be enhanced by a factor $12/5 \times f_{\text{NL}} \times \zeta_L = 6 \times \zeta_L$. Therefore the non-Gaussian response enhances the local power spectrum by an order unity quantity unless the long-wavelength mode is large, i.e.

$$\zeta_L \gtrsim 10^{-1}. \quad (3.19)$$

However, the scale invariance of USR would then imply

$$\Delta_S^2 = \Delta_L^2 \sim \langle \zeta_L^2(x) \rangle \gtrsim 10^{-2}, \quad (3.20)$$

which satisfies the criterion Eq. (3.2) for PBH formation, and therefore PBHs would already be produced at scale k_S even before accounting for the non-Gaussian response. Note that the conversion from Δ_ζ^2 to spatial variance involves a summation over k_L and gives a logarithmic factor which depends on the total e -folds of USR. In a realistic model this logarithmic factor must be finite so as to also satisfy constraints from the CMB.

This result is the second main conclusion of this work: in a USR model which does

than comoving gauge in temporal synchronization (see also App. B of Ref.[3]). In spatially flat gauge, a superhorizon field fluctuation $\delta\phi$ can be absorbed into a new, nearly conformally flat FLRW background, as we exploit with the δN formalism.

not produce a significant PBH abundance under the Gaussian approximation, the squeezed non-Gaussian response enhances the local power spectrum by at most

$$\frac{\Delta P_\zeta}{P_\zeta} \lesssim 1, \quad (3.21)$$

and therefore the squeezed non-Gaussian response does not qualitatively change Gaussian conclusions. Of course as they originate from rare fluctuations, PBHs can change in their abundance but these changes can be reabsorbed into model parameters that make no more than an order unity change in the power spectrum. In particular squeezed non-Gaussianity cannot make a model that falls far short of making PBHs the dark matter under the Gaussian assumption into one that does.

Since inflation has to end and observational constraints should be satisfied on CMB scales, the simple picture presented here must be modified to account for transitions into and out of USR. In §3.2 we shall explore whether even this level of enhancement still holds in such models of transient USR inflation.

3.2 Transient Ultra-Slow Roll and non-Gaussianity

In addition to a graceful exit problem, USR inflation is incompatible with the measured tilt of the CMB power spectrum [7] and is in tension with constraints on local non-Gaussianities in the CMB [74], and therefore any USR phase must begin after CMB modes exit the horizon and must take care not to grow those modes after horizon exit.

One model proposed in the literature for PBH production with a transient USR phase is inflection-point inflation [8, 75]. In §3.2.1, we show that the transition out of USR in inflection-point inflation induces

$$f_{\text{NL}}(k_L, k_S, k_S) \ll 1, \quad (3.22)$$

and therefore non-Gaussianities do not enable PBHs to be the dark matter in inflection-point inflation.

This numerical result can be understood from Ref. [71]’s analytic study of infinitely sharp potential transitions between USR and SR, which we review briefly in §3.2.2. Transitions where the inflaton velocity monotonically decreases to reach an attractor solution lead to squeezed non-Gaussianity that is proportional to the potential slow-roll parameters on the attractor. Conversely, transitions where the inflaton instantly goes from having too much kinetic energy for the potential it evolves on to suddenly having insufficient kinetic energy for a now much steeper potential conserve the USR non-Gaussianity. We call the latter transitions **large**, which we will define specifically below [see Eq. (3.39)].

In §3.2.3, we generalize the analysis of Ref. [71] to potentials which do not have an infinitely sharp break, and in particular we study how quickly the inflaton must traverse the potential feature to reproduce the USR result. We show that to conserve the USR result Eq. (3.17) the transition must be **fast** in that it completes in a small fraction of an e -fold [see Eq. (3.45)].

We conclude that the large f_{NL} of USR will only be preserved if the transition to SR is both large and fast. For all other cases, the enhancement to the local power spectrum

$$\frac{\Delta P_\zeta}{P_\zeta} \ll 1, \tag{3.23}$$

and so squeezed non-Gaussianity in transient USR does not generally affect the conclusions on PBH formation.

3.2.1 Slow–Small Transition: Inflection-Point Inflation

Inflection-point inflation is characterized by a potential which supports a slow-roll phase when CMB scales exit the horizon followed by a slow-roll violation and subsequent ultra-

slow-roll phase which enhances the power spectrum at small scales. This USR phase is generally unstable and lasts just a few e -folds before the inflaton loses enough kinetic energy to lock onto the attractor solution of the potential and slow-roll inflation resumes [76]. We call this transition **slow** because the inflaton kinetic energy decreases monotonically to the slow-roll value, and **small** because the potential slow-roll parameters on the attractor are comparable to the kinetic energy at the end of the USR phase.

We consider an inflection potential of the form explored in Ref. [8] following Ref. [75],

$$V(\phi) = \frac{\lambda v^4 x^2 (6 - 4ax + 3x^2)}{12 (1 + bx^2)^2}, \quad (3.24)$$

where $x = \phi/v$. We study this model with the parameters

$$\{a, b - 1, \lambda, v\} = \{3/2, 4 \times 10^{-5}, 7 \times 10^{-8}, 0.658\}. \quad (3.25)$$

In terms of the auxiliary variables of Refs. [8, 75], this model has

$$\{\beta, \Delta N_{\text{SR}}\} = \{4 \times 10^{-5}, 125\}. \quad (3.26)$$

These parameters are finely tuned to significantly suppress ϵ_H after the CMB scale $k_0 = 0.05 \text{ Mpc}^{-1}$ exits the horizon 55 e -folds before the end of inflation while also preventing the inflection point from trapping the inflaton for too many e -folds. Nonetheless our qualitative results for the non-Gaussianity are not sensitive to the specific functional form of the potential nor to the parameter set above.

Note that even with fine-tuning, this model does not fit observational constraints from the CMB (e.g., [7]) because the power spectrum is too red (scalar slope $n_s = 0.91$) due to the proximity of the inflection point to CMB scales. This additional red tilt implies a larger value of ϵ at CMB scales and hence a larger relative suppression of ϵ and growth of the power

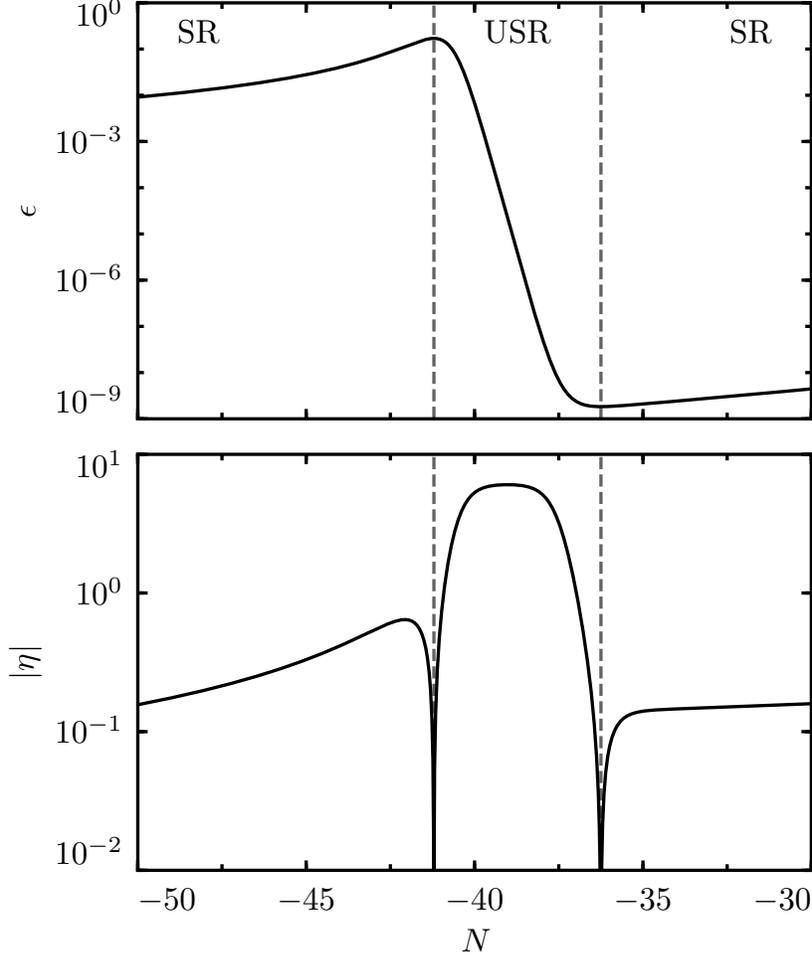


Figure 3.2: Inflection-point model background parameters $\epsilon_H \equiv -d \ln H/dN$ and $\eta \equiv d \ln \epsilon_H/dN$. η experiences two zero-crossings as ϵ reaches critical points entering and exiting the USR period, which we use to delineate the USR phase from the SR phases. Here the transient period lasts for ~ 5 e -folds, but only achieves $\eta \simeq -6$ for a shorter period.

spectrum during the USR phase. Without this enhancement, the inflection model falls far short of forming PBHs as the dark matter [8] and so we choose these parameters to study whether models on the threshold of forming sufficient PBHs for Gaussian fluctuations can be made to do so through non-Gaussianity in the model.

Along the inflaton trajectory, the potential (3.24) has a single inflection point, where $d^2V/d\phi^2 = 0$ is satisfied, between two close points where $dV/d\phi = 0$. In this region the slope of the potential is tiny, and hence the USR condition $|dV/d\phi| \ll |\phi'|H^2$ is satisfied briefly, after which slow roll quickly resumes. The evolution of the slow-roll parameters ϵ_H and η in

this model is shown in Fig. 3.2. The model exhibits a transient period where $\epsilon_H \propto a^{-6}$ and thus the USR result $\eta \simeq -6$ is temporarily achieved.

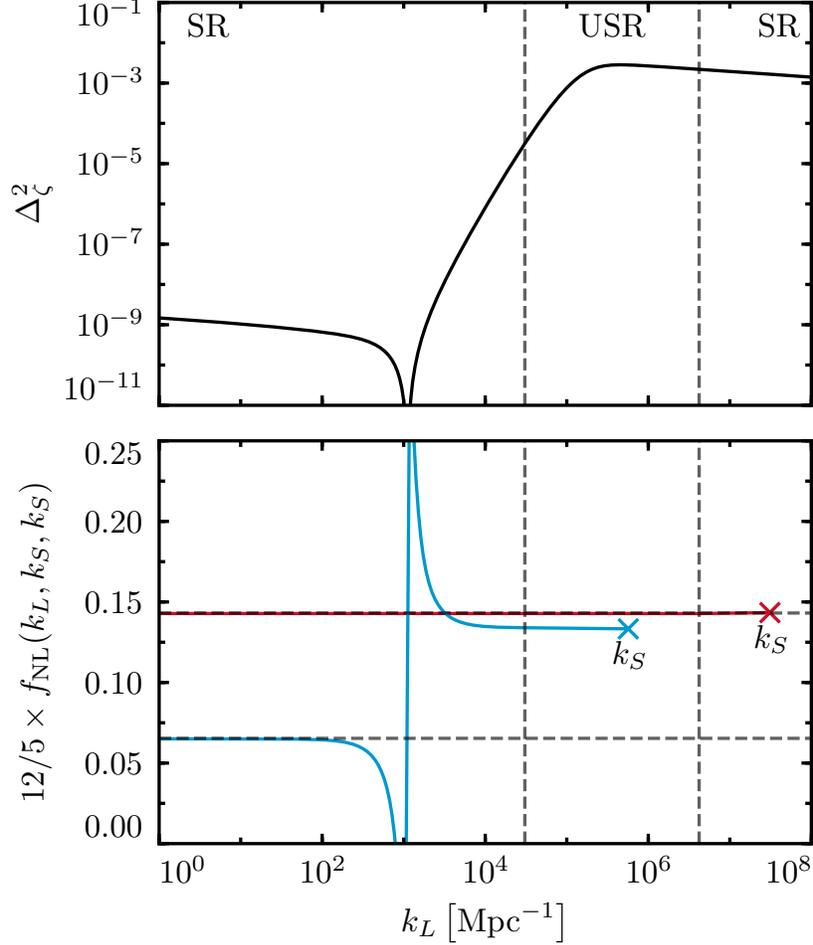


Figure 3.3: Inflection-point model power spectrum Δ_ζ^2 and non-Gaussianity parameter $f_{\text{NL}}(k_L, k_S, k_S)$. The vertical dashed lines delineate modes which cross the horizon during the SR and USR periods (see Fig. 3.2). For f_{NL} , \times 's denote values for $k_S > k_L$ with blue a mode that exits the horizon during USR and red after USR. The horizontal dashed lines denote the consistency relation expectation for the two modes and the spike in f_{NL} reflects the near zero in Δ_ζ^2 rather than a large bispectrum (see §3.2.1 for further discussion).

The upper panel of Fig. 3.3 shows the power spectrum Δ_ζ^2 produced by inflection-point inflation with the potential (3.24), computed by numerically solving the Mukhanov-Sasaki equation,

$$\frac{1}{a^2 \epsilon} \frac{d}{ds} \left(a^2 \epsilon \frac{d\zeta_k}{ds} \right) + k^2 \zeta_k = 0, \quad (3.27)$$

where $s \equiv \int_t^{t_{\text{end}}} dt/a$, with Bunch-Davies initial conditions at $ks \gg 1$ of the form

$$\zeta_k^0 = \frac{1}{2a\sqrt{k\epsilon}} \left(1 + \frac{i}{ks} \right) e^{iks}. \quad (3.28)$$

Modes which exit the horizon well before the USR phase do not grow outside the horizon and their power spectrum satisfies the slow-roll result (3.3). Modes which exit the horizon shortly before the USR phase do, however, grow outside the horizon leading to the rise before peak power in the USR phase.

This behavior can be understood in more detail from the exact, but formal, solution of the Mukhanov-Sasaki equation (3.27) [77]

$$\zeta' = -\frac{1}{a^3 \epsilon_H H} \left[\int \frac{da}{a} a^3 \left(\frac{k}{aH} \right)^2 (\epsilon_H H) \zeta + \text{const.} \right]. \quad (3.29)$$

In the SR phase, ϵ_H is roughly constant and the growing integrand provides the leading contribution

$$\zeta' \simeq - \left(\frac{k}{aH} \right)^2 \zeta, \quad (\text{SR}), \quad (3.30)$$

and hence the curvature perturbation $\zeta \propto e^{\frac{1}{2} \left(\frac{k}{aH} \right)^2}$ freezes out to a constant for $k/(aH) \ll 1$ as in Eq. (3.15). On the other hand, in the USR phase, since $\zeta \propto a^3$ outside the horizon from (3.15), it immediately holds that

$$\zeta' \propto a^3, \quad (\text{USR}), \quad (3.31)$$

outside the horizon. One can also see that (3.31) is consistent with (3.29) as follows. With $\zeta \propto a^3$, the integral in (3.29) acts as $\simeq \int d \ln a a^{-2}$, which is dominated by early times and hence converges to a constant, whereas the prefactor grows as $\propto a^3$, resulting in (3.31). Thus for a mode which spends N_{SR} e -folds outside the horizon in slow roll, it takes $N_{\text{USR}} = 2/3 \times N_{\text{SR}}$ e -folds of USR inflation to raise ζ' back to order unity. Therefore at a fixed duration N_{USR} of USR inflation, modes which exit the horizon more than $3/2 \times N_{\text{USR}}$ e -folds before USR remain constant while modes which exit within $3/2 \times N_{\text{USR}}$ grow outside the horizon. After the USR phase, modes freeze in and the smooth change in the slope of the potential assures a slow increase in ϵ and a smooth transition of the power spectrum to the final SR phase.

The power spectrum shown in the upper panel of Fig. 3.3 exhibits a near-zero minimum $\Delta_\zeta^2 \sim 8 \times 10^{-16}$. Similar behavior occurs in other models in which the growing mode overtakes the constant mode (see, e.g., Refs. [78, 79, 80] and §3.2.3). This phenomenon can also be understood in detail from the formal solution Eq. (3.29), in which it can be seen that in slow roll the superhorizon mode approaches its slow-roll freezeout value with decreasing amplitude, i.e. with

$$\text{Arg} \left[\frac{\zeta'}{\zeta} \right] = \pi + \mathcal{O} \left(\frac{k}{aH} \right), \quad (\text{SR}), \quad (3.32)$$

where the order of the correction follows from using the approximate SR form (3.28) in (3.29).

While $\zeta \propto a^3$ in the USR superhorizon limit, at the onset of USR, the curvature perturbation must reach this limit from the SR side. Let $a = a_*$ at the onset of the USR phase, then the curvature evolves as

$$\zeta' = \zeta' \Big|_{a_*} \times \left(\frac{a}{a_*} \right)^3, \quad (\text{USR}), \quad (3.33)$$

with the boundary condition $\zeta' \Big|_{a_*}$ given approximately by the SR solution for a smooth

transition. Given the relative sign in the leading order SR expression (3.30), this represents an increase in the decay rate of $|\zeta|$ and thus before modes can grow as $\zeta \propto a^3$ in USR they must reverse sign.

There is a mode which experiences just enough evolution outside the horizon by the end of the USR phase to go from its freezeout value to near-zero. The corresponding value of the power spectrum at the minimum is determined by the small out-of-phase component, i.e. how close Eq. (3.32) is to π and therefore how far outside the horizon this mode is when USR begins. Thus the longer the USR phase is, the deeper the minimum is.

Modes which exit after this minimum are dominated by their superhorizon growth, and as modes exit the horizon closer to the USR phase they grow for a longer period and thus the power spectrum grows with increasing k . While a prolonged USR phase leads to a constant Δ_ζ^2 for modes which exit the horizon during USR [cf. Eq. (3.16)], the inflection model touches the $\eta \simeq -6$ phase only briefly and the power spectrum therefore exhibits a peak $\Delta_\zeta^2(k_{\text{Peak}}) = 2.8 \times 10^{-3}$. This peak falls a factor of a few short of the value $\Delta_\zeta^2 \sim 10^{-2}$ required for PBHs to form all the dark matter (see §3.1.1 and note that a model with the right tilt at CMB scales must fall much further short of this requirement [8]). After the USR phase, the model returns to the slow-roll attractor and Δ_ζ^2 is once more described by Eq. (3.3).

It is now interesting to ask whether the power spectrum of the upper panel of Fig. 3.3 can be locally enhanced by a factor of a few to exceed the threshold (3.2) for PBH dark matter. According to Eqs. (3.4) and (3.6), for the power spectrum at a short-wavelength scale k_S to be significantly enhanced, we require a large long-wavelength perturbation ζ_L and a large correlation f_{NL} .

In the lower panel of Fig. 3.3, we plot $f_{\text{NL}}(k_L, k_S, k_S)$ as a function of the long-wavelength mode k_L for two different values of the short-wavelength mode k_S . The red upper curve shows f_{NL} for a short-wavelength mode which exits the horizon after the end of USR, while

the blue lower curve shows f_{NL} for a short-wavelength mode which exits the horizon during USR. The upper and lower horizontal dashed lines show the consistency relation expectation f_{NL} in the limit $k_L/k_S \rightarrow 0$.

The numerically computed bispectrum for a short-wavelength mode which exits the horizon after USR, the red upper curve, agrees with the consistency relation. In other words, the short-wavelength perturbation ζ_S retains no memory that, while it was inside the horizon, the long-wavelength perturbation ζ_L outside the horizon grew in USR. This is because $f_{\text{NL}}(k_L, k_S, k_S)$ is set when k_S exits the horizon and the modes ζ_L are already frozen at this time. Because the transformation of the bispectrum to $\overline{\text{FNC}}$ involves a subtraction of the consistency relation component, Eq. (3.11), we conclude that short-wavelength modes which exit after the USR phase show no response to long-wavelength modes in local coordinates and therefore no enhancement of local PBH abundance.

For a short-wavelength mode which exits the horizon during USR, the blue curve, the above logic does not hold. The numerically computed bispectrum $f_{\text{NL}}(k_L, k_S, k_S)$ does not agree with the consistency relation when ζ_L grows outside the horizon. For such triangles, $12/5 \times f_{\text{NL}} \simeq 0.13$ while the consistency relation predicts $12/5 \times f_{\text{NL}} \simeq 0.065$.

Conversely, for the frozen ζ_L modes that correspond to modes that exited the horizon well before USR, the consistency relation for f_{NL} does hold. This is a successful test of our numerical computation, since in this limit the long-wavelength mode remains constant outside the horizon and just shifts the local coordinates for the small-wavelength mode along the background trajectory.

Fig. 3.3 also shows that when k_S exits the horizon during USR, the near-zero of Δ_ζ^2 induces a feature on $f_{\text{NL}}(k_L, k_S, k_S)$. This is due to the division by the power spectrum in the definition of f_{NL} , Eq. (3.7). In particular, when k_S exits the horizon ζ_L has not yet reached its final (tiny) value set at the end of USR and thus a non-zero bispectrum $B_\zeta(k_L, k_S, k_S)$ can be obtained. After the end of USR, ζ_L is very small and thus f_{NL} is amplified. The

physical effect of this feature is negligible since, to obtain the power spectrum response, f_{NL} should be multiplied by ζ_L , which has a minimum at this feature.

More generally, the USR phase does enhance f_{NL} relative to the consistency relation value. Hence the non-Gaussianity in $\overline{\text{FNC}}$, \bar{f}_{NL} , is non-zero. There is therefore an enhancement of the PBH abundances due to squeezed non-Gaussianity, which is not the case in single-field inflation on the attractor. However, for the k_S shown in Fig. 3.3 in blue, both f_{NL} and $d \ln \Delta_\zeta^2 / d \ln k$ are so small that, once multiplied by $\zeta_L \ll 1$, the position-dependent effect on ζ_S is insignificant.

Quantitatively, we can summarize the PBH abundance enhancement in this model by choosing k_S and k_L in the USR phase, where f_{NL} is nearly constant. In particular, to eliminate the tilt-dependent coordinate effects on the abundance and to maximize the Gaussian part of the power spectrum, we can choose $k_S = k_L = k_{\text{Peak}}$. This triangle is not squeezed but since in USR f_{NL} is the same for all triangle shapes, this triangle does serve as a summary statistic for local non-Gaussianity in inflection-point inflation.

Doing so, we compute numerically that $12/5 \times f_{\text{NL}}(k_{\text{Peak}}, k_{\text{Peak}}, k_{\text{Peak}}) = 0.13$. To obtain the response, we set ζ_L to the peak value $\zeta^{\text{RMS}}(k_{\text{Peak}}) = \sqrt{\Delta_\zeta^2(k_{\text{Peak}})} \simeq 0.05$. Squeezed non-Gaussianity can therefore enhance the local power spectrum by a factor of at most ~ 0.006 and so in inflection-point inflation, its ability to enhance the local power spectrum is negligible,

$$\frac{\Delta P_\zeta}{P_\zeta} \ll 1. \quad (3.34)$$

In this model, we do not recover the USR squeezed limit result $12/5 \times f_{\text{NL}} \simeq 6$ and therefore do not enhance the small-scale power spectrum by an order-unity quantity. This is a reflection of the analytic result of Ref. [71] that transitions from a USR phase to a SR phase which are monotonically decreasing in the field velocity suppress the USR non-Gaussianity, and similar results were found numerically in Ref. [81].

In the following sections, we will show that this suppression of non-Gaussianity is generic

to transition models, except for the special case where the transition is both fast and large.

3.2.2 *Infinitely Fast Transitions*

Infinitely fast transitions from USR to SR were considered in Refs. [71] and [76], and Ref. [71] established analytically that the final level of non-Gaussianity is sensitive to the way USR is exited.

Specifically Ref. [71] considered the case where a pure ultra-slow-roll potential is joined to a slow-roll potential V_{SR} at a field position which we label ϕ_2 for ease of generalization later:

$$V(\phi) = \begin{cases} V_{\text{SR}}(\phi_2), & \phi > \phi_2 \quad (\text{USR}) \\ V_{\text{SR}}(\phi), & \phi \leq \phi_2 \quad (\text{SR}) \end{cases} \quad (3.35)$$

and hence the potential has an infinitely sharp discontinuity in slope at ϕ_2 . We call this an infinitely fast transition from USR to SR because the inflaton rolls over this discontinuity instantaneously. $V_{\text{SR}}(\phi)$ can be characterized in general by the potential slow-roll parameters for $\phi \leq \phi_2$

$$\epsilon_V \equiv \frac{1}{2} \left(\frac{1}{V} \frac{dV}{d\phi} \right)^2, \quad \eta_V \equiv \frac{1}{V} \frac{d^2V}{d\phi^2}. \quad (3.36)$$

The transition can be characterized by the strictly positive amplitude parameter²

$$h \equiv \sqrt{\frac{\epsilon_V(\phi_2^-)}{\epsilon(\phi_2^+)}} \quad (3.37)$$

where

$$\phi_2^\pm \equiv \lim_{\omega \rightarrow 0} \phi_2 \pm \omega, \quad (3.38)$$

which is a ratio between the potential slow-roll parameter at the beginning of the SR phase and the Hubble slow-roll parameter ϵ at the end of the USR phase.

2. The h defined in (3.37) is equivalent to Ref. [71]'s $-h/6$.

If $h = 1$, the kinetic energy at the end of USR is just enough to keep the field on the attractor of the SR phase. The $h \ll 1$ limit therefore corresponds to the small transition, a monotonic transition from USR to the SR attractor where the inflaton continues to slow down before hitting the attractor and hence the power spectrum continues to evolve. Conversely for $h \gg 1$, Ref. [71] showed that the perturbations freeze out at $N(\phi_2)$. We call this a **large** transition because the inflaton instantly goes from having too much kinetic energy for the potential it evolves on to suddenly having insufficient kinetic energy for a now much steeper potential

$$h \gg 1 \implies \epsilon_V \gg \epsilon = \frac{1}{2}\phi'^2, \quad (3.39)$$

Since perturbations do not freezeout immediately for a finite value of h the final level of non-Gaussianity is not given by Eq. (3.17) but rather can be shown analytically to be [71]

$$\lim_{k_L/k_S \rightarrow 0} \frac{12}{5} f_{\text{NL}}(k_L, k_S, k_S) = 2 \frac{h(3h + \eta_V)}{(h + 1)^2}, \quad (3.40)$$

for scales k_L, k_S which cross the horizon during USR. Eq. (3.40) yields the USR result Eq. (3.17) only in the limit $h \gg 1$, and thus for infinitely fast transitions the USR non-Gaussianity is conserved only when the transition is large.

The enhancement of the local power spectrum is suppressed for small transitions, but we shall next see that it is also suppressed if the transition is not sufficiently fast. Therefore, the transition needs to be large and fast to recover the USR non-Gaussianity. In contrast, the inflection model of the previous section is an example where the transition is both small and slow.

3.2.3 *Fast/Slow–Large/Small Transitions*

In order to study in more detail the phenomenology of transient USR inflationary phases beyond the slow-small transition of §3.2.1 and the infinitely fast limit of §3.2.2, we construct

a toy inflationary model which begins in SR, enters a USR phase, and then transitions back to SR. We implement this with a potential where the slope of an otherwise linear potential makes two transitions across adjustable widths in field space

$$\frac{dV}{d\phi}(\phi) = \frac{\beta}{2} \left[1 + \tanh \left(\frac{\phi - \phi_1}{\delta_1} \right) \right] + \frac{\gamma}{2} \left[1 + \tanh \left(\frac{\phi_2 - \phi}{\delta_2} \right) \right], \quad (3.41)$$

and hence

$$\begin{aligned} V(\phi) = V_0 + \frac{\beta}{2} \left[\phi + \delta_1 \log \left\{ \cosh \left(\frac{\phi - \phi_1}{\delta_1} \right) \right\} \right] \\ + \frac{\gamma}{2} \left[\phi - \delta_2 \log \left\{ \cosh \left(\frac{\phi_2 - \phi}{\delta_2} \right) \right\} \right]. \end{aligned} \quad (3.42)$$

This potential describes three phases with finite transitions, which is a natural generalization of the model (3.35) with two phases with instant transition considered in §3.2.2. The model parameters $\{\delta_1, \delta_2\}$ and $\{\phi_1, \phi_2\}$ determine the widths and positions of two transitions, respectively. The limit $\delta_1, \delta_2 \rightarrow 0$ amounts to instant transitions, where the potential is composed of a flat plateau of amplitude V_0 for $\phi_2 < \phi < \phi_1$ in between two linear pieces of slope $\{\beta, \gamma\}$, which we set positive. By modifying these parameters, we can set the duration of the USR phase as well as the circumstances of its beginning and end. By constructing the transition in $dV/d\phi(\phi)$ rather than in $V(\phi)$ directly, we trivially obtain a monotonic potential where the field always rolls downhill.

Since this is a toy model that we introduce to illustrate the fast/slow and large/small distinction, we do not attempt to accurately fit measurements at CMB scales or to appropriately end inflation. Once all modes which we are interested in have frozen out in the latter slow-roll phase, we end inflation by hand. By adjusting the potential before and after the plateau, one could turn this toy model into a model which can fit CMB observations and produce PBHs while ending inflation gracefully without changing the conclusions we draw below.

For $\phi > \phi_1$ the potential has a positive slope $\simeq \beta$ and the inflaton follows the slow-roll attractor. In order for the inflaton to leave the slow-roll attractor and enter a USR phase, the transition must be sufficiently sharp that the inflaton enters the flat region of the potential with excess kinetic energy. Thus we fix the entry parameters $\{\beta, \phi_1, \delta_1\} = \{10^{-14}, 0, 10^{-2}\}$ to guarantee such a transition. By having inflation start on the slow-roll attractor, we are freed from having to specify initial conditions during USR.

The region $\phi_2 < \phi < \phi_1$ marks the USR phase where the potential is approximately flat. We fix the amplitude of the potential in the flat plateau $V_0 = 2 \times 10^{-14}$, chosen to ensure that USR modes in our fiducial model are still perturbative, i.e. $\Delta_\zeta^2 \lesssim 1$ for the durations of USR we consider here.

Finally, for $\phi < \phi_2$ the potential has a positive tilt $\simeq \gamma$ and the inflaton returns to the slow-roll attractor. Among the remaining parameters $\{\phi_2, \delta_2, \gamma\}$, ϕ_2 determines the duration of the USR period, and $\{\delta_2, \gamma\}$ set the circumstances of the exit from USR. ϕ_2 in particular must be very finely tuned to allow several e -folds of USR inflation while still reaching the transition point in a reasonable amount of time. The instant transition of §3.2.2 corresponds to taking $\delta_2 \rightarrow 0$ and to focusing on the inflaton behavior around $\phi = \phi_2$.

We generalize Ref. [71]’s analysis to transitions of finite width between the flat and slow-roll potentials by allowing $\delta_2 \neq 0$. We start by generalizing the definition for the start and end of the transition, Eq. (3.38). We choose the end of the transition ϕ_2^- from the potential through

$$\phi_2^- \equiv \phi_2 - 2\delta_2. \quad (3.43)$$

The beginning of the transition, ϕ_2^+ , is not simply $\phi_2 + 2\delta_2$ since the USR phase persists while $\epsilon_V \ll \epsilon$. Instead, we choose to define the beginning of the transition through the deviation from the USR analytic solution,

$$1 - \frac{\phi'_{\text{USR}}}{\phi'} \Big|_{\phi_2^+} = 0.05 \times \left(1 - \frac{\phi'_{\text{USR}}}{\phi'} \Big|_{\phi_2^-} \right), \quad (3.44)$$

where ϕ'_{USR} is the analytic solution for the field velocity in USR, and ϕ' is the actual field velocity, which is evaluated numerically. By computing the field velocity deviation relative to the change at the end of the transition ϕ_2^- we guarantee that ϕ_2^+ can be defined even for small and fast transitions.

In other words, ϕ_2^- is roughly where the potential completes its transition, and ϕ_2^+ is roughly where the field velocity begins to leave the USR solution. While the specific criteria chosen here are arbitrary, they are useful for classifying transition regimes and in the $\delta_2 \rightarrow 0$ limit the choices here return the limit Eq. (3.38) up to percent-level factors.

From these definitions for ϕ_2^+ and ϕ_2^- we compute h by evaluation of Eq. (3.37) and we quantify the duration of the transition from USR inflation to the beginning of the relaxation process

$$d_N \equiv N(\phi_2^-) - N(\phi_2^+). \quad (3.45)$$

The situation of §3.2.2 corresponds to the limit $d_N \rightarrow 0$ and we now generalize this result by exploring the impact of the duration d_N on the resultant non-Gaussianity. In Fig. 3.4, we show the background parameters ϵ_H and η for two models with a large transition $h = 2.5$, one fast ($d_N = 5 \times 10^{-3}$) and the other slow ($d_N = 1$). For these models γ is fixed at 6×10^{-21} , while the fast transition has $\{\phi_2, \delta_2\} = \{-0.1580281699, 2.12 \times 10^{-10}\}$ and the slow transition has $\{\phi_2, \delta_2\} = \{-0.1580282187, 3.6 \times 10^{-8}\}$. Notice the amount of fine-tuning in ϕ_2 required to achieve subpercent-level control of the transition amplitude and duration.

Defining $n \equiv N - N(\phi_2^-)$ as the positive increasing number of e -folds elapsed since the potential transition, in the fast transition limit $d_N \rightarrow 0$ the solution for η after the transition point behaves according to the analytic result [71]

$$\eta(n > 0) = \frac{6(h-1)}{1+h(e^{3n}-1)}. \quad (3.46)$$

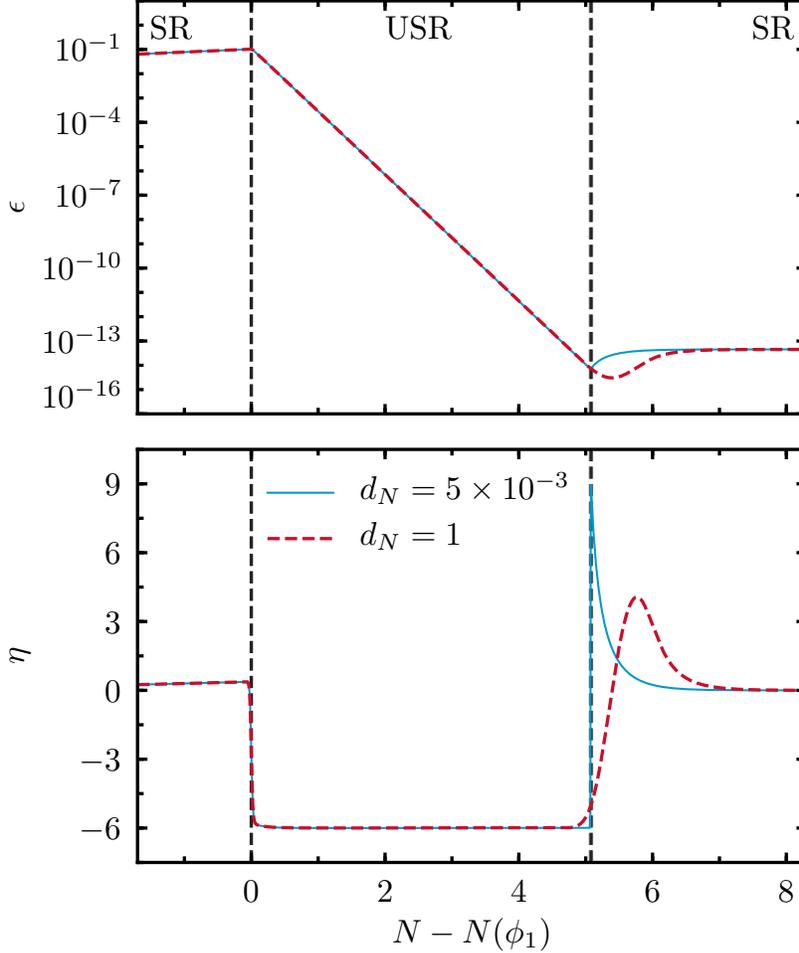


Figure 3.4: Transition model background parameters $\epsilon_H \equiv -d \ln H / dN$ and $\eta \equiv d \ln \epsilon_H / dN$ for a fast vs. slow transition, denoted by the e -fold width d_N , with a fixed large transition ($h = 2.5$). The vertical lines mark $N(\phi_1)$ and $N(\phi_2^+)$. The specific parameter choices used for these models are described in §3.2.3.

This is reflected in the behavior of the blue solid curve in Fig. 3.4, which behaves as Eq. (3.46) up to a $\sim 0.3\%$ difference in the h parameter as defined here compared to the h parameter in the exact $\delta_2 \rightarrow 0$ limit.

On the other hand, in the case where d_N is large, the red dashed curve of Fig. 3.4, the numerical solution for η deviates significantly from this analytic form. This can be understood by Taylor expanding Eq. (3.46) around the transition point $n = 0$,

$$\eta(n > 0) = 6(h - 1)(1 - 3hn) + \mathcal{O}(n^2), \quad (3.47)$$

from which we can see that after ϕ_2^- , η evolves on a timescale $n \sim h^{-1}$. Thus if the transition timescale d_N is larger than this timescale, the evolution of η will differ from the analytic solution (3.46).

The behavior of η is important because it controls the freezeout of perturbations through the evolution of ϵ in Eq. (3.29), and it comes directly into the source of squeezed non-Gaussianity in the in-in formalism, where the cubic Lagrangian relevant in the squeezed limit is

$$\begin{aligned} \mathcal{L}_3 \equiv & a^3 \epsilon_H \frac{d}{dt} \left(\epsilon_H + \frac{\eta}{2} \right) \zeta^2 \dot{\zeta} - \frac{d}{dt} \left[a^3 \epsilon_H \left(\epsilon_H + \frac{\eta}{2} \right) \zeta^2 \dot{\zeta} \right] \\ & + \epsilon_H \zeta (\mathcal{H}_2 + 2\mathcal{L}_2) \\ & - \frac{d}{dt} \left[\frac{a^3 \epsilon_H}{H} \zeta \dot{\zeta}^2 + a^3 \frac{\epsilon_H^2}{2H} \dot{\zeta} \partial_a \zeta \partial_a (\partial^{-2} \dot{\zeta}) \right]. \end{aligned} \quad (3.48)$$

\mathcal{H}_2 here is the quadratic Hamiltonian density

$$\mathcal{H}_2 = a^3 \epsilon_H \left[\dot{\zeta}^2 + \frac{1}{a^2} (\partial \zeta)^2 \right]. \quad (3.49)$$

Therefore, the timescale d_N plays an important role in changing the non-Gaussianity produced in USR.

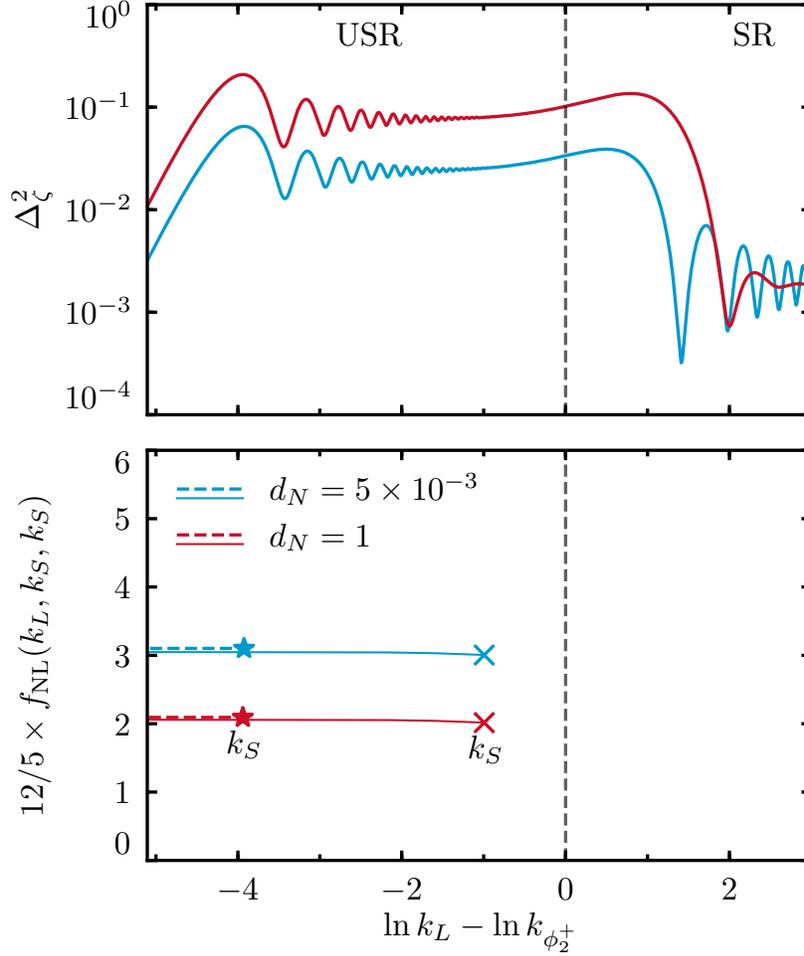


Figure 3.5: Transition model power spectrum Δ_ζ^2 and non-Gaussianity parameter $f_{\text{NL}}(k_L, k_S, k_S)$ for the large-fast (blue) and large-slow (red) transition models of Fig. 3.4. The left edge corresponds roughly to a mode which exits the horizon at the beginning of USR. Conventions for displaying f_{NL} are the same as in Fig. 3.3 except both k_S values for each model cross the horizon before the end of USR (vertical line, $k_{\phi_2^+}$) (see §3.2.3 for further discussion).

The upper panel of Fig. 3.5 shows the power spectra for the same large-fast and large-slow models as Fig. 3.4. Once more, the red upper curve shows the slow transition, while the blue lower curve shows the fast transition. The vertical dashed line shows a mode which exits the horizon at the onset of the transition $N(\phi_2^+)$.

The power spectra in these models show many of the same features as the inflection-point power spectrum, Fig. 3.3, and therefore we focus on the USR and transition regions of the plot, where unlike the inflection model these models have power spectrum plateaus for modes which exit during USR. This plateau is modulated by small oscillations sourced by the first feature in the potential. The USR to SR transition feature in the potential also induces power spectrum oscillations, with the fast transition model having more oscillations due to the sharper source. The slow transition model has a larger power spectrum than the fast transition as ϵ_H reaches a lower level in this model (see Fig. 3.4) and thus the modes grow more.

The lower panel of Fig. 3.5 shows the non-Gaussianity for these models for two different values of the short-wavelength mode, one marked with a cross and the other a star and both exiting during the USR phase, as a function of the long-wavelength mode, curves with correspondingly solid and dashed lines. All triangles yield the same value of f_{NL} when the legs exit during the USR phase up to corrections of order $k_S/k_{\phi_2^+}$, consistent with the result in the inflection model §3.2.1 and the exact USR result of §3.1.2. However, for neither model does the level of the non-Gaussianity agree with the analytic result for USR Eq. (3.17), $5f_{\text{NL}}/12 = 6$. This is the result of Ref. [71], that the residual level of f_{NL} depends on the value of the transition parameter h , and in particular the fast transition model yields the result expected from Eq. (3.40) for a transition with $h = 2.5$, $f_{\text{NL}} \simeq 3.1$. However, the slow model has the same $h = 2.5$ as the fast model, yet a smaller value of f_{NL} . This is due to the slow nature of the transition in the $d_N = 1$ model. Fast transitions yield Eq. (3.40), while slow transitions suppress the non-Gaussianity.

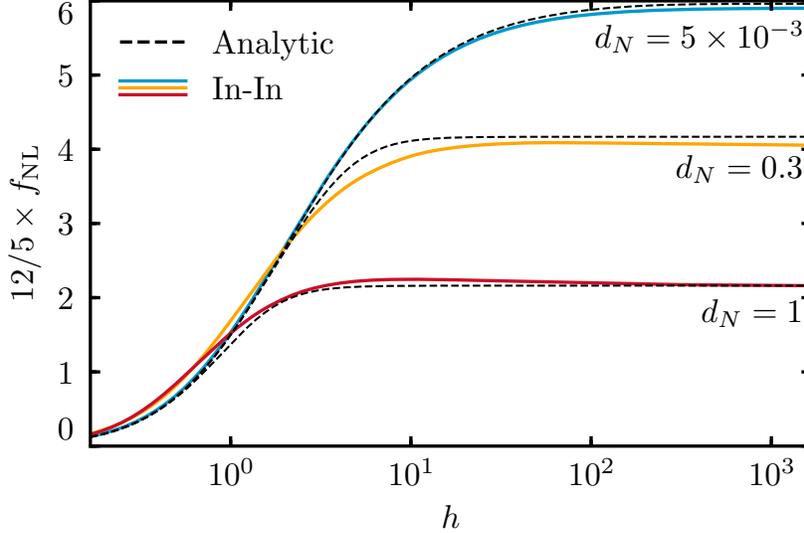


Figure 3.6: Transition model non-Gaussianity parameter f_{NL} as a function of amplitude h for various e -fold widths d_N with both k_L and k_S exiting the horizon during USR. Colored lines show the numerical in-in computation and dashed black lines show the calibrated analytic prediction from Eqs. (3.50) and (3.51). To reach the USR result $12f_{\text{NL}}/5 = 6$, the transition must be large $h \gg 1$ and fast $d_N \ll 1$ (see §3.2.3 for further discussion).

We model this effect with the ansatz that a given transition length d_N sets an upper bound to the transition amplitude, independent of h . We can define an effective transition amplitude

$$h_{\text{eff}} \equiv \left[\left(\frac{1.5}{d_N} \right)^{-3} + h^{-3} \right]^{-1/3}, \quad (3.50)$$

where the exponent serves merely to interpolate between the two limits and the factor of 1.5 comes from calibrating the results to the form of (3.40)

$$\frac{12}{5} f_{\text{NL}}(h_{\text{eff}}) = 2 \frac{h_{\text{eff}}(3h_{\text{eff}} + \eta_V)}{(h_{\text{eff}} + 1)^2}, \quad (3.51)$$

where in this toy model with a linear SR potential we set $\eta_V = 0$.

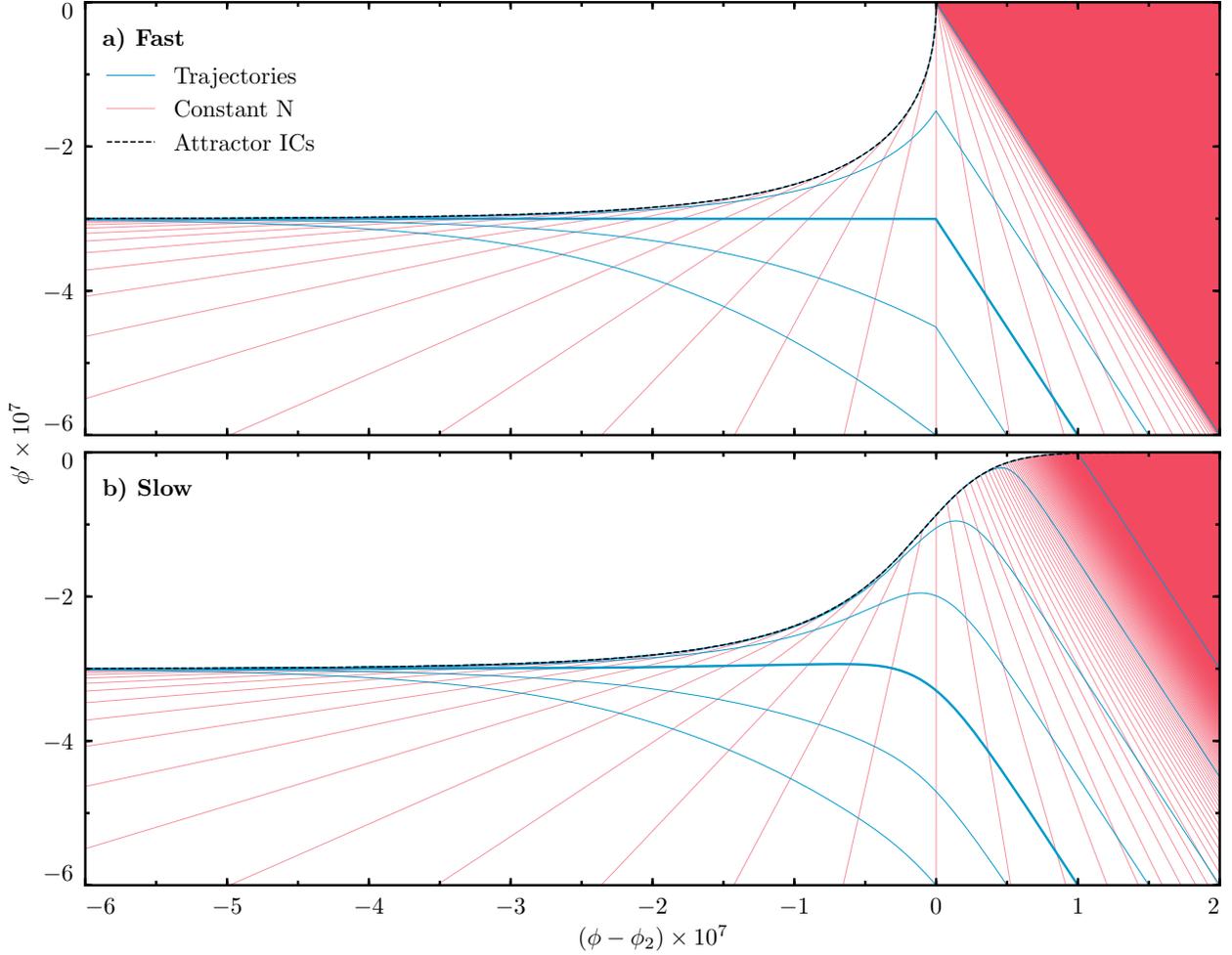


Figure 3.7: Phase-space diagram for the USR to SR transition models with model trajectories (blue lines) and constant e -fold surfaces (red lines) relative to the transition at ϕ_2 . The top panel a) has $\delta_2 = 2.12 \times 10^{-10}$, such that most trajectories correspond to fast transitions, while the bottom panel b) has $\delta_2 = 3.6 \times 10^{-8}$, such that most trajectories correspond to slow transitions. In each panel, higher trajectories represent larger transitions, with $h = 1$ as the large-small dividing line (thick blue). Constant N surfaces become space-filling in the top right corners. A large f_{NL} requires a fast-large transition as can be visualized by δN , the change in e -folds given a shift in the initial field $\delta\phi_i$ that takes the local background to a new trajectory (see §3.2.3 for further discussion).

In Fig. 3.6, we compare the ansatz (3.50) coupled with the analytic formula (3.51) (black dashed curves) to the full numerical in-in computation of the bispectrum (colored solid curves) for modes that exit the horizon during USR as a function of the transition amplitude h for three different transition speeds d_N . Every point along these lines corresponds to a different set of parameters for our toy model.

The analytic formula (3.50) agrees well with the in-in computation for all values of h and d_N . For small-slow transitions (small h and large d_N), where f_{NL} is small and becomes proportional to the potential slow-roll parameters on the attractor, there is a slight difference between the numerical and analytic f_{NL} . This we attribute to small differences in the value of h as defined for the infinitely fast transitions and as defined for slow transitions since the errors decrease for smaller values of d_N . For large transitions $h \gg 1$, where f_{NL} is largest, the analytic ansatz produces a slight overestimate of f_{NL} as the transitions become faster. This is due to nonlinearities between the true h_{eff} and $1/d_N$ in the large h limit which our ansatz does not model. Since these differences are minor, we conclude from the analytic formula that to produce a large level of non-Gaussianity after the transition from USR to SR requires $h_{\text{eff}} \gg 1$ and thus the transition must be large, $h \gg 1$, and fast, $d_N \ll 1$.

Just as in the SR (§3.1.1) and exact USR case (§3.1.2), the effect on f_{NL} of the transition from USR to SR can be understood visually from the way phase-space trajectories intersect constant N surfaces. Fig. 3.7a shows the phase-space trajectories (blue lines) and constant N surfaces (red lines) for a narrow $\delta_2 = 2.12 \times 10^{-10}$ and therefore faster transitions, such that the fast model of Fig. 3.4 and Fig. 3.5 corresponds to a trajectory in this space. Fig. 3.7b shows the phase space for a wider $\delta_2 = 3.6 \times 10^{-8}$ and therefore slower transitions, and the slow model of Fig. 3.4 and Fig. 3.5 evolves through this space.

Trajectories near the top of each panel have the inflaton speed up after the transition and hence have a large h , with the large-small dividing line of $h = 1$ denoted by thick blue lines. Notice also that the union of the panels of Fig. 3.1 gives the limit of infinitely fast transitions,

with the exception that here constant N surfaces are plotted relative to the transition feature $N(\phi_2)$ rather than the end of inflation $N = 0$. Trajectories are evenly spaced in ϕ at the point where they cross the bottom edge at $\phi' = -6$ in a range that reflects a reasonable amount of USR e -folds as we describe next.

Due to the smooth nature of the potential (3.42), for any finite $\phi > \phi_2$ the potential slope $dV/d\phi$ has a finite positive value. Thus, unlike in the exact USR case (Fig. 3.1b) or the infinitely fast case, all trajectories with any finite ϕ' for $\phi > \phi_2$ will eventually cross ϕ_2 .

Black dashed curves in Fig. 3.7 depict the envelope of such trajectories, neglecting stochastic effects, and correspond to initial conditions on the attractor on the very nearly flat potential. Consequently, the constant e -fold surfaces become increasingly tightly packed and eventually space-filling, in contrast to the empty upper right triangle in Fig. 3.1b. We choose not to continue showing trajectories which take such large numbers of e -folds to traverse the nearly flat plateau of the potential.

By the same δN arguments of §3.1.2 we can immediately see from these phase spaces why a large ($h \gg 1$) transition is necessary to conserve the USR non-Gaussianity. Here δN refers to the change in the total number of e -folds elapsed to a fixed field position on the SR side for a shift in the initial field position $\delta\phi_i$ on the USR side which then shifts the whole trajectory. Note that δN combines the change from the USR and SR sides.

Let us first consider the fast case in the top panel. Around a central trajectory with large h (upper trajectories), the crossing rate $\partial N/\partial\phi_i$ is strongly asymmetric to the sign of $\delta\phi$, i.e. there is a large second derivative $\partial^2 N/\partial\phi_i^2$ and hence a large f_{NL} according to Eq. (3.18). This is due to the much larger contribution to the rate of surfaces crossed in the USR side where the asymmetry is larger than the SR side where the asymmetry is small.

On the other hand, the asymmetry around a small h trajectory (lower trajectories) is small and therefore the non-Gaussianity is small. This is due to the smaller contribution to $\partial N/\partial\phi_i$ on the USR side relative to the SR side. In other words the power spectrum

continues to grow on the SR side at small h , which suppresses the non-Gaussianity from the USR side.

By comparing the fast and slow cases, we can visually see that the transition duration sets an effective maximum transition amplitude h_{eff} . Above a certain value of h , the trajectory joins the slow-roll attractor before the transition and therefore will have a highly suppressed non-Gaussianity comparable to the small-slow transition of §3.2.1.

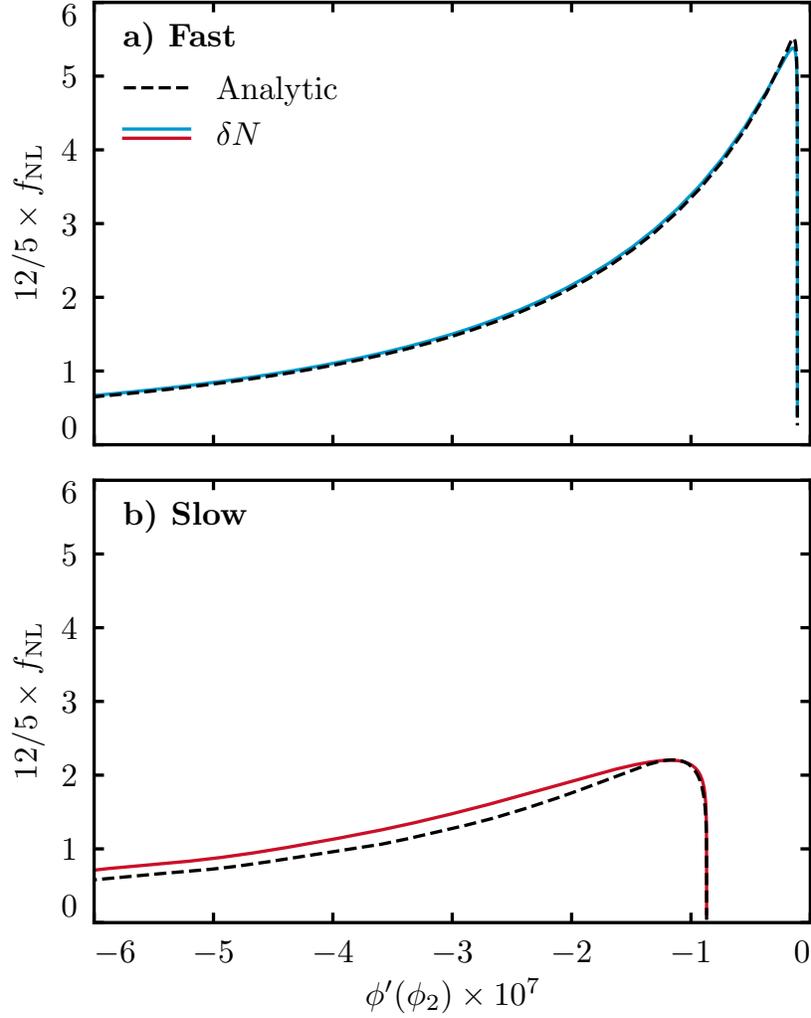


Figure 3.8: Non-Gaussianity for trajectories in the corresponding phase spaces of Fig. 3.7 computed via the δN formula (3.18) compared to our calibrated analytic result Eqs. (3.50) and (3.51). The horizontal axis here corresponds to the vertical axis of Fig. 3.7. Only for fast and large transitions does the transition model reproduce the USR result, and for fixed δ_2 there is a maximum value of f_{NL} attainable.

In Fig. 3.8 we formalize these heuristic arguments by computing the δN formula (3.18) for different trajectories in these phase spaces. We organize the trajectories by their velocity at ϕ_2 , and thus the horizontal axis of Fig. 3.8 corresponds to the vertical axis of the Fig. 3.7 at the transition point. We then also compare the δN result to our analytic expectation for f_{NL} , Eqs. (3.50) and (3.51).

The δN computation cross-validates our analytic formula which was calibrated to the in-in calculations, showing excellent agreement for all methods of computation across the fast-slow and large-small transition space. The δN computation thus also confirms that a large f_{NL} requires a large-fast transition.

For large $\phi'(\phi_2)$, after the peak non-Gaussianity, there is a sharp cliff in Fig. 3.8 beyond which f_{NL} becomes suppressed. This cliff corresponds to trajectories which reach the attractor before the transition, thus inhabiting the space-filling regions of Fig. 3.7, and the edge of the cliff is given by the black dashed envelope of Fig. 3.7. These trajectories are slow transitions even though they have the same narrow field space width δ_2 and appear in the ‘fast’ panel. These cases behave in the same way as those in the ‘slow’ panel once the duration of the transition in e -folds d_N is accounted for.

Of course even for a large and fast transition, for which $5/12 f_{\text{NL}} \rightarrow 6$, the response of the small scale power spectrum to the long-wavelength mode is still dependent on the value of the long-wavelength mode, and in particular the argument of §3.1.2 still holds, that

$$\frac{\Delta P_\zeta}{P_\zeta} \lesssim 1, \tag{3.52}$$

unless the model already produces PBHs with a Gaussian distribution.

3.3 Discussion

Canonical slow-roll inflation cannot produce primordial black holes in a large enough quantity to be the dark matter. While perturbations do exhibit a small level of local non-Gaussianity which couples short-wavelength PBH fluctuations to the long-wavelength modes they live in and can in principle enhance local abundances at peaks of long-wavelength modes, transforming to a freely-falling coordinate system shows that locally measured PBH abundances are completely insensitive to this non-Gaussianity because it is generated by the reverse coordinate transformation to begin with.

Any confirmation that the dark matter is in the form of PBHs would rule out canonical slow-roll inflation. The only way to rescue canonical inflation would be to violate slow roll, and a phase of ultra-slow-roll inflation after CMB scales exit the horizon is the natural way to do this. We showed by gauge transformation and by the δN formalism, which can be illustrated graphically and contrasted to the SR case, that local non-Gaussianities are large in the USR phase when perturbations freeze out instantly at some fixed field position.

The same coordinate transformation machinery as in slow roll confirms that squeezed USR non-Gaussianities can locally enhance PBH abundances. However the effect is very mild, giving at most an order unity enhancement of the local power spectrum. Such enhancements can only make models that are already on the borderline of succeeding to produce PBHs as the dark matter under Gaussian assumptions actually succeed. For such cases, generally a small change in parameters that prolong the USR phase would equally well produce PBHs under Gaussian assumptions.

Even more importantly, the USR phase has to end in some way. Ref. [71] established that the non-Gaussianity is very sensitive to how this period ends using cases where the transition is infinitely fast. By exact computation in the in-in formalism and validation with the δN formalism, we mapped the entire range of possible endings to USR to show that only a small class of transitions conserves the large USR non-Gaussianity through the transition to slow

roll. These are the transitions which are fast, in that the potential exhibits a sharp feature that is traversed by the inflaton in much less than an e -fold, and large, in that the inflaton needs to gain significant velocity after transiting the feature. All other types of transitions suppress the non-Gaussianity significantly.

Note that while we have computed the full local non-Gaussianity, it is only the squeezed non-Gaussianity which we have shown has a negligible effect on whether a model produces PBH dark matter. Understanding the effect of other mode couplings requires a full knowledge of the probability distribution function of the density perturbation taking into account the contributions from all orders of non-Gaussian correlators of every shape. Work in this direction has been pursued in a variety of contexts (see, e.g., [82, 83, 84, 85, 86, 87, 88]), but it remains a challenging problem and direction for future work.

Nonetheless the conclusion that producing primordial black holes as dark matter in canonical single-field inflation requires a complicated and fine-tuned potential shape with a transient violation of slow roll is robust to the inclusion of local squeezed non-Gaussian effects.

CHAPTER 4

THE HIGGS FIELD DURING INFLATION

We have dealt in detail with single-field inflation and its generalizations, and we can now complete this thesis by studying the behavior of extra spectator fields during inflation and the conditions that must be satisfied for them to contribute to primordial black hole dark matter production, based on Ref. [4].

We will focus our study on the Standard Model Higgs field, which exhibits the unusual property that its effective potential turns over at large field values. This instability can lead to complicated behavior during and after inflation which has been claimed to lead to PBH DM production.

We will first provide a general overview how the Higgs instability could possibly lead to PBH production. We will then focus first on the Higgs' evolution during inflation and compute its spectrum of perturbations on CMB scales, which should be small relative to its value on PBH scales. We then track the Higgs and its fluctuations nonlinearly through reheating to compute whether Higgs fluctuations are converted to large curvature fluctuations. We will find that the Higgs fails on both counts and therefore does not produce PBH DM.

4.1 The Higgs Instability Mechanism for Primordial Black Holes

In this section, we review the general features of the Higgs instability mechanism for producing PBH dark matter and the principles governing the spectrum of perturbations it generates.

The Higgs field acts as a spectator during inflation, which is driven by an inflaton field, but converts its quantum fluctuations into curvature fluctuations once it decays after inflation. If these quantum fluctuations are amplified into large enough curvature fluctuations by the Higgs instability, they will form PBHs when they re-enter the horizon in the radiation

dominated epoch.

In particular, after inflation the spatial metric on a uniform total energy density hypersurface,

$$g_{ij} = a^2 e^{2\zeta} \delta_{ij}, \quad (4.1)$$

possesses a curvature perturbation ζ that can be decomposed as

$$\zeta = \left(1 - \frac{\rho'_h}{\rho'_{\text{tot}}}\right) \zeta_r + \frac{\rho'_h}{\rho'_{\text{tot}}} \zeta_h, \quad (4.2)$$

in linear theory. Here ζ_r is the curvature perturbation on hypersurfaces of uniform energy density ρ_r of reheat products from the inflaton, ζ_h is the curvature perturbation on hypersurfaces of uniform Higgs energy density ρ_h , and $\rho_{\text{tot}} \equiv \rho_r + \rho_h$ is the total energy density. Throughout primes $'$ denote a derivatives with respect to the e -folds N , where $N = 0$ marks the end of inflation.

For this mechanism to succeed observationally, the variance per logarithmic interval in k of the Fourier modefunction ζ^k ,

$$\Delta_\zeta^2(k) \equiv \frac{k^3}{2\pi^2} |\zeta^k|^2, \quad (4.3)$$

must be small on the large scales probed by anisotropies in the CMB,

$$\Delta_\zeta^2(k_{\text{CMB}}) \sim 10^{-9}, \quad (4.4)$$

while on small scales associated with primordial black holes the power spectrum must reach

$$\Delta_\zeta^2(k_{\text{PBH}}) \sim 10^{-2}, \quad (4.5)$$

so that enough regions are over the collapse threshold $\zeta_c \sim 1$ to form PBHs in sufficient

abundance to be the dark matter [8, 3]. Moreover, the large-scale modes should be sourced predominantly by a single degree of freedom in order to comply with isocurvature constraints from the CMB. The working assumption for a successful model is that inflaton perturbations lead to a ζ which dominates on CMB scales while Higgs fluctuations produce one which dominates on PBH scales. We use a superscript k to denote relations that are exclusively in Fourier space as opposed to real space quantities or linear relations that apply to both.

Although the Higgs field h is a spectator during inflation, the mechanism works by enhancing its impact on the total ζ by exploiting the unstable, unbounded nature of the Higgs potential $V(h)$ at large field values $h > h_{\text{max}}$ in the Standard Model. In particular, the effective Higgs potential at field values far larger than its electroweak vacuum expectation value can be approximated as [89, 90]

$$V(h) = \frac{1}{4}\lambda h^4, \quad (4.6)$$

with $\lambda = \lambda^{\text{SM}}$ and

$$\lambda^{\text{SM}} \simeq -b \ln \left(\frac{h^2}{h_{\text{max}}^2 \sqrt{e}} \right), \quad (4.7)$$

where h_{max} is the location of the maximum of the Higgs potential which separates the familiar metastable electroweak vacuum from the unstable region, and b controls the flatness of the potential around the maximum. h_{max} and b are computable given the parameters of the SM, and in this work, we choose to fix them at representative values $h_{\text{max}} = 4 \times 10^{12}$ GeV and $b = 0.09/(4\pi)^2$, corresponding to a top quark mass $M_t \simeq 172$ GeV, following Refs. [91, 92, 89] to facilitate comparisons. The Higgs instability exists for $M_t \gtrsim 171$ GeV [93], which includes the range from the most recent constraints by the Tevatron and the LHC [94, 95, 96, 97]. Here we have neglected an effective mass term for the Higgs generically generated by a nonminimal Higgs coupling to the Ricci scalar, since at the level expected from quantum corrections it does not change the qualitative features of the mechanism [89].

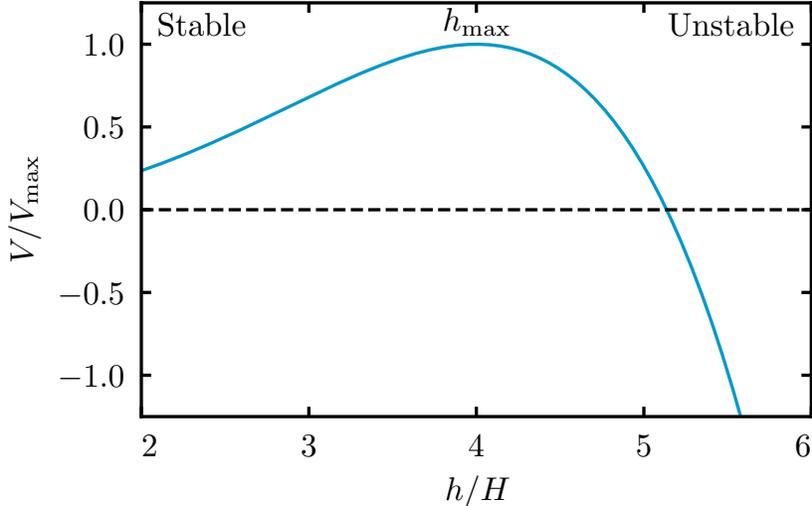


Figure 4.1: Beyond the field value h_{\max} , the Standard Model Higgs effective potential turns over and decreases rapidly, with an unbounded true vacuum of negative energy density. In the PBH scenario, the spectator Higgs rolls down the unstable potential, amplifying its stochastic field fluctuations.

We show the potential just around its maximum in Fig. 4.1, and across a wider range of scales in Fig. 4.2 on the unstable side in order to illustrate the field values that will be crucial to the Higgs instability phenomenology detailed in the following sections. Specifically, for a representative choice for the Hubble scale during inflation $H = 10^{12}$ GeV which we employ throughout for illustration, the potential maximum is at

$$h_{\max} = 4H. \quad (4.8)$$

When $h < h_{\max}$, the minimum corresponds to our familiar electroweak vacuum, while for $h > h_{\max}$, the potential decreases and is unbounded from below.

On the unstable side of the potential, the Higgs at first rolls slowly relative to the Hubble rate before accelerating as it rolls down the instability. The location where the Higgs' roll in one e -fold becomes comparable to $H/2\pi$ defines the classical roll scale h_{cl} , which we shall

define precisely in §4.2. For our parameter choices, it lies at

$$h_{\text{cl}} \simeq 8.3H. \quad (4.9)$$

At the scale k_{cl} which crosses the horizon at $N_{\text{cl}} \equiv N(h_{\text{cl}})$, the power spectrum of ζ_h at horizon crossing becomes order unity,

$$\Delta_{\zeta_h}^2(k_{\text{cl}}) \sim 1. \quad (4.10)$$

Our working assumption is that the classical-roll scale k_{cl} will be the one to produce primordial black holes, $k_{\text{cl}} = k_{\text{PBH}}$. We will show in §4.2 that this implies that the Higgs is in fact on the unstable side of the potential during all phases of inflation relevant for observation. In particular, the CMB scales left the horizon during inflation a few e -folds after our Hubble patch crossed the horizon, which we will assume is 60 e -folds before the end of inflation. At this time, the Higgs is on the unstable side of the potential at

$$h_{60} \simeq 5.8H, \quad (4.11)$$

if we take the field value at the end of inflation to be

$$h_{\text{end}} \simeq 1200H, \quad (4.12)$$

which we will see below is approximately the largest value possible. This also leads to $N_{\text{cl}} \sim -20$. If PBHs are produced on that scale then they have a small mass $M_{\text{PBH}} \simeq 10^{-15}M_{\odot}$ at formation, and after mergers and accretion could today lie in the region $M_{\text{PBH}} \simeq 10^{-12}M_{\odot}$ where all the dark matter could be in the form of PBHs [9, 98].

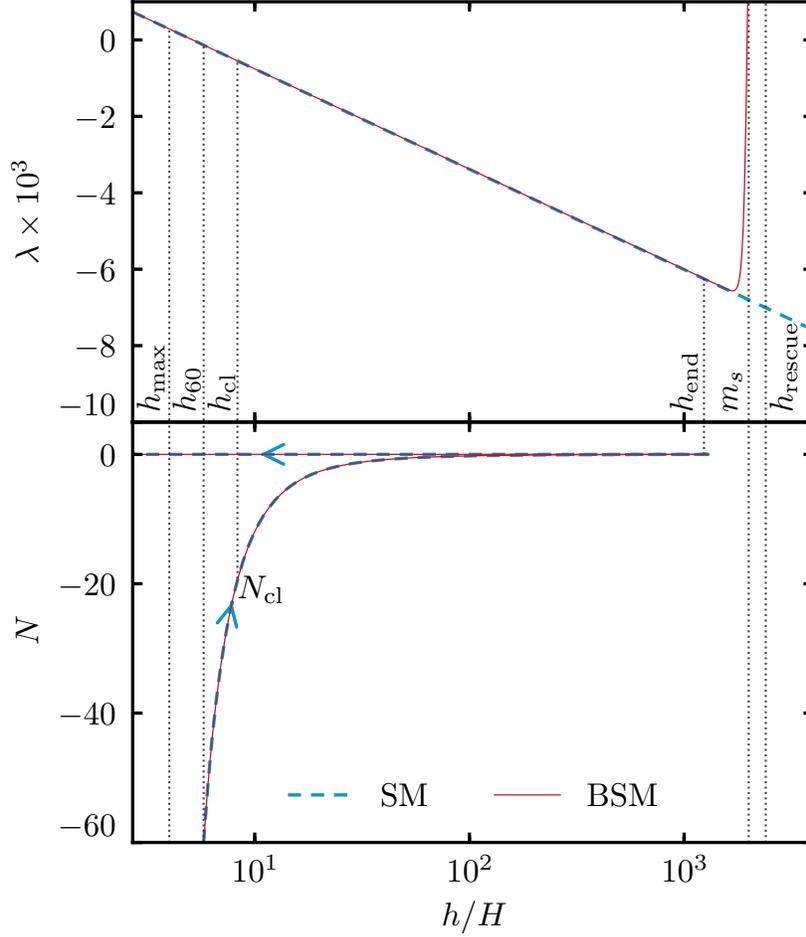


Figure 4.2: Quartic coupling λ and e -folds N corresponding to the Higgs field position h , with marked special values as computed in §4.2 and §4.3: h_{\max} , the maximum of the Higgs potential; h_{60} , its position 60 e -folds before the end of inflation; h_{cl} , its classical roll position; h_{end} , its position at the end of inflation; h_{rescue} , beyond which the SM Higgs cannot be “rescued” by reheating so that it rolls back and oscillates around the origin after inflation. The BSM Higgs adds a coupled scalar of suitable mass m_s to eliminate the runaway instability.

The Higgs continues to roll to larger field values until the end of reheating when interactions with the thermal bath lift the effective Higgs potential. If the Higgs lies within a maximum rescuable distance h_{rescue} ,

$$h_{\text{rescue}} \simeq 2400H, \quad (4.13)$$

which we compute in §4.3, then after inflation it rolls safely back to the metastable electroweak vacuum, oscillating in a roughly quadratic potential with a temperature-dependent mass $\propto T \propto 1/a$ until it decays to radiation on a uniform Hubble surface.

The behavior of Higgs perturbations through reheating is complicated. However, all relevant physical scales are at this stage far outside the horizon. So long as a gradient expansion holds, under which such perturbations can be absorbed into an approximate FLRW background for a local observer, then the curvature field after horizon crossing evolves locally, with no explicit scale-dependence [99]. Therefore the Higgs contribution to the fully nonlinear curvature field ζ on a uniform total density slice after the Higgs decays is related to the curvature ζ_h on constant Higgs slices on superhorizon scales during inflation by

$$\zeta|_{\text{decay}} = R(\zeta_h) \times \zeta_h|_{\text{inflation}}, \quad (4.14)$$

with all of the complicated physics of reheating absorbed into a local remapping $R(\zeta_h)$. More generally, this local remapping would also involve inflaton curvature fluctuations but these are statistically independent and can be calculated separately in the usual way. In linear theory, the mapping becomes a simple rescaling factor

$$R \equiv \lim_{\zeta_h \rightarrow 0} R(\zeta_h). \quad (4.15)$$

The power spectrum after Higgs decay is then related mode-by-mode to the power spectrum

during inflation in a scale-independent fashion

$$\Delta_{\zeta}^2(k)\Big|_{\text{decay}} = R^2 \Delta_{\zeta_h}^2(k)\Big|_{\text{inflation}}, \quad (4.16)$$

in which the number R encodes all of the details of reheating.

It is therefore important to emphasize here that if the Higgs instability mechanism is successful, such that $\Delta_{\zeta}^2 \sim 10^{-2}$ on PBH scales after the Higgs decays, then in linear theory R must reach at least 0.1 for the order unity Higgs perturbations (4.10) to be converted to curvature perturbations with the correct amplitude (4.5) to form sufficient PBHs and the details of how this is achieved through reheating and Higgs decay are irrelevant for the prediction of the linear power spectrum on other scales. In particular, its value on CMB scales depends solely on the inflationary $\Delta_{\zeta_h}^2(k)$. This form is controlled by the Higgs potential itself and by the evolution of the Hubble rate during inflation. Therefore, the viability with respect to CMB anisotropies of the PBH formation scenarios introduced in the literature [91, 89, 92], which all assume mode evolution can be calculated linearly through reheating, can be assessed independently of the details of the reheating model.

On the other hand, we shall see that the nonlinear nature of the Higgs instability plays an important role in the mapping between ζ_h and ζ in Eq. (4.14). Here, though the mapping remains local in that a given value of $\zeta_h(\vec{x})$ at a given position \vec{x} is mapped into a specific value of $\zeta(\vec{x})$, Fourier modes no longer evolve independently. Instead, we will compute the mapping using the nonlinear δN formalism. This mapping does depend on the specifics of how inflation ends, but is independent of physical scale. CMB scale fluctuations are in principle calculable from the spatial field $\zeta_h(\vec{x})$ determined by modes that froze out during inflation.

To illustrate these concepts, we make a few simplifying assumptions about how inflation and reheating proceed. We show in §4.2 that the most optimistic case for the scenario occurs when H is effectively constant through inflation (see Eq. (4.44)). Therefore rather

than introducing a specific inflaton potential we assume that inflation occurs at a fixed H and ends after an appropriate number of e -foldings. The constant Hubble scale during inflation H and the position h_{end} of the Higgs at the end of inflation then together control the number of e -folds between the classical-roll scale h_{cl} and the end of inflation, and therefore they control the physical scales on which PBHs are formed.

We then assume that at the end of inflation, the inflaton decays instantly into radiation and that the Higgs later also suddenly decays into radiation, as in the model proposed in Ref. [91]. Maximizing R in linear theory requires that the position of the Higgs at the end of inflation, h_{end} , is as close as possible to the maximum rescuable distance h_{rescue} . This criticality requirement motivates the various choices of scale in Eqs. (4.8)-(4.12), following Ref. [91]. Once h_{end} is set in this way, the value $H = 10^{12}$ GeV is chosen to give a certain mass scale to PBHs by fixing $N_{\text{cl}} \sim -20$.

However, evolving the Higgs on the unstable side of its potential during inflation is dangerous, and the required proximity of h_{end} to h_{rescue} aggravates the situation beyond linear theory. Due to quantum fluctuations of the Higgs during inflation, there are regions in which the local Higgs value at the end of inflation exceeds the background value, overshoots h_{rescue} , and cannot be restored by reheating to the metastable electroweak vacuum created thermally. Such vacuum decay bubbles, with infinitely growing $|\rho_h|$, expand even after the end of inflation and eventually engulf our current horizon. These quantum fluctuations occur independently in the e^{120} causally disconnected regions at $N_{\text{cl}} \sim -20$ which make up our current horizon, and therefore avoiding the vacuum decay bubbles requires extreme fine-tuning [89]. In §4.3.4, we will cast this fine-tuning in terms of a breakdown in linear theory at the end of inflation, and we will show using the nonlinear δN formalism that fine-tuning away the vacuum decay bubbles directly tunes away the PBH abundance.

Vacuum decay bubbles can be avoided by stabilizing the Higgs at some large field value between h_{end} and h_{rescue} . By adding a singlet heavy scalar of mass m_s with appropriate

couplings to the theory, a threshold effect can be exploited to lift the Higgs effective potential during inflation and induce a new true minimum at $h \sim m_s$, preventing unbounded runaway [100, 92]. For the purposes of this mechanism this Higgs potential beyond the Standard Model can be modeled as

$$\lambda^{\text{BSM}} \simeq \lambda^{\text{SM}} + \frac{\delta\lambda}{2} \left[1 + \tanh \left(\frac{h - m_s}{\delta} \right) \right], \quad (4.17)$$

such that for $h \ll m_s$ the potential is as in the SM, given in (4.7), while for $h \gg m_s$ the potential is increased by $\delta\lambda(h^4/4)$. The step height $\delta\lambda$ should be such that the Higgs potential is stabilized, the step position m_s should be close to h_{rescue} , and the step width δ sufficiently narrow to not interfere with h_{end} . In Fig. 4.2, we plot the Higgs potential with the representative choices $\{\delta\lambda, m_s, \delta\} = \{0.02, 2000H, 100H\}$. While the BSM potential does not suffer from vacuum decay bubbles, it still experiences a breakdown in linearity at the end of inflation. We will therefore also use the nonlinear δN formalism to compute the conversion of ζ_h to ζ in this case.

Despite this difference at the end of inflation, the SM and BSM potentials are identical until large field values and therefore $\Delta_{\zeta_h}^2$ during inflation is the same in both potentials. Fluctuations at this stage can be locally remapped onto Δ_{ζ}^2 . For a successful PBH model, this remapping must still achieve $\Delta_{\zeta}^2 \sim 10^{-2}$ in both cases. We will therefore focus on the SM potential until we begin discussing nonlinear effects at the end of inflation in §4.3.

4.2 Higgs During Inflation

In this section, we compute the power spectrum of the Higgs fluctuations from their production inside the horizon through to a common epoch when all modes relevant for observation are superhorizon in scale.

In §4.2.1, we present the equation of motion for the Higgs and describe the local com-

petition between stochastic kicks and classical roll which governs its evolution. In §4.2.2, we argue that a well-defined background for the Higgs exists during the inflationary epochs relevant for observations, and that Higgs fluctuations during inflation can be computed by linearizing around this background mode-by-mode. In §4.2.3, we follow this procedure and compute the Higgs power spectrum during inflation at all scales relevant for observations, regarding the Higgs field as a spectator and hence dropping metric perturbations.

Combined, these arguments will show that during inflation the Higgs power spectrum at CMB scales is larger than the Higgs power spectrum at the primordial black hole scales

$$\Delta_{\zeta_h}^2(k_{\text{CMB}}) > \Delta_{\zeta_h}^2(k_{\text{PBH}}). \quad (4.18)$$

After the conversion of these superhorizon Higgs fluctuations during inflation into curvature fluctuations after inflation through Eq. (4.14), this leads to

$$\Delta_{\zeta}^2(k_{\text{CMB}}) > \Delta_{\zeta}^2(k_{\text{PBH}}), \quad (4.19)$$

in linear theory. Accounting for nonlinearity, we shall see that a similar relation between the scales holds as long as the mapping between the Higgs and curvature fluctuations is local. Therefore the first conclusion of the present paper is that in the Higgs vacuum instability scenario, a large amplitude of the power spectrum on small scales generating PBHs is ruled out by the CMB normalization. Conversely, if one chooses a different set of parameters in this scenario in order to satisfy the CMB normalization, one ends up with a small-scale power spectrum of at most $\mathcal{O}(10^{-9})$ which fails to form PBHs.

4.2.1 Classical Roll vs Stochastic Kicks

The equation of motion for the position- and time-dependent Higgs field $h(\vec{x}, N)$ is the Klein-Gordon equation

$$\square h(\vec{x}, N) = \left. \frac{\partial V}{\partial h} \right|_{h(\vec{x}, N)} \equiv V_{,h}|_{h(\vec{x}, N)}, \quad (4.20)$$

where here and throughout we denote partial derivatives with comma subscripts for compactness.

An important scale in this equation is the classical-roll scale h_{cl} , defined as follows. Every e -fold, the potential derivative leads $h(\vec{x}, N)$ to roll by

$$\Delta h \simeq -\frac{1}{3H^2} V_{,h}|_{h(\vec{x}, N)}. \quad (4.21)$$

Meanwhile, if one splits the field into a piece averaged on scales larger than a fixed proper distance $\sim 1/H$ and small scale modes which continually cross the averaging scale, the small scale modes can be viewed as providing a local stochastic noise term to the equation for the coarse-grained superhorizon field [101, 102]. The rms of this noise term each e -fold is

$$\langle \Delta h^2 \rangle^{\frac{1}{2}} \simeq \frac{H}{2\pi}. \quad (4.22)$$

In the language of perturbation theory, this is the per e -fold rms of the free field fluctuation δh and leads to a stochastic behavior of $h(\vec{x}, N)$. There are no subtleties involved in using N as a time coordinate since the number of e -folds is not a stochastic quantity so long as the Higgs remains a spectator.

The location h_{cl} in the potential where the roll contribution and the stochastic contribution are equal,

$$-\frac{1}{3H^2} V_{,h}|_{h_{\text{cl}}} = \frac{H}{2\pi}, \quad (4.23)$$

defines the ‘classical roll’ scale h_{cl} beyond which the classical term dominates the evolution of $h(\vec{x}, N)$. We show this scale in Fig. 4.2, where it lies at $h_{\text{cl}} \simeq 8.3H$.

The classical-roll scale is important because in slow-roll, Higgs modes which cross the horizon when the background satisfies Eq. (4.23) generically have a large power spectrum. In particular, the Higgs power spectrum at the scale k_{cl} which crosses the horizon at $N_{\text{cl}} = N(h_{\text{cl}})$ is order one at horizon crossing,

$$\Delta_{\zeta_h}^2(k_{\text{cl}}) \sim \frac{\langle \Delta h^2 \rangle}{(\Delta h)^2} \sim 1. \quad (4.24)$$

If these Higgs fluctuations are converted to large curvature fluctuations, they can satisfy the requirements of §4.1 such that $k_{\text{cl}} = k_{\text{PBH}}$.

4.2.2 Background and Linearization

We split $h(\vec{x}, N)$ equation into a background and perturbations

$$h(\vec{x}, N) = h(N) + \delta h(\vec{x}, N). \quad (4.25)$$

Here we define the background to be the part of the field representing the spatial average over our Hubble patch. Therefore at $N \sim -60$, the spatial average for the perturbation vanishes, $\langle \delta h(\vec{x}, -60) \rangle = 0$. The fluctuations are then generated by kicks from quantum fluctuations at $N > -60$.

To evolve the Higgs field under Eq. (4.20) we need to evaluate its position on the potential after $N = -60$, as established by its classical roll or quantum kicks. To do this, we need to establish whether the perturbations $\delta h(\vec{x}, N)$ are linear around the background $h(N)$.

For the mechanism to work, there should be a well-defined classical roll to the Higgs field at N_{cl} , and we can linearize the Higgs fluctuations at that epoch as usual [91, 89]. Between $-60 \lesssim N \lesssim N_{\text{cl}}$, there is a competition between the local stochastic kicks and the bulk

classical roll, and we need to check whether the kicks destabilize the average field in our Hubble patch.

In linear theory, stochastic kicks at the same \vec{x} but subsequent times evolve independently from each other. In particular, the potential term in the Klein-Gordon equation controls their interactions. When the Higgs is a spectator field, and we can expand this term around the homogeneous piece as

$$V_{,h}|_{h(\vec{x},N)} = V_{,h}|_{h(N)} + \delta h(\vec{x}, N) V_{,hh}|_{h(N)} + \dots \quad (4.26)$$

where ‘...’ contains terms higher order in $\delta h(\vec{x}, N)$. If we neglect the higher order terms, then each subsequent kick evolves as a free field and, as previously mentioned, has rms $H/2\pi$ at horizon crossing. The higher order terms then are suppressed relative to the linear term by

$$\frac{1}{2} \frac{V_{,hhh}}{V_{,hh}} \sqrt{\langle \delta h^2 \rangle} \simeq \frac{1}{2} \frac{V_{,hhh}}{V_{,hh}} \frac{H}{2\pi}, \quad (4.27)$$

in which we approximate all modes by their value at horizon crossing, which is appropriate while both the inflaton and the Higgs fields are slowly rolling. We plot this quantity for the Standard Model Higgs potential in Fig. 4.3 and show that it is less than one at all scales in the unstable region.

Therefore, the distance traveled due to stochastic kicks between $N = -60$ and N_{cl} accumulates as a random walk. For $N_{\text{cl}} = -20$, they therefore lead to a displacement

$$|\Delta h|_{\text{stochastic}} = \frac{H}{2\pi} \times \sqrt{40} \simeq H. \quad (4.28)$$

This is significantly less than the distance between the classical-roll scale $h_{\text{cl}} = 8.3H$ and the maximum of the potential $h_{\text{max}} = 4H$. This means that stochastic kicks do not, over 40 e -folds, kick our horizon into the other side of the potential. Therefore our whole Hubble volume was on the unstable side of the potential when it crossed the inflationary horizon at

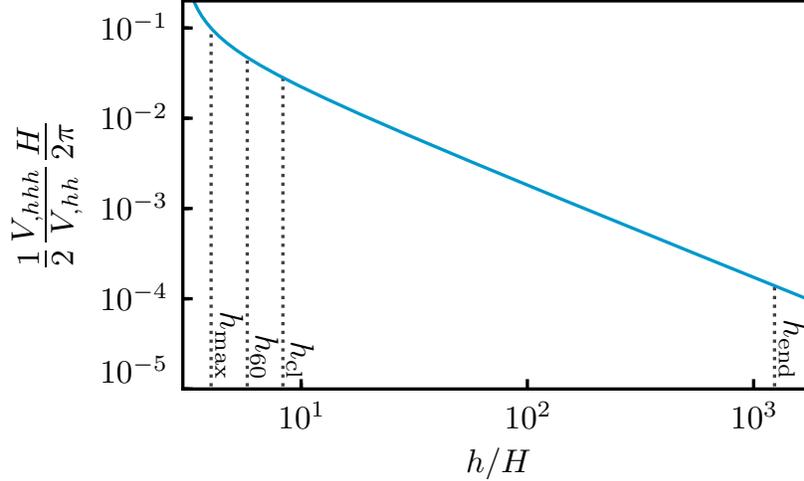


Figure 4.3: Interaction strength for modes of the typical horizon crossing amplitude $H/2\pi$ from (4.27). Since it is far less than one at the relevant scales h_{60} and h_{cl} , modes evolve independently and the cumulative background roll (4.31) dominates over the stochastic displacement (4.28). Linear perturbation theory holds until δh grows much larger than $H/2\pi$, near h_{end} (see §4.3.4).

$N = -60$.

Moreover, the total displacement from stochastic kicks (4.28) is less than the amount the background field rolls in these 40 e -folds as we shall now see. For a homogeneous field $h(N)$, Eq. (4.20) becomes

$$h'' + (3 - \epsilon_H)h' + \frac{V_{,h}}{H^2} = 0, \quad (4.29)$$

where $\epsilon_H = -H'/H$ is the first Hubble slow-roll parameter, which is zero during the exact de Sitter inflation in our fiducial model.

The initial conditions h_{60} and h'_{60} for this equation should be such that the Higgs reaches the desired field position h_{end} at the end of inflation close enough to h_{rescue} such that R is maximized. With one constraint and two initial values, a range of h'_{60} and h_{60} can lead to $h(N=0) = h_{\text{end}}$. Assuming attractor initial conditions for the Higgs, we choose

$$h'_{60} = -\frac{1}{(3 - \epsilon_H)H^2} V_{,h}|_{h_{60}}, \quad (4.30)$$

making the initial field position given by Eq. (4.11) $h_{60} \simeq 5.8H$, when h_{end} is set by Eq. (4.12).

Therefore the classical roll from -60 to N_{cl} is

$$|\Delta h|_{\text{roll}} = |h_{60} - h_{\text{cl}}| = 2.5H, \quad (4.31)$$

which is significantly larger than the stochastic displacement (4.28) but is nonetheless safely on the unstable side of the potential. This occurs despite the fact that each stochastic kick is larger than the per e -fold roll because the roll is coherent across our Hubble volume while the kicks are random.

Therefore we have a consistent picture where if we begin with an average field in our horizon volume around $h_{60} \sim 5.8H$, then our local background will reach h_{cl} at $N \sim -20$, unspoiled by stochastic kicks. Between these scales perturbations are linear, thanks to Fig. 4.3, and subdominant over the background roll. Our background will then continue to roll to h_{end} , where the Higgs will be uplifted. We plot this background in the lower panel of Fig. 4.2. App. A of Ref. [4] shows that this picture is consistent with the creation of our background from superhorizon stochastic fluctuations.

We can now use this background to solve for δh mode by mode during inflation for all relevant observational scales as in linear theory. This linearization depends on ignoring the interaction of Higgs fluctuations rather than the full machinery of linear perturbation theory for the metric and the matter, and in particular its validity does not assume $|\zeta_h| \ll 1$. Higgs nonlinearities become important in the last e -fold of inflation and beyond as the field fluctuations are amplified by the Higgs instability. Such nonlinear effects will affect the superhorizon CMB and PBH modes equally as we shall show in §4.3.

4.2.3 Higgs Power Spectrum

The linearized Klein-Gordon equation for the Fourier mode $\delta h^k(N)$ of $\delta h(\vec{x}, N)$ in spatially flat gauge is

$$\left(\frac{d^2}{d\eta^2} + 2\frac{\dot{a}}{a} \frac{d}{d\eta} + k^2 \right) \delta h^k + a^2 \delta V_{,h}^k \simeq 0, \quad (4.32)$$

where $\delta V_{,h}^k = V_{,hh} \delta h^k$ during inflation, and note that, unlike in the rest of this thesis, in this chapter an overdot denotes a derivative with respect to the conformal time η . Here we have dropped metric perturbations, which are suppressed when the Higgs is a spectator; these are restored in App. B of Ref. [4] for completeness.

The Klein-Gordon equation can then be conveniently expressed in terms of the auxiliary variable $u^k \equiv a\delta h^k$,

$$\ddot{u}^k + \left(k^2 - \frac{\ddot{a}}{a} + a^2 V_{,hh} \right) u^k = 0, \quad (4.33)$$

which holds at all orders of background and Higgs slow roll parameters. To order $\mathcal{O}(\epsilon_H^2)$ and $\mathcal{O}(\epsilon_H \eta_H)$, where η_H is the second Hubble roll parameter (see, e.g., Ref. [51]), but fully general in terms of the Higgs roll, we can write

$$\ddot{u}^k + \left(k^2 - \frac{\ddot{z}}{z} \right) u^k = 0, \quad (4.34)$$

where

$$z \equiv H\dot{h}. \quad (4.35)$$

This equation, of the Mukhanov-Sasaki type, is conveniently solved in the variable $s \equiv \eta_{\text{end}} - \eta$, the positive decreasing conformal time to the end of inflation (see, e.g., Refs. [103, 10]).

First, let us focus on the evolution in the superhorizon regime. In that limit, the analytic

solution for the Mukhanov-Sasaki equation (4.34) is given by

$$\frac{u^k}{z} = c_0 + c_1 \int \frac{d\eta}{z^2}, \quad (4.36)$$

where c_0 and c_1 are constants. So long as the second mode is decaying, we therefore have that on superhorizon scales

$$\frac{\delta h^k}{h'} = c_0 H^2 + \mathcal{O}(\epsilon_H, \eta_H). \quad (4.37)$$

In linear theory, the curvature perturbation on uniform Higgs density slices is obtained by gauge transformation as

$$\zeta_h^k = -\frac{\delta \rho_h^k}{\rho_h'} \simeq -\frac{\delta h^k}{h'}, \quad (4.38)$$

where first the approximate equality indicates that when the Higgs is slowly rolling, uniform Higgs density and uniform Higgs field slicing coincide to order ϵ_H (see [4]). More generally ζ_h is defined as the change in e -folds from a spatially flat surface to a constant density Higgs surface. This linear approximation holds so long as $\rho_h''/\rho_h'^2 \delta \rho_h \ll 1$, as it is here (see Eq. (4.61)).

Using the superhorizon evolution equation (4.37), we therefore have that if H evolves during inflation, the curvature on uniform Higgs density slices is not conserved on superhorizon scales and in particular decays according to

$$\frac{\zeta_h^{k'}}{\zeta_h^k} = -2\epsilon_H, \quad (4.39)$$

at leading order in ϵ_H . This estimate of superhorizon evolution assumes only background slow roll.

This superhorizon evolution is due to a pressure perturbation on uniform density slices for the Higgs, in other words a nonadiabatic pressure, induced because the uniform Higgs density slicing is not a uniform Hubble slicing when $\epsilon_H \neq 0$. This phenomenon is studied in

detail in App. B of Ref. [4]. Conversely, if H is constant, then the fact that the Higgs field evolves onto an attractor solution implies that nonadiabatic stress vanishes thereafter and ζ_h is conserved nonlinearly. In this case, much like single-field slow-roll inflation, the Higgs field supplies the only clock and field perturbations are equivalent to changing that clock on the background trajectory.

Next, let us focus on the evolution from subhorizon scales to the superhorizon regime. This evolution can be tracked by solving the Mukhanov-Sasaki equation (4.34) with Bunch-Davies initial conditions deep inside the horizon

$$u^k(s) = \frac{1}{\sqrt{2k}} \left(1 + \frac{i}{ks} \right) e^{iks}. \quad (4.40)$$

For analytic estimates, we can assume slow-roll evolution of z , in which case we can take the de Sitter modefunction (4.40) to the superhorizon limit, and find that each field fluctuation crosses the horizon with amplitude

$$\delta h^k \simeq \frac{iH}{\sqrt{2k^3}}, \quad (4.41)$$

and the field fluctuation power spectrum at that time is

$$\Delta_{\delta h}^2(k) = \frac{k^3}{2\pi^2} |\delta h^k|^2 \simeq \left(\frac{H}{2\pi} \right)^2. \quad (4.42)$$

Using the gauge transformation Eq. (4.38) with the field fluctuation Eq. (4.41) and the field velocity from the slow-roll solution of Eq. (4.29), the curvature perturbation on uniform

Higgs density hypersurfaces at horizon crossing is

$$\begin{aligned}\zeta_h^k(\eta_k) &\simeq -\frac{iH}{\sqrt{2k^3}} \frac{1}{h'} \Big|_{\eta_k} \\ &\simeq -\frac{iH}{\sqrt{2k^3}} \frac{(3 - \epsilon_H)H^2}{-V_{,h}} \Big|_{\eta_k}.\end{aligned}\tag{4.43}$$

To lowest order in Higgs- and background-slow-roll η_k is chosen to be the epoch of horizon-crossing $k\eta_k = 1$, but to next order can be optimized to $k\eta_k = \exp[7/3 - \ln 2 - \gamma_E]$, with γ_E the Euler-Mascheroni constant [103, 10].

We can now estimate the relative amplitude of $\Delta_{\zeta_h}^2$ on CMB and PBH scales. Choosing some comparison time η_* once both scales have exited the horizon but far enough from the end of inflation that slow-roll parameters are still small, we have

$$\begin{aligned}\frac{\Delta_{\zeta_h}^2(k_{\text{CMB}})}{\Delta_{\zeta_h}^2(k_{\text{PBH}})} \Big|_{\eta_*} &\simeq \left(\frac{H_{\text{PBH}}}{H_{\text{CMB}}}\right)^4 \frac{k_{\text{CMB}}^3 \left|\zeta_h^{k_{\text{CMB}}}\right|_{\eta_k}^2}{k_{\text{PBH}}^3 \left|\zeta_h^{k_{\text{PBH}}}\right|_{\eta_k}^2} \\ &\simeq \left(\frac{H_{\text{CMB}}}{H_{\text{PBH}}}\right)^2 \left(\frac{V_{,h}|_{\text{PBH}}}{V_{,h}|_{\text{CMB}}}\right)^2,\end{aligned}\tag{4.44}$$

where we have used at horizon crossing the Higgs slow-roll expression (4.43) and outside the horizon the Hubble slow-roll expression (4.37). Thus in the generic situation where H is decreasing and the Higgs rolls downhill, we find that $\Delta_{\zeta_h}^2(k_{\text{CMB}})/\Delta_{\zeta_h}^2(k_{\text{PBH}}) > 1$. The most optimistic case for the scenario is therefore the one where H is strictly constant between the CMB and PBH scales, and it still results in $\Delta_{\zeta_h}^2(k_{\text{CMB}})/\Delta_{\zeta_h}^2(k_{\text{PBH}}) > 1$.

We show in Fig. 4.4 the Higgs power spectrum computed by solving the Mukhanov-Sasaki equation exactly in the optimistic case where H is constant between CMB and classical-roll scales, evaluated at the convenient time η_* when all relevant modes have crossed the horizon. We compare this exact solution to the slow-roll expression (4.43) with the optimized freeze-out epoch. For decreasing H , the ratio $\Delta_{\zeta_h}^2(k_{\text{CMB}})/\Delta_{\zeta_h}^2(k_{\text{PBH}})$ would be larger than the

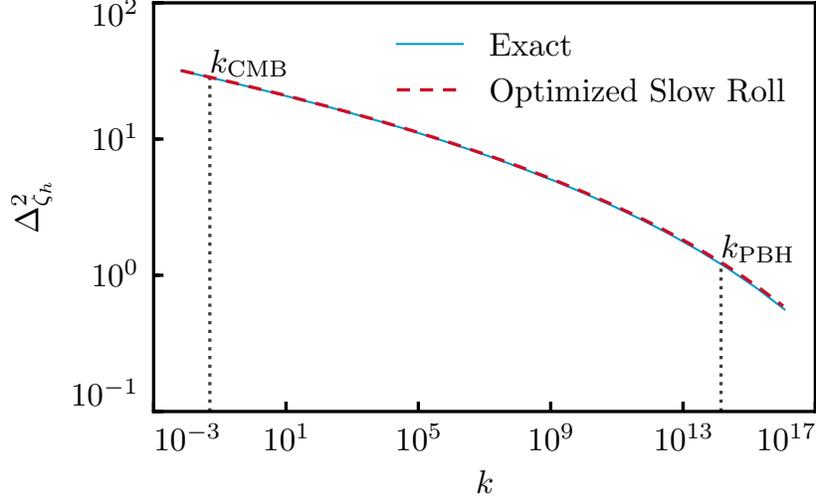


Figure 4.4: The Higgs power spectrum during inflation computed as described in §4.2.3 on a uniform Higgs energy density slice by exact solution of the Mukhanov-Sasaki equation (4.34) (solid blue) and by the optimized slow-roll approximation (4.43) (dashed red), at some time η_* after the relevant modes have crossed the horizon in the optimistic scenario where H is constant until η_* . The Higgs power spectrum is larger on CMB scales than on the classical-roll scales.

one estimated from Fig. 4.4.

These Higgs fluctuations at η_* will be converted to a total curvature fluctuations after Higgs decay by the factor $R(\zeta_h)^2$ which we discuss in §4.3. The key is that this mapping affects all mode contributions to ζ_h uniformly. For example in linear theory R is a constant whose value must be ~ 0.1 for successful PBH formation at k_{PBH} . Eq. (4.44) determines the total curvature power relative to this scale. In particular, the power spectrum at CMB scales is an order of magnitude larger than the power spectrum at the classical-roll scale. It is simply a feature of the Higgs potential that the field slope increases as the Higgs goes farther into the unstable region, and therefore that the Higgs fluctuation shrinks as k increases. Thus, a model with $\Delta_{\zeta_h}^2(k_{\text{CMB}})/\Delta_{\zeta_h}^2(k_{\text{PBH}}) > 1$ that forms PBHs at k_{PBH} will necessarily violate CMB constraints.

The results of this section hold equally for the SM and BSM potential. The difference between the two potentials enters only into $R(\zeta_h)$ which converts these results into the final curvature perturbation after Higgs decay. More generally as long as this mapping depends

only on field amplitude and not on k explicitly, PBHs cannot be formed from the Higgs instability without violating CMB constraints.

* * *

In summary, we have shown that during inflation

$$\Delta_{\zeta_h}^2(k_{\text{CMB}}) > \Delta_{\zeta_h}^2(k_{\text{PBH}}). \quad (4.45)$$

As we argued in §4.1, the conversion of the inflationary ζ_h to the final ζ depends only on the amplitude of ζ_h and thus all the information about reheating can be encoded in a scale-independent function $R(\zeta_h)$.

This means that in linear theory, where R is a constant, if primordial black holes are produced on small scales then on large scales

$$\Delta_{\zeta}^2(k_{\text{CMB}}) > \Delta_{\zeta}^2(k_{\text{PBH}}) \gtrsim 10^{-2}, \quad (4.46)$$

which is incompatible with measurements of the CMB.

Nonlinearly, when CMB and PBH modes cannot be tracked independently through the final e -folding of inflation and reheating, the $\Delta_{\zeta_h}^2(k)$ results in this section provide the super-horizon initial conditions which can be mapped to the final ζ . Given that this local mapping $R(\zeta_h)$ does not distinguish between Higgs fluctuations of different physical scales, we will argue in §4.3 that even nonlinearly, Higgs induced curvature fluctuations on CMB scales will be larger than those on PBH scales.

We now study in §4.3 the specific values taken by the conversion function $R(\zeta_h)$ itself. This will allow us to determine whether or not the Higgs fluctuations computed here can be transferred into large enough curvature perturbations to form PBHs, regardless of the compatibility with the CMB. Moreover, given that we have produced large Higgs fluctuations

on CMB scales, this will allow us to determine under which conditions Higgs criticality is incompatible with the small curvature fluctuations observed in the CMB.

4.3 Higgs and Reheating

We now track the curvature perturbations ζ_h on uniform Higgs density slices during inflation, computed in §4.2, through the end of inflation, reheating, and Higgs decay to compute the final curvature perturbations ζ on uniform total density hypersurfaces relevant for PBH formation.

In §4.3.1, we discuss how the Higgs evolves near the end of inflation and present the basic features of the instantaneous reheating model proposed by Ref. [91].

In §4.3.2, we discuss how to use the nonlinear δN formalism to convert ζ_h to ζ for any local reheating scenario, and we discuss jump conditions which must be satisfied during instantaneous reheating. We also present linearized δN formulae which yield results corresponding to those of linear perturbation theory, allowing an important crosscheck of the computation.

In §4.3.3, we follow the assumption of Refs. [91, 89, 92, 104] that linear theory holds through reheating and we compute explicitly the conversion of the inflationary ζ_h to the final ζ . We show that energy conservation at reheating, neglected in previous works, prevents the model from achieving the required $R = 0.1$ for both the SM and BSM potentials and thus PBH are not produced in sufficient quantities to be the dark matter in linear theory.

In §4.3.4 we show that linear theory is in fact violated at the end of inflation and we explicitly compute the full nonlinear conversion $R(\zeta_h)$ for the SM and BSM Higgs effective potentials. We show that PBHs in sufficient abundance to be the dark matter are never formed, second-order gravitational waves are suppressed, and only for a special class of criticality scenarios can observable perturbations be produced on CMB scales.

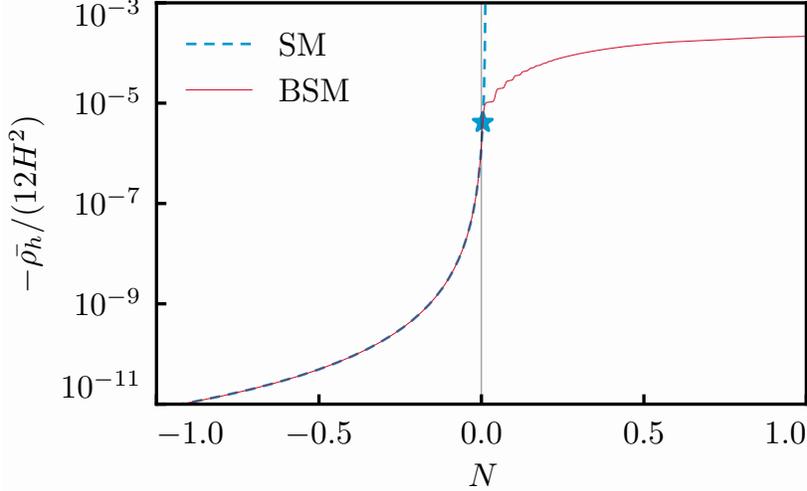


Figure 4.5: The (negative) Higgs energy density evolution during inflation, with $H = \text{const.}$ $N = 0$ corresponds to the end of inflation for the background, and $N > 0$ shows how the Higgs energy would evolve in local regions in advance of the background. The SM Higgs (dashed blue) at $N = 0$ is close to saturating the rescue condition (4.53) (blue star). The BSM Higgs (solid red) hits a wall during inflation and therefore is always rescued. Due to the Higgs attractor behavior, the exact value of $\zeta(0^+)$ for any given local shift $\delta N = \zeta_h$ can be read off using Eq. (4.59).

4.3.1 Instantaneous Reheating

As the Higgs travels farther and farther on the unstable side of the SM potential, it rolls faster and faster and if inflation never ended its energy density would diverge in finite time. In Fig. 4.5, we show this ρ_h during this last phase of the instability as a function of the number of e -folds. Of course inflation does end and for our chosen example this occurs at $N = 0$ and a field value h_{end} in the background. In this case, $N > 0$ then shows how the Higgs energy would continue to evolve if inflation did not end at $N = 0$. This range around $N = 0$ will also be useful when we consider perturbations that can be ahead of or behind the background value.

As we can see, the fiducial position of the background SM Higgs we have chosen is near-critical. Its energy is increasing rapidly and if inflation lasts much longer it will gain sufficient energy such that it is no longer a spectator. On this edge, the *background* Higgs field experiences the same evolution in the SM and BSM cases by construction (see Fig. 4.5).

Once inflation ends, the process of reheating transfers the inflaton's energy to radiation. The Higgs then no longer evolves in vacuum and the Klein-Gordon equation becomes

$$\square h = V_{,h}^T, \quad (4.47)$$

where V^T is the thermal effective potential in a bath of temperature T [90],

$$V^T = V + \frac{1}{2}M_{\text{T}}^2 h^2 e^{-h/2\pi T}, \quad M_{\text{T}}^2 \simeq 0.12T^2. \quad (4.48)$$

If reheating is instantaneous, as proposed by Ref. [91], then the total energy is conserved across it,

$$(\rho_\phi + \rho_h)(0^-) = (\rho_{\text{r}} + \rho_h)(0^+). \quad (4.49)$$

The division of the total energy density after inflation into a radiation piece ρ_{r} and a thermal component to the Higgs ρ_h is somewhat arbitrary. Since to leading order in ρ_h/ρ_{tot} , the curvature perturbation after the Higgs has decayed depends only on the conservation of the total energy, for convenience we choose to define ρ_{r} as a separately conserved thermal piece obeying the equation of motion

$$\rho_{\text{r}}' = -4\rho_{\text{r}}, \quad (4.50)$$

and the thermal bath temperature T is then

$$T = \left(\frac{30\rho_{\text{r}}}{\pi^2 g_*} \right)^{1/4}, \quad (4.51)$$

where $g_* = 106.75$ is the number of degrees of freedom in the Standard Model. Conservation of the total stress-energy along with the Higgs equation of motion (4.47) then imply that

the separately conserved Higgs energy is

$$\rho_h \equiv \frac{1}{2} \frac{\dot{h}^2}{a^2} + V^T. \quad (4.52)$$

Changes in V^T as the Higgs rolls down the effective potential then provide kinetic energy for the field as if it were a true potential energy.

Since $3H^2(0^-) = (\rho_\phi + \rho_h)(0^-)$, neglecting the Higgs' energy density contribution to H during inflation but including it after entails a dynamically negligible $\mathcal{O}(\rho_h/\rho_\phi) \sim \mathcal{O}(\rho_h/3H^2)$ discontinuity in the Hubble rate at the end of inflation H_{end} (see Fig. 4.5). On the other hand, strict energy conservation (4.49) at reheating is important because we will evaluate perturbations on constant density surfaces.

After reheating, the thermal effective potential in Eq. (4.47) will rescue the Higgs from the unbounded SM minimum so long as

$$\text{Max}[h] < h_{\text{rescue}}, \quad (4.53)$$

where $\text{Max}[h]$ is the maximum displacement of the Higgs field and h_{rescue} is the peak of the thermal potential V^T . Due to the nonzero kinetic energy of the Higgs at the end of inflation, $\text{Max}[h]$ is larger than the field displacement at the end of inflation h_{end} . We mark the maximal point which saturates this bound with a star in Fig. 4.5.

Neglecting the exponential term, we find that the peak of the uplifted potential at reheating is

$$h_{\text{rescue}}^{(0)} = \frac{M_T}{\sqrt{|\lambda^{\text{SM}}|}}, \quad (4.54)$$

a solution which can be iterated to account for the exponential term, yielding

$$h_{\text{rescue}}^{(1)} = h_{\text{rescue}}^{(0)} e^{-h_{\text{rescue}}^{(0)}/4\pi T} \sim 1.6 \sqrt{H_{\text{end}} M_{\text{Pl}}}, \quad (4.55)$$

where we have used $|\lambda| \sim 0.007$ and which with $H_{\text{end}} = 10^{12}$ GeV evaluates to $\sim 2500H$. The value of the uplifted Higgs potential at this approximate maximum is

$$V^T(h_{\text{rescue}}^{(1)}) \sim 0.02H_{\text{end}}^2 M_{\text{Pl}}^2. \quad (4.56)$$

These scalings are in good agreement with the exact calculation for $h_{\text{rescue}}^{(0)}/4\pi T \ll 1$ and they serve to highlight the dependence of the results with parameter choices. For our fiducial parameter set, the exact calculation yields $h_{\text{rescue}} \simeq 2400H_{\text{end}}$, $V^T(h_{\text{rescue}}) \simeq 0.02H_{\text{end}}^2 M_{\text{Pl}}^2$.

If the Higgs is rescued, then it oscillates in its uplifted temperature-dependent potential, redshifting as radiation on the cycle-average up to corrections from the non-quadratic components of its potential, until it decays on the e -fold timescale on constant Hubble surfaces. The rescue point is therefore relevant even for the BSM potential. In our example shown in Fig. 4.5, we set the m_s barrier close to h_{rescue} to maximize the instability while ensuring that the field returns to the electroweak vacuum after reheating.

We now describe in §4.3.2 how to track perturbations through the end of inflation and this instantaneous reheating epoch.

4.3.2 *Nonlinear Curvature Evolution*

The PBH abundance depends on the probability that the local horizon averaged density field exceeds some collapse threshold δ_c . We approximate this by the Gaussian probability that the curvature on uniform total density slices ζ lies above some threshold ζ_c . We therefore need to compute ζ after the Higgs decays.

Nonlinearly in the Higgs field perturbations, the transformation of the curvature on uniform Higgs density slicing during inflation ζ_h to the curvature on uniform total density slicing after inflation ζ can be performed in the δN formalism [66, 67, 68, 105], which allows us to evolve superhorizon perturbations by counting the number of e -folds of expansion from

an initial flat slice at some convenient initial time N_i to a uniform total density slice at a final time N ,

$$\zeta(N, \vec{x}) = \mathcal{N}(\rho_\alpha(N_i, \vec{x}); \rho_{\text{tot}}(N)) - \bar{\mathcal{N}}(\bar{\rho}_\alpha(N_i); \rho_{\text{tot}}(N)), \quad (4.57)$$

where \mathcal{N} is the local number of e -folds of expansion from the initial flat hypersurface at N_i on which any fields α have energy densities $\rho_\alpha(N_i, \vec{x}) = \bar{\rho}_\alpha(N_i) + \delta\rho_\alpha(N_i, \vec{x})$ to a final surface of uniform total density $\rho_{\text{tot}}(N)$, and $\bar{\mathcal{N}}$ is the corresponding expansion of the unperturbed universe. Using the separate universe assumption, \mathcal{N} can be computed in terms of background FLRW equations for a universe with the labeled energy contents.

If we chose N_i to be some time during inflation when we know the superhorizon density fluctuation $\delta\rho_h$ (or ζ_h) from our computation in §4.2, then since we are considering only perturbations sourced by the Higgs we can make the inflaton density at N_i implicit and keep only the dependence on the initial $\delta\rho_h$. We will perform this fully general calculation in §4.3.4.

To understand these results, it is also useful to have a simple analytic approximation for the impact of reheating given conditions just before reheating. In this case we can set the initial time just after reheating at $N_i = 0^+$. The Higgs energy density is a small component of the total energy budget, and therefore to leading order in ζ we can evaluate the δN formula assuming that $\rho_{\text{tot}} \propto a^{-4}$ to find

$$\begin{aligned} \zeta(0^+, \vec{x}) &\simeq \frac{1}{4} \ln \left(\frac{\rho_{\text{tot}}(0^+, \vec{x})}{\bar{\rho}_{\text{tot}}(0^+)} \right) \\ &\simeq \frac{\rho_{\text{tot}}(0^+, \vec{x}) - \bar{\rho}_{\text{tot}}(0^+)}{12H_{\text{end}}^2}. \end{aligned} \quad (4.58)$$

Conservation of energy at reheating (4.49) then implies the $N = 0$ jump condition

$$\zeta(0^+, \vec{x}) = \frac{\rho_h(0^-, \vec{x}) - \bar{\rho}_h(0^-)}{12H_{\text{end}}^2}. \quad (4.59)$$

Note that this condition applies to nonlinear Higgs density fluctuations $|(\rho_h - \bar{\rho}_h)/\bar{\rho}_h| \gg 1$ so long as $|(\rho_h - \bar{\rho}_h)/\bar{\rho}_{\text{tot}}| \ll 1$.

We can see from (4.59) that the post-inflationary energy partitioning chosen in (4.49) does not enter into the total curvature just after reheating. Instead, $\zeta(0^+, \vec{x})$ is determined solely by energy conservation.

We can further simplify this condition by noting that to the extent that H is constant during inflation, which we have shown by (4.44) is the most optimistic scenario for PBH production, the shift in e -folds to a constant Higgs energy density $\delta N_h = \zeta_h$ is conserved nonlinearly.¹ Therefore $\rho_h(0^-, \vec{x}) = \bar{\rho}_h(-\delta N_h)$, with this Higgs density computed as though inflation did not end at $N = 0$ as in Fig. 4.5. We can therefore read off $\zeta(0^+, \vec{x})$ for a given ζ_h from $\bar{\rho}_h(N)$ as

$$\zeta(0^+, \vec{x}) = \left. \frac{\bar{\rho}_h(-\zeta_h) - \bar{\rho}_h(0^-)}{12H^2} \right|_{\text{inf}}, \quad (4.60)$$

where $|_{\text{inf}}$ denotes this convention of evaluating the background as if inflation never ends.

Before using these nonlinear formulae (4.57) and (4.60) in §4.3.4, we will in §4.3.3 perform the calculation using linear perturbation theory. To validate the linear theory calculations, below we derive linear approximations to the δN formulae.

The full δN formula (4.57) can be linearized in $\delta\rho_{h_i} \equiv \delta\rho_h(N_i, \vec{x})$ to obtain

$$\zeta \simeq \frac{\partial\mathcal{N}(\rho_{h_i}; \rho_{\text{tot}})}{\partial\rho_{h_i}} \delta\rho_{h_i} \simeq -\frac{\partial\mathcal{N}(\rho_{h_i}; \rho_{\text{tot}})}{\partial\rho_{h_i}} \rho'_{h_i} \zeta_{h_i}, \quad (4.61)$$

where ρ_{h_i} and the ρ'_{h_i} are Higgs density and its derivative at N_i and ζ_{h_i} is the Higgs curvature

1. If H evolves, the leading order effect will be simply to shrink the Higgs δN_h .

at that time. Likewise, a linear Taylor expansion of the jump condition (4.60) is given by

$$\zeta(0^+, \vec{x}) \simeq -\zeta_h(0^-) \times \frac{\bar{\rho}'_h(0^-)}{12H^2}. \quad (4.62)$$

The ratio $-\rho'_h(0^-)/12H^2$ is the rescaling factor $R(0^+)$ if H is constant through to the end of inflation.

As we shall see below, these linear δN formulae provide an important point of contact between the nonlinear δN and linear perturbation theory approaches.

4.3.3 Linear Conversion

We now follow the assumption of Refs. [91, 89, 92, 104] that linear theory holds through reheating, and we show that under this assumption primordial black holes cannot be the dark matter.

While Higgs field values h and δh and their derivatives h' and $\delta h'$ are all continuous through reheating, the Higgs potential and its slope change instantaneously when the Higgs potential is uplifted. Therefore the Higgs energy density (4.52), its derivative

$$\rho'_h = -3\frac{\dot{h}^2}{a^2} - V_{,T}^T T, \quad (4.63)$$

and its perturbation

$$\delta\rho_h = \frac{1}{a^2}\dot{h}\delta\dot{h} + V_{,h}^T\delta h + V_{,T}^T\delta T, \quad (4.64)$$

are not continuous with their values at $N = 0^-$. δT here is any perturbation in the bath temperature correlated with the Higgs, which we shall see is generically induced at reheating. For simplicity, we have omitted here a contribution to the perturbed energy density coming from the metric lapse perturbation. Again, these are restored in App. B of [4]. Its relative contribution is negligible.

The jump in ρ'_h

$$\Delta [\rho'_h] = -V_{,T}^T T, \quad (4.65)$$

and in the energy density perturbation

$$\Delta [\delta\rho_h] = \delta h \left(V_{,h}^T - V_{,h} \right) + V_{,T}^T \delta T, \quad (4.66)$$

imply that the curvature perturbation (4.38) on constant Higgs energy density slices is discontinuous at reheating. This instantaneous change in ζ_h is due to an instantaneous source in the conservation equation from the interaction of the Higgs with the thermal bath.

However, the instantaneous increase in the Higgs energy density perturbation $\delta\rho_h$ does not come for free. Conservation of energy, which we imposed at the level of the background in Eq. (4.49), also holds locally. It implies that the increase in the Higgs energy density is counterbalanced by an induced perturbation in the radiation field

$$\delta\rho_r(0^+) = -\Delta [\delta\rho_h], \quad (4.67)$$

and therefore that the Higgs and radiation energy densities after uplift are nearly canceling. In other words, the uplift creates a Higgs-radiation isocurvature fluctuation rather than a net curvature fluctuation. To the extent that the Higgs fluctuation then redshifts like radiation, the isocurvature mode does not subsequently contribute to the curvature fluctuation.

The conserved curvature on uniform radiation density slices which corresponds to this induced radiation perturbation is

$$\zeta_r = -\frac{\delta\rho_r}{\rho'_r} = \frac{\delta T}{T}, \quad (4.68)$$

and solving for the radiation perturbation (4.67) using the jump in Higgs energy (4.66) and

Eq. (4.68), we find

$$\delta\rho_r(0^+) = -\delta h \left(V_{,h}^T - V_{,h} \right) \left(1 - V_{,T}^T \frac{T}{\rho_r'} \right)^{-1}. \quad (4.69)$$

This induced radiation perturbation comes from the direct interaction of the Higgs with the radiation during the thermal uplift. It is distinct from radiation perturbations corresponding to intrinsic inflaton fluctuations, which are uncorrelated and can be computed separately, or to inflaton perturbations produced by the gravitational influence of the Higgs perturbations during inflation, which are suppressed.

This radiation perturbation was omitted in Refs. [91, 104, 92, 89], though it in fact its role in conserving total energy has a large impact on the final curvature perturbation ζ . In particular, on a constant total density surface, the curvature perturbation is given by Eq. (4.2), reproduced here for convenience,

$$\zeta = \left(1 - \frac{\rho_h'}{\rho_{\text{tot}}'} \right) \zeta_r + \frac{\rho_h'}{\rho_{\text{tot}}'} \zeta_h. \quad (4.70)$$

Immediately after the uplift of the Higgs potential, ζ therefore satisfies

$$\begin{aligned} \zeta(0^+) &= \zeta_h(0^-) \times \frac{\rho_h'(0^-)}{\rho_{\text{tot}}'(0^+)} \\ &\simeq -\zeta_h(0^-) \times \frac{\rho_h'(0^-)}{12H_{\text{end}}^2}, \end{aligned} \quad (4.71)$$

by virtue of Eq. (4.67). This equation is nothing but the linear jump condition (4.62), this time derived from linear perturbation theory rather than the δN formalism.

To evolve $\zeta(N)$ from its value at $\zeta(0^+)$, we must solve for the Higgs perturbations after uplift. Again the most important aspect of this calculation is energy conservation. Energy conservation guarantees that the cancellation responsible for the suppression in the starting value $\zeta(0^+)$ is maintained on a timescale short compared to an e -fold.

Although during inflation we assumed the Higgs is a spectator, after potential uplift we

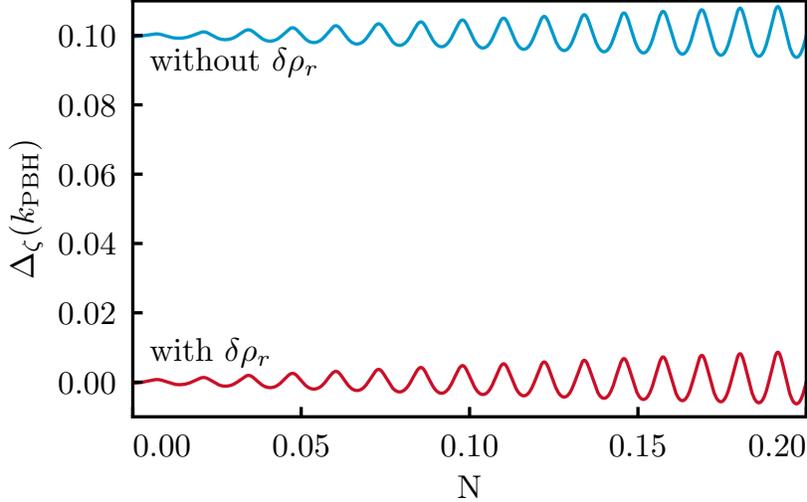


Figure 4.6: The curvature perturbation on uniform total energy slices computed in linear perturbation theory with and without including radiation perturbations required by energy conservation at reheating. These induced perturbations suppress the cycle-averaged curvature by several orders of magnitude (see §4.3.3).

jointly solve the Higgs background equation (4.29), now with $V \rightarrow V^T$, and the radiation background equation (4.50), with the Hubble rate determined by Friedmann equation. For the perturbations we solve the linearized Klein-Gordon equation (4.32) for the Higgs, now with a perturbed potential

$$\delta V_{,h}^{Tk} = V_{,hh}^T \delta h^k + V_{,hT}^T \delta T^k, \quad (4.72)$$

which accounts for the effect of temperature perturbations. Again the metric terms are negligible and we exploit here that ζ_{r} is constant to avoid solving perturbation equations for the radiation component; both are shown to be good approximations in App. B of [4].

To quantify the importance of the induced radiation perturbation (4.67), we chose h_{end} such that the post-inflationary Higgs contribution to the total $\Delta_{\zeta}^2(k_{\text{PBH}})$ is $\sim (0.1)^2$. This is the calculation performed in the literature which suggests that primordial black holes can be formed at the classical-roll scale. We then add the induced radiation contribution and see how Δ_{ζ} is affected.

We show these numerical results in Fig. 4.6. Here we plot $\Delta_{\zeta}(k_{\text{PBH}})$ with a phase

convention φ such that the analogous superhorizon Higgs fluctuation

$$\Delta_{\zeta_h} \equiv e^{i\varphi} \sqrt{\Delta_{\zeta_h}^2}, \quad (4.73)$$

is negative real during inflation. Note that Δ_{ζ_h} changes sign at the potential uplift and becomes positive real. After inflation, $\Delta_{\zeta}(k_{\text{PBH}})$ oscillates between negative and positive values but stays real.

It is immediately clear from Fig. 4.6 that the induced radiation perturbation $\delta\rho_r$ suppresses the amplitude of the total curvature ζ by orders of magnitude, making it much more difficult to achieve the required $R = 0.1$ in this model. For the fiducial background, which was claimed to produce $R = 0.1$, by taking into account $\delta\rho_r$ we instead have $R(0^+) \simeq 3 \times 10^{-4}$.

In Fig. 4.7, we validate our calculation of $\Delta_{\zeta}(k_{\text{PBH}})$ by also computing ζ from the linearized δN equation (4.61). The δN result relies solely on the behavior of the background equations and thus is an independent check on the rather involved perturbation theory calculations. The δN result agrees closely with our perturbation theory calculation and confirms that the induced radiation perturbation is crucial in this mechanism. This test would fail if the radiation compensation in Eq. (4.69) were omitted as in Refs. [91, 104, 92, 89].

ζ is not conserved after reheating, and in particular it oscillates due to the changing nonadiabatic pressure induced as the Higgs oscillates. Even though oscillations in $\delta\rho_h$ are relatively small, the initial near cancellation between $\delta\rho_h$ and $\delta\rho_r$ make them prominent in ζ . Moreover, deviations of the Higgs potential from a simple quadratic with a temperature-dependent mass make these oscillations grow in time. These effects are discussed in detail in App. B of Ref. [4]. So long as the Higgs decay time, generally of order an e -fold, is larger than the oscillation timescale, it is the cycle-averaged ζ that matters. Because of the initial outgoing trajectory of the Higgs, the instantaneous value of $|\zeta|$ at $N = 0^+$ is always larger than the cycle average of the first oscillations.

To the extent that the cycle-averaged Higgs energy redshifts as radiation, the cycle-

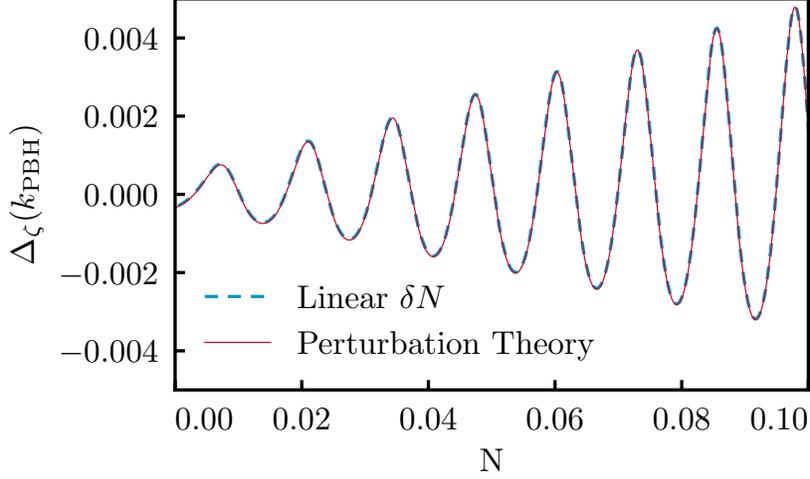


Figure 4.7: The linear perturbation theory calculation including induced radiation fluctuations agrees well with the linearized δN result based on (4.61), validating our result that energy conservation suppresses curvature fluctuations (see §4.3.3).

averaged value of ζ will be conserved. However, the nonquadratic terms in the Higgs potential also cause deviation from this behavior which leads the near cancellation between the Higgs and radiation energy densities gradually to break down.

In particular, Higgs perturbations redshift slightly slower than radiation on the cycle average,²

$$\langle \delta \rho_h^k \rangle \propto a^{-4-3\Delta w}, \quad (4.75)$$

with $\Delta w \equiv \langle w \rangle - 1/3 \sim -0.004$. Therefore the cancellation between the radiation piece and the Higgs piece gradually becomes undone,

$$\zeta^k = -\frac{\delta \rho_r^k + \delta \rho_h^k}{\rho'_{\text{tot}}} \sim -\frac{\delta \rho_r^k + \delta \rho_h^k}{\rho'_r} = \zeta_r^k + \frac{\delta \rho_h^k}{4\rho_r}, \quad (4.76)$$

and $\langle \zeta^k \rangle$ grows gradually.

2. The rate at which Higgs perturbations redshift (4.75) is different from the rate at which the background Higgs redshifts,

$$\langle \rho_h \rangle \propto a^{-4-3\Delta \bar{w}}, \quad (4.74)$$

with $\Delta \bar{w} \sim -0.002$. This means that there is also internal nonadiabatic stress in the Higgs field itself and so the cycle averaged ζ_h^k would also evolve.

Once the Higgs piece dominates, the cycle-average becomes

$$\langle \zeta^k \rangle(N) \sim -\frac{3}{4} \frac{\delta \rho_h^k(0^+)}{\rho_r(0^+)} \Delta w \times N \sim 10^{-3} N. \quad (4.77)$$

Thus the curvature grows to $\mathcal{O}(10^{-1})$ only on a timescale

$$\Delta N \sim 100 \text{ } e\text{-folds}. \quad (4.78)$$

The Higgs must decay to radiation well before this, and therefore the curvature perturbations cannot become large enough in this scenario to form PBHs.

Note that the details of the post-reheating evolution and in particular the de-cancellation rate derived here do depend on the radiation-Higgs split in (4.49) through the temperature dependence of the thermal potential (4.48). However, conservation of total energy on the timescale of Higgs oscillations imposes that the total curvature can grow only on the e -fold timescale, and only due to deviations in the redshifting rate of the different components. Therefore the qualitative result that this growth will take many e -folds is robust to our specific implementation here.

In summary, under the assumption that linear theory is valid through reheating, the Higgs instability mechanism falls far short of being able to form PBHs as the dark matter. Models that were previously thought to achieve the required $R = 0.1$ in fact produce $R \lesssim 10^{-3}$ once the radiation density perturbations required by energy conservation at reheating are properly accounted for.

4.3.4 *Nonlinear Conversion*

In §4.3.3, we computed curvature fluctuations assuming that the Higgs perturbations remain linear through reheating. In fact, the Higgs instability induces a breakdown of linear theory when the background position of the Standard Model Higgs at the end of inflation h_{end} is

close to the maximum rescue scale h_{rescue} .

This breakdown can be seen immediately from Fig. 4.5. With $\delta N = \zeta_h \sim \pm 1$, a typical outwardly perturbed region of the Standard Model Higgs field crosses h_{rescue} during inflation, gains exceedingly large negative energy and will inevitably backreact on the background trajectory. Reheating will be disrupted, the perturbed Higgs will not be rescued from the unbounded vacuum, and our universe will be destroyed. Even for smaller δN which do not cross h_{rescue} , the perturbed Higgs energy density is not well represented by the linear Taylor expansion (4.62) around the background value due to the extremely rapid evolution of ρ_h .

In terms of field interactions, linear theory itself also reveals its own breakdown. At the end of inflation, an order unity ζ_h leads to an RMS Higgs fluctuation of roughly

$$\delta h_{\text{end}} \simeq h'_{\text{end}} \simeq -\frac{1}{3H^2} \lambda h_{\text{end}}^3, \quad (4.79)$$

where we have used the Higgs slow-roll approximation throughout. With $h_{\text{end}} \sim 10^3 H$ and $|\lambda| \sim 10^{-2}$ we have

$$\delta h_{\text{end}} \simeq 10^7 H, \quad (4.80)$$

which as we have seen is orders of magnitude larger than the distance between $h_{\text{end}} \simeq 1200H$ and $h_{\text{rescue}} \simeq 2400H$. Moreover, the potential interaction ratio (4.27) is

$$\frac{1}{2} \delta h \frac{V_{,hhh}}{V_{,hh}} \simeq \frac{\delta h}{4h} \simeq 10^4 \gg 1, \quad (4.81)$$

and thus field fluctuations interact. Nevertheless, this breakdown has no effect on our previous computation of $\Delta_{\zeta_h}^2$ during inflation since ζ_h is conserved nonlinearly so long as $H \sim \text{const}$.

Though it was not phrased in terms of a breakdown of linear theory, Ref. [89] noted that the Standard Model Higgs is generally not rescued in this scenario. It was argued in Ref. [92] that the background h_{end} can be placed near h_{rescue} while multiverse and an-

thropic considerations justify tuning the local Higgs field at the end of inflation such that $\text{Max}[h(\vec{x})] < h_{\text{rescue}}$ everywhere. However, tuning $\delta h(\vec{x})$ at the end of inflation is equivalent to tuning $\delta N = \zeta_h$ to be small at the end of inflation and so it directly tunes away the ability to form PBHs as we shall now show.

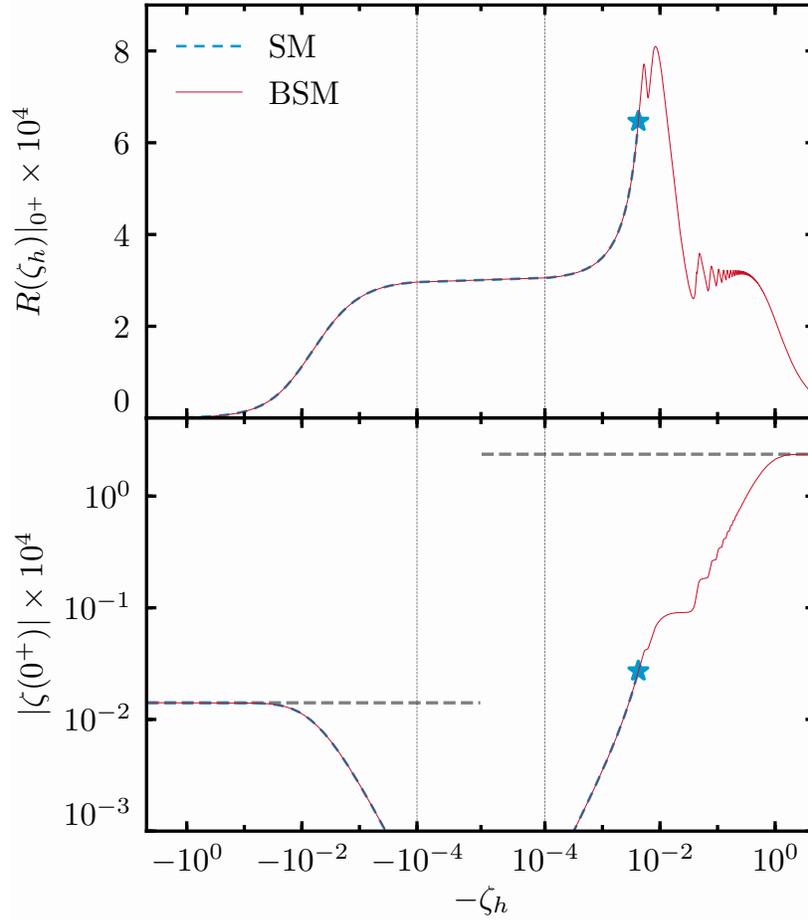


Figure 4.8: The nonlinear mapping of the inflationary ζ_h to the post-reheating $\zeta(0^+) = R(\zeta_h)|_{0^+} \zeta_h$. Fluctuations outwards, toward the instability, correspond to $-\zeta_h > 0$. The horizontal axis scale is linear between $\pm 10^{-4}$ and logarithmic elsewhere. R deviates from the linear theory value $\simeq 3 \times 10^{-4}$ for fluctuations larger than about $\pm 10^{-3}$. The horizontal dashed lines indicate analytic saturation values (4.85) and (4.86). For the SM Higgs, the maximal ζ satisfying the rescue condition (4.53) is marked with a star. For the BSM Higgs, ζ saturates to a maximum. Neither value is large enough to form PBHs

In Fig. 4.8 we show $\zeta(0^+)$ as a function of ζ_h as computed using the nonlinear δN formalism using Eq. (4.60) and Fig. 4.5. Linear theory holds for small enough inflationary $|\zeta_h| \lesssim 10^{-3}$, but breaks down for perturbations of the typical amplitude produced during inflation.

Large inwards perturbations away from the instability, shown on the left-hand side of Fig. 4.8, saturate to a constant $\zeta(0^+)$ that is independent of ζ_h . These uphill kicks produce a local Higgs energy density at $N = 0^-$ that has a much smaller magnitude than its background value as can be seen in Fig. 4.5. Using Eq. (4.59), the left-hand side saturation can therefore be written as

$$\zeta_{\text{SM}}^{\text{in}}(0^+) = \frac{-\bar{\rho}_h(0^-)}{12H^2} \simeq +1.4 \times 10^{-6}, \quad (4.82)$$

which we show as a horizontal dashed line on the left-hand side in Fig. 4.8.

Outward perturbations of the SM Higgs toward the instability, shown on the right-hand side of Fig. 4.8, are enhanced relative to linear theory. This is because the amplitude of the energy density of the Higgs shown in Fig. 4.5 grows much faster than expected from a linear approximation. The largest outward perturbations that satisfy the rescue condition (4.53) produce a curvature

$$\zeta_{\text{SM}}^{\text{out}}(0^+) = \frac{\rho_h(h_{\text{rescue}}) - \bar{\rho}_h(0^-)}{12H^2} \simeq -2.7 \times 10^{-6}. \quad (4.83)$$

Despite the enhancement of the Higgs perturbation relative to linear theory, $\zeta_{\text{SM}}^{\text{out}}$ evolves after inflation much like the linear theory ζ computed in §4.3.3. The cycle average of $\zeta_{\text{SM}}^{\text{out}}(N)$ is smaller than $\zeta_{\text{SM}}^{\text{out}}(0^+)$. So long as the Higgs redshifts like radiation after inflation, this value is then conserved. Nonlinear evolution does not change the conclusion of linear theory on PBHs with the SM potential.

So far in this nonlinear calculation we have kept the background trajectory of the Higgs fixed. We might wonder whether a different background Higgs trajectory, at fixed H , can

achieve a larger value for the saturating ζ . Moving the background away from the instability suppresses $\bar{\rho}_h(0^-)$, and so the largest curvature perturbation which can be produced in the SM, maximized over the choice of background trajectory, is

$$\zeta_{\text{SM}}^{\text{max}}(0^+) = \zeta_{\text{SM}}^{\text{out}}(0^+) - \zeta_{\text{SM}}^{\text{in}}(0^+) \simeq -4.1 \times 10^{-6}, \quad (4.84)$$

which is still insufficient to form PBHs. Therefore since ζ is uniquely determined by ζ_h in this way, anthropically tuning away $\zeta_h(\vec{x})$ or setting it to the edge of rescueability forbids the Standard Model Higgs from forming PBHs in sufficient abundance to be the dark matter.

Ref. [92] proposed that the mechanism could function with a BSM potential derived from the addition of a massive scalar as detailed in §4.1. The massive scalar adds a wall in the potential between the field value where the background Higgs ends inflation h_{end} and the maximum rescuable point h_{rescue} , preventing the local Higgs from reaching parts of the potential from which it cannot be rescued by thermal uplift at reheating.

When CMB and PBH modes cross the horizon, the BSM potential behaves like the SM potential and as we have seen in §4.2 this means that it generates larger inflationary ζ_h perturbations on CMB scales than on PBH scales. In addition, so long as the background trajectory never encounters the wall, this model behaves like the SM in linear theory and yields $R \ll 0.1$ as shown in §4.3.3.

However, whereas typical regions with outward field fluctuations were not rescued in the SM, in the BSM case such regions oscillate in a new minimum of the potential during inflation and then can be safely rescued at reheating.

To evolve fluctuations through this highly nonlinear process, we again use the nonlinear δN formalism described in §4.3.2, just as in the Standard Model case, and the BSM results for $\zeta(0^+)$ are also shown in Fig. 4.8. We compute results using the representative parameter set for the BSM scalar described in §4.1, and we will later show how our results scale with different choices of model parameters.

Inward perturbations of the BSM Higgs act just like inward perturbations in the SM and thus again lead to the same saturation

$$\zeta_{\text{BSM}}^{\text{in}}(0^+) = \zeta_{\text{SM}}^{\text{in}}(0^+) \simeq +1.4 \times 10^{-6}. \quad (4.85)$$

Large outward Higgs perturbations hit the BSM potential wall, become trapped in the new minimum at $h \sim m_s$, lose their kinetic energy, and end inflation with a potential dominated Higgs with energy V_{min} . This leads to a saturating curvature

$$\zeta_{\text{BSM}}^{\text{out}}(0^+) = \frac{V_{\text{min}} - \bar{\rho}_h(0^-)}{12H_{\text{end}}^2} \simeq -2.4 \times 10^{-4}, \quad (4.86)$$

which we show with a horizontal dashed line on the right-hand side in Fig. 4.8. This value is still too small to form enough PBHs to be the DM.

For perturbations which do not fully saturate this limit, $\zeta(0^+)$ has a stepped behavior and $R(0^+)$ an oscillatory one as depicted in Fig. 4.8. These features correspond to the energy density oscillations for the BSM Higgs seen in Fig. 4.5, induced because the Higgs has large oscillations around the potential minimum before settling on the e -fold timescale. Note that the approximate equality of the linear theory $R(0^+)$ and the nonlinear $R(0^+)$ corresponding to $|\zeta_h| \simeq 0.1$ is a coincidence: changes to the background position change the linear theory $R(0^+)$ while leaving $R(0^+)$ on this brief plateau fixed.

Once more we might wonder if a different choice of background trajectory could enhance the saturation value, but in fact the maximum over background trajectories can again be computed as

$$\zeta_{\text{BSM}}^{\text{max}}(0^+) = \zeta_{\text{BSM}}^{\text{out}}(0^+) - \zeta_{\text{BSM}}^{\text{in}}(0^+) \simeq \zeta_{\text{BSM}}^{\text{out}}(0^+), \quad (4.87)$$

and therefore PBHs cannot be formed for our fiducial BSM potential no matter the position of the background or the size of initial fluctuation.

Note that so far we have only computed $\zeta(0^+)$ for BSM. We should check whether ζ

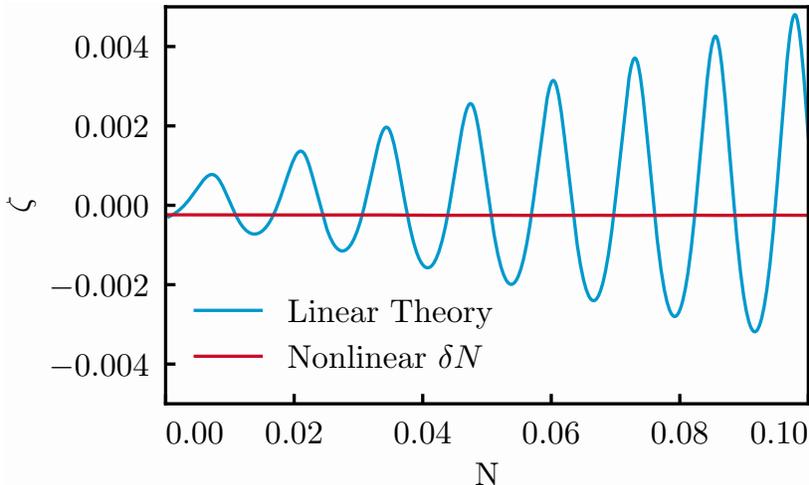


Figure 4.9: Curvature evolution after reheating. The BSM nonlinear δN result for the local $\zeta(\vec{x})$, from an example inflationary $\zeta_h(\vec{x}) = \Delta_{\zeta_h}(k_{\text{PBH}})$, is compared to the linear theory approximation $\Delta_{\zeta}(k_{\text{PBH}})$ of Fig. 4.7. The nonlinear evolution of ζ after reheating is too small to form PBHs (see §4.3.4).

evolves significantly after $N = 0^+$. We do so again with the δN formalism, using the full Eq. (4.57). Since inward fluctuations lead to a negligible $\zeta(0^+)$, we can select a typical outward field fluctuation with $\zeta_h(\vec{x}) = \Delta_{\zeta_h}(k_{\text{PBH}}) \simeq -1$ as an example.

We show this case in Fig. 4.9. As expected from Fig. 4.8, the nonlinearly evolved ζ is small, comparable in amplitude to the linear $\zeta(0^+)$ but not in its evolution. In fact ζ evolves negligibly after $N = 0^+$ and we can robustly conclude that PBHs are generically not formed nonlinearly in this case.

This lack of nonlinear evolution can be explained by the difference in the impact of uplift on the perturbations. In linear theory the small amplitude of ζ resulted from large cancellations between the Higgs and radiation perturbations due to energy conservation and the large impact of uplift. Nonlinearly the impact of uplift is much smaller, bounded by the BSM modification, and so the Higgs energy density fluctuations after reheating are no longer as dominated by the uplift contribution. In particular,

$$\frac{\rho_h(0^+, \vec{x}) - \bar{\rho}_h(0^+)}{\rho_h(0^-, \vec{x}) - \bar{\rho}_h(0^-)} \ll \left. \frac{\delta\rho_h(0^+)}{\delta\rho_h(0^-)} \right|_{\text{linear}}, \quad (4.88)$$

where the right-hand side is in linear theory. Therefore the cancellation with radiation is less dramatic than in linear theory. As discussed in §4.3.3, the cancellation and subsequent decancellation are responsible for the linear theory oscillations and slow secular drift (4.77), and therefore all these effects are suppressed in the nonlinear case. For the same reason, the details of the split of the total energy density into Higgs and radiation pieces are also less important for the nonlinear curvature. Our nonlinear results therefore essentially depend only on energy conservation during reheating.

Finally, the fiducial BSM potential used to compute the results of Fig. 4.8 was constructed according to the specifications of Ref. [92]: it uplifts the Standard Model potential somewhere between h_{end} and h_{rescue} . We might wonder whether PBHs could be formed by optimizing the position of the uplift so that it is as close as possible to h_{rescue} , maximizing the criticality of the scenario.

The maximum position of m_s will be just before h_{rescue} . This leads to a maximum curvature for this entire scenario of

$$\text{Max}[\zeta_{\text{BSM}}^{\text{max}}] = \frac{V(h_{\text{rescue}})}{12H_{\text{end}}^2}. \quad (4.89)$$

Using the approximate maximum rescueable field value (4.55) and $\lambda = \lambda^{\text{SM}} \sim -0.007$, we have

$$V(h_{\text{rescue}}^{(1)}) = \frac{1}{4}\lambda^{\text{SM}} \left(h_{\text{rescue}}^{(1)}\right)^4 = -0.012H_{\text{end}}^2, \quad (4.90)$$

which yields

$$\zeta_{\text{BSM}}^{\text{max}} \simeq -1.0 \times 10^{-3}, \quad (4.91)$$

which depends on H only through the logarithmic evolution of λ^{SM} evaluated at h_{rescue} . This estimate is in good agreement with the computation using the exact value of h_{rescue} which yields $\zeta_{\text{BSM}}^{\text{max}} = -8.2 \times 10^{-4}$.

Therefore no matter the size of inflationary Higgs perturbations, the position of the

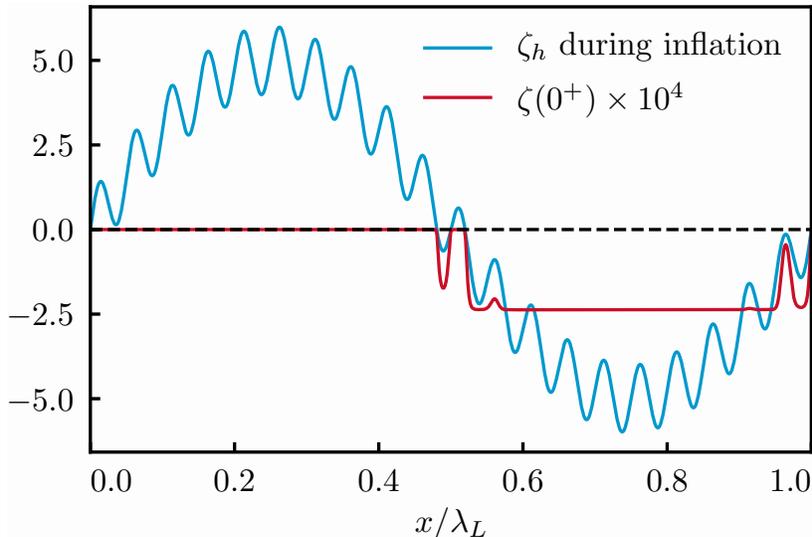


Figure 4.10: The real-space post-inflationary curvature field ζ (red) produced by the nonlinear transformation of the inflationary ζ_h (blue). Perturbations on small scales are suppressed when they occur where the long-wavelength λ_L mode has saturated the curvature field, and therefore CMB scale perturbations are larger than PBH scale perturbations fully nonlinearly in this mechanism (see §4.3.4).

background Higgs, the SM or BSM nature of the Higgs potential, the position of the BSM wall, or the Hubble rate, this mechanism does not produce perturbations large enough to form PBHs in sufficient abundance to be the dark matter.

Moreover, the largest possible curvature perturbations produced by this model, Eq. (4.91), are so small that the second-order gravitational waves predicted by Ref. [104] will be undetectable with LISA.

We can also ask what happens to CMB scale fluctuations nonlinearly in this model. We showed in §4.2 that during inflation $\Delta_{\zeta_h}^2(k_{\text{CMB}}) > \Delta_{\zeta_h}^2(k_{\text{PBH}})$. In Fig. 4.10, we show the effect of the highly nonlinear local transformation of ζ_h to ζ shown for the BSM potential in Fig. 4.8 on a cartoon realization of the curvature field on a constant Higgs density surface during inflation.

For visualization purposes, we have generated two modes apart by a factor of only 20 in scale rather than the ~ 35 e -folds which separate the CMB and PBH modes. We see

that fluctuations on the long-wavelength scales cause a saturation of the short-wavelength fluctuations, and therefore fully nonlinearly we have that perturbations are larger on long wavelengths than on small wavelengths in this model as long as this is true of the inflationary ζ_h itself. Therefore if the reheating scenario is changed somehow to achieve large PBH scale fluctuations, the CMB scale fluctuations will still be larger than PBH modes unless the functional form of the transformation shown in Fig. 4.8 is radically altered.

Beyond the specific motivation of PBH DM formation, we can now return to the question of whether Higgs instability is compatible with the small curvature fluctuations observed in the CMB.

For the SM Higgs, suppressing $\zeta_{\text{SM}}^{\text{out}}$ requires placing the background h_{end} far from h_{rescue} . Specifically, since the typical CMB scale perturbation has $|\zeta_h| \sim 5$, h_{end} should be moved at least ~ 5 e -folds backwards along its trajectory.

For the BSM potential, with the scalar mass set near h_{rescue} to ensure that no regions ever fall into the unrescuable region, there are two situations which are compatible with the CMB and one which is not.

If h_{end} does not approach h_{rescue} , then we return to the linear theory, Standard Model result, where Higgs fluctuations on CMB scales do not lead to large curvature fluctuations since ρ_h/H^2 decreases sharply in Fig. 4.5 and predicts ζ through Eq. (4.60).

Conversely, if the Higgs travels far down the unstable region early in inflation, the Higgs becomes uniform in the potential well during inflation and thus leads to no curvature perturbations after inflation. In this region as well the Higgs instability is compatible with the CMB.

It is only in the region of parameter space near our fiducial model, where the background Higgs h_{end} approaches but does not reach the minimum induced by the BSM massive scalar near h_{rescue} , that Higgs fluctuations on CMB scales can be converted to curvature fluctuations which are large enough to disturb the CMB.

To avoid this possibility completely, one should set the mass m_s of the scalar field to be slightly smaller than h_{rescue} . In particular to achieve $|\zeta| \lesssim 10^{-5}$, using Eq. (4.89) and Eq. (4.91), one requires

$$\frac{m_s}{h_{\text{rescue}}} \lesssim \left(\frac{1}{100}\right)^{\frac{1}{4}} \lesssim \frac{1}{3}. \quad (4.92)$$

In summary, it is only a special class of Higgs criticality scenarios where the parameters are arranged so that the regions of the universe fluctuate near the edge of rescuable instability which would be testable in the CMB and even that class cannot form PBHs as the majority of the dark matter, nor generate second-order gravitational waves at an amplitude detectable with LISA.

4.4 Discussion

We have definitively shown that the dark matter is not composed exclusively of primordial black holes produced by the collapse of density perturbations generated by a spectator Higgs field during inflation.

While a spectator Higgs evolving on the unstable side of its potential can generate large Higgs fluctuations on PBH scales, even larger Higgs fluctuations are produced on CMB scales. This result is obtained using linear perturbation theory during inflation, which we show holds even though the CMB modes cross the horizon at an epoch when the Higgs' per e -fold classical roll is smaller than its per e -fold stochastic motion, because the stochastic motion is incoherent and does not backreact on the Higgs background.

Inflation ends well after all relevant modes have crossed the horizon, and when reheating occurs the Higgs potential is uplifted by the interaction between the Higgs and the thermal bath. If the Higgs is rescued from the unstable region by this thermal uplift then the Higgs redshifts as, and eventually decays to, radiation. The CMB and PBH modes are superhorizon at these epochs and evolve in the same way through these processes. Therefore if the Higgs

fluctuations are converted into sufficiently large curvature perturbations such that enough PBHs were produced to explain the dark matter, CMB constraints would necessarily be violated.

In fact though, a sufficient abundance of PBHs is never produced and CMB constraints are only violated in cases of near criticality. We first showed that this is true under the assumption that linear theory holds through reheating, where we correct an error in local energy conservation made in the literature.

We then showed that linear theory is violated because the model requires the Higgs to be as close as possible to the maximum value beyond which it cannot be rescued at reheating. This criticality condition leads typical perturbations to evolve nonlinearly, and using the nonlinear δN formalism we also show that the Standard Model Higgs, regardless of fine-tuning or anthropic arguments, can never produce enough PBHs to be the DM. Modifying the Higgs potential at large field values can eliminate fine-tuning or anthropic issues, but cannot enhance curvature perturbations significantly enough to explain the dark matter.

CHAPTER 5

CONCLUSION

Primordial black holes are extreme objects that push inflationary theory to its limits. Over the course of this thesis, we have endeavored to use the breaking points to understand the more general principles that govern the physics of inflation.

To do so, we had to take a wide variety of theoretical approaches. We used the geometric quantities of ADM to build an effective field theory for perturbations, and we solved equations of motion using generalized slow-roll; We computed non-Gaussianities with in-in, and we evolved regions nonlinearly using $\delta\mathcal{N}$; And we explored the arcane secrets of perturbation theory in various gauges and reference frames.

Along the way we broached some fascinating topics, like the history of the Higgs in our universe, while other ideas were touched on only briefly, like reheating, stochasticity, tensor modes, and true multi-field effects. These, and further frontiers, are certainly worthy of more detailed study.

Yet in the end, only observation can bring definitive answers in our quest to understand inflation. Guided by theory, existing and upcoming experiments around the world are searching for primordial non-Gaussianity, hunting for primordial black holes, and hoping to catch inflationary gravitational waves. With these clues, or as yet unexpected ones, we will uncover the birth of our universe.

REFERENCES

- [1] Samuel Passaglia and Wayne Hu. Scalar Bispectrum Beyond Slow-Roll in the Unified EFT of Inflation. *Phys. Rev.*, D98(2):023526, 2018.
- [2] Héctor Ramírez, Samuel Passaglia, Hayato Motohashi, Wayne Hu, and Olga Mena. Reconciling tensor and scalar observables in G-inflation. *JCAP*, 1804(04):039, 2018.
- [3] Samuel Passaglia, Wayne Hu, and Hayato Motohashi. Primordial black holes and local non-Gaussianity in canonical inflation. *Phys. Rev.*, D99(4):043536, 2019.
- [4] Samuel Passaglia, Wayne Hu, and Hayato Motohashi. Primordial Black Holes as Dark Matter through Higgs Criticality. *Phys. Rev. D*, 101:123523, 2020.
- [5] Arno A. Penzias and Robert Woodrow Wilson. A Measurement of excess antenna temperature at 4080-Mc/s. *Astrophys. J.*, 142:419–421, 1965.
- [6] George F. Smoot et al. Structure in the COBE differential microwave radiometer first year maps. *Astrophys. J. Lett.*, 396:L1–L5, 1992.
- [7] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters.
- [8] Hayato Motohashi and Wayne Hu. Primordial Black Holes and Slow-Roll Violation. *Phys. Rev.*, D96(6):063503, 2017.
- [9] Paulo Montero-Camacho, Xiao Fang, Gabriel Vasquez, Makana Silva, and Christopher M. Hirata. Revisiting constraints on asteroid-mass primordial black holes as dark matter candidates. 2019.
- [10] Hayato Motohashi and Wayne Hu. Generalized Slow Roll in the Unified Effective Field Theory of Inflation. *Phys. Rev.*, D96(2):023502, 2017.
- [11] Xian Gao. Unifying framework for scalar-tensor theories of gravity. *Phys. Rev.*, D90:081501, 2014.
- [12] Xian Gao. Hamiltonian analysis of spatially covariant gravity. *Phys. Rev.*, D90:104033, 2014.
- [13] Petr Horava. Quantum Gravity at a Lifshitz Point. *Phys. Rev.*, D79:084008, 2009.
- [14] D. Blas, O. Pujolas, and S. Sibiryakov. On the Extra Mode and Inconsistency of Horava Gravity. *JHEP*, 10:029, 2009.
- [15] Diego Blas, Oriol Pujolas, and Sergey Sibiryakov. Models of non-relativistic quantum gravity: The Good, the bad and the healthy. *JHEP*, 04:018, 2011.
- [16] Jibril Ben Achour, Marco Crisostomi, Kazuya Koyama, David Langlois, Karim Noui, and Gianmassimo Tasinato. Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order. *JHEP*, 12:100, 2016.

- [17] David Langlois, Michele Mancarella, Karim Noui, and Filippo Vernizzi. Effective Description of Higher-Order Scalar-Tensor Theories. *JCAP*, 1705(05):033, 2017.
- [18] Hayato Motohashi and Wayne Hu. Effective field theory of degenerate higher-order inflation. *Phys. Rev. D*, 101:083531, 2020.
- [19] Hayato Motohashi, Teruaki Suyama, and Kazufumi Takahashi. Fundamental theorem on gauge fixing at the action level. *Phys. Rev.*, D94(12):124021, 2016.
- [20] Juan Martin Maldacena. Non-Gaussian features of primordial fluctuations in single field inflationary models. *JHEP*, 05:013, 2003.
- [21] Xingang Chen, Min-xin Huang, Shamit Kachru, and Gary Shiu. Observational signatures and non-Gaussianities of general single field inflation. *JCAP*, 0701:002, 2007.
- [22] Enrico Pajer, Guilherme L. Pimentel, and Jaap V. S. Van Wijck. The Conformal Limit of Inflation in the Era of CMB Polarimetry. *JCAP*, 1706(06):009, 2017.
- [23] Frederico Arroja and Takahiro Tanaka. A note on the role of the boundary terms for the non-Gaussianity in general k-inflation. *JCAP*, 1105:005, 2011.
- [24] Gerasimos Rigopoulos. Gauge invariance and non-Gaussianity in Inflation. *Phys. Rev.*, D84:021301, 2011.
- [25] Clare Burrage, Raquel H. Ribeiro, and David Seery. Large slow-roll corrections to the bispectrum of noncanonical inflation. *JCAP*, 1107:032, 2011.
- [26] Sebastien Renaux-Petel. On the redundancy of operators and the bispectrum in the most general second-order scalar-tensor theory. *JCAP*, 1202:020, 2012.
- [27] David Seery and James E. Lidsey. Primordial non-Gaussianities from multiple-field inflation. *JCAP*, 0509:011, 2005.
- [28] Peter Adshead, Wayne Hu, Cora Dvorkin, and Hiranya V. Peiris. Fast Computation of Bispectrum Features with Generalized Slow Roll. *Phys. Rev.*, D84:043519, 2011.
- [29] Clifford Cheung, A. Liam Fitzpatrick, Jared Kaplan, and Leonardo Senatore. On the consistency relation of the 3-point function in single field inflation. *JCAP*, 0802:021, 2008.
- [30] Paolo Creminelli and Matias Zaldarriaga. Single field consistency relation for the 3-point function. *JCAP*, 0410:006, 2004.
- [31] Paolo Creminelli, Guido D’Amico, Marcello Musso, and Jorge Norena. The (not so) squeezed limit of the primordial 3-point function. *JCAP*, 1111:038, 2011.
- [32] Peter Adshead, Wayne Hu, and Vinícius Miranda. Bispectrum in Single-Field Inflation Beyond Slow-Roll. *Phys. Rev.*, D88(2):023507, 2013.

- [33] Clifford Cheung, Paolo Creminelli, A. Liam Fitzpatrick, Jared Kaplan, and Leonardo Senatore. The Effective Field Theory of Inflation. *JHEP*, 03:014, 2008.
- [34] Richard P. Woodard. Ostrogradsky’s theorem on Hamiltonian instability. *Scholarpedia*, 10(8):32243, 2015.
- [35] Adam R. Solomon and Mark Trodden. Higher-derivative operators and effective field theory for general scalar-tensor theories. *JCAP*, 1802(02):031, 2018.
- [36] Gregory Walter Horndeski. Second-order scalar-tensor field equations in a four-dimensional space. *Int. J. Theor. Phys.*, 10:363–384, 1974.
- [37] Jérôme Gleyzes, David Langlois, Federico Piazza, and Filippo Vernizzi. Exploring gravitational theories beyond Horndeski. *JCAP*, 1502:018, 2015.
- [38] Ryotaro Kase and Shinji Tsujikawa. Effective field theory approach to modified gravity including Horndeski theory and Hořava–Lifshitz gravity. *Int. J. Mod. Phys.*, D23(13):1443008, 2014.
- [39] Antonio De Felice and Shinji Tsujikawa. Shapes of primordial non-Gaussianities in the Horndeski’s most general scalar-tensor theories. *JCAP*, 1303:030, 2013.
- [40] Antonio De Felice and Shinji Tsujikawa. Inflationary non-Gaussianities in the most general second-order scalar-tensor theories. *Phys. Rev.*, D84:083504, 2011.
- [41] Xian Gao and Daniele A. Steer. Inflation and primordial non-Gaussianities of ‘generalized Galileons’. *JCAP*, 1112:019, 2011.
- [42] Matteo Fasiello and Sébastien Renaux-Petel. Non-Gaussian inflationary shapes in G^3 theories beyond Horndeski. *JCAP*, 1410(10):037, 2014.
- [43] Steven Weinberg. Quantum contributions to cosmological correlations. *Phys. Rev.*, D72:043514, 2005.
- [44] Peter Adshead, Richard Easther, and Eugene A. Lim. The ‘in-in’ Formalism and Cosmological Perturbations. *Phys. Rev.*, D80:083521, 2009.
- [45] Peter Adshead, Richard Easther, and Eugene A. Lim. Cosmology With Many Light Scalar Fields: Stochastic Inflation and Loop Corrections. *Phys. Rev.*, D79:063504, 2009.
- [46] Ewan D. Stewart. The Spectrum of density perturbations produced during inflation to leading order in a general slow roll approximation. *Phys. Rev.*, D65:103508, 2002.
- [47] Jeongyeol Choe, Jinn-Ouk Gong, and Ewan D. Stewart. Second order general slow-roll power spectrum. *JCAP*, 0407:012, 2004.
- [48] Cora Dvorkin and Wayne Hu. Generalized Slow Roll for Large Power Spectrum Features. *Phys. Rev.*, D81:023518, 2010.

- [49] Wayne Hu. Generalized Slow Roll for Non-Canonical Kinetic Terms. *Phys. Rev.*, D84:027303, 2011.
- [50] Kenji Kadota, Scott Dodelson, Wayne Hu, and Ewan D. Stewart. Precision of inflaton potential reconstruction from CMB using the general slow-roll approximation. *Phys. Rev.*, D72:023510, 2005.
- [51] Vinicius Miranda, Wayne Hu, and Peter Adshead. Warp Features in DBI Inflation. *Phys. Rev.*, D86:063529, 2012.
- [52] Peter Adshead and Wayne Hu. Fast Computation of First-Order Feature-Bispectrum Corrections. *Phys. Rev.*, D85:103531, 2012.
- [53] Junko Ohashi and Shinji Tsujikawa. Potential-driven Galileon inflation. *JCAP*, 1210:035, 2012.
- [54] Vinicius Miranda, Wayne Hu, Chen He, and Hayato Motohashi. Nonlinear Excitations in Inflationary Power Spectra. *Phys. Rev.*, D93(2):023504, 2016.
- [55] Daniel Baumann and Daniel Green. Equilateral Non-Gaussianity and New Physics on the Horizon. *JCAP*, 1109:014, 2011.
- [56] Leonardo Senatore, Kendrick M. Smith, and Matias Zaldarriaga. Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data. *JCAP*, 1001:028, 2010.
- [57] Nicola Bartolo, Matteo Fasiello, Sabino Matarrese, and Antonio Riotto. Large non-Gaussianities in the Effective Field Theory Approach to Single-Field Inflation: the Trispectrum. *JCAP*, 1009:035, 2010.
- [58] Nicola Bartolo, Dario Cannone, and Sabino Matarrese. The Effective Field Theory of Inflation Models with Sharp Features. *JCAP*, 1310:038, 2013.
- [59] William H. Kinney. Horizon crossing and inflation with large eta. *Phys. Rev.*, D72:023515, 2005.
- [60] F. K. Manasse and C. W. Misner. Fermi normal coordinates and some basic concepts in differential geometry. *J. Math. Phys.*, 4(6):735, 1963.
- [61] Leonardo Senatore and Matias Zaldarriaga. The constancy of ζ in single-clock Inflation at all loops. *JHEP*, 09:148, 2013.
- [62] Leonardo Senatore and Matias Zaldarriaga. A Note on the Consistency Condition of Primordial Fluctuations. *JCAP*, 1208:001, 2012.
- [63] Enrico Pajer, Fabian Schmidt, and Matias Zaldarriaga. The Observed Squeezed Limit of Cosmological Three-Point Functions. *Phys. Rev.*, D88(8):083502, 2013.

- [64] Mohammad Hossein Namjoo, Hassan Firouzjahi, and Misao Sasaki. Violation of non-Gaussianity consistency relation in a single field inflationary model. *EPL*, 101(3):39001, 2013.
- [65] Jerome Martin, Hayato Motohashi, and Teruaki Suyama. Ultra Slow-Roll Inflation and the non-Gaussianity Consistency Relation. *Phys. Rev.*, D87(2):023514, 2013.
- [66] Alexei A. Starobinsky. Multicomponent de Sitter (inflationary) stages and the generation of perturbations. *JETP Lett.*, 42:152–155, 1985. *Pis'ma v ZhETF*. 42, 124 (1985).
- [67] D. S. Salopek and J. R. Bond. Nonlinear evolution of long wavelength metric fluctuations in inflationary models. *Phys. Rev.*, D42:3936–3962, 1990.
- [68] Misao Sasaki and Ewan D. Stewart. A General analytic formula for the spectral index of the density perturbations produced during inflation. *Prog. Theor. Phys.*, 95:71–78, 1996.
- [69] Naonori S. Sugiyama, Eiichiro Komatsu, and Toshifumi Futamase. δN formalism. *Phys. Rev.*, D87(2):023530, 2013.
- [70] Xingang Chen, Hassan Firouzjahi, Eiichiro Komatsu, Mohammad Hossein Namjoo, and Misao Sasaki. In-in and δN calculations of the bispectrum from non-attractor single-field inflation. *JCAP*, 1312:039, 2013.
- [71] Yi-Fu Cai, Xingang Chen, Mohammad Hossein Namjoo, Misao Sasaki, Dong-Gang Wang, and Ziwei Wang. Revisiting non-Gaussianity from non-attractor inflation models. *JCAP*, 1805(05):012, 2018.
- [72] Chris Pattison, Vincent Vennin, Hooshyar Assadollahi, and David Wands. Quantum diffusion during inflation and primordial black holes. *JCAP*, 1710(10):046, 2017.
- [73] Giovanni Cabass, Enrico Pajer, and Fabian Schmidt. How Gaussian can our Universe be? *JCAP*, 1701(01):003, 2017.
- [74] P. A. R. Ade et al. Planck 2015 results. XVII. Constraints on primordial non-Gaussianity. *Astron. Astrophys.*, 594:A17, 2016.
- [75] Juan García-Bellido and Ester Ruiz Morales. Primordial black holes from single field models of inflation. *Phys. Dark Univ.*, 18:47–54, 2017.
- [76] Chris Pattison, Vincent Vennin, Hooshyar Assadollahi, and David Wands. The attractive behaviour of ultra-slow-roll inflation. 2018.
- [77] Wayne Hu and Austin Joyce. Separate Universes beyond General Relativity. *Phys. Rev.*, D95(4):043529, 2017.

- [78] Michele Cicoli, Victor A. Diaz, and Francisco G. Pedro. Primordial Black Holes from String Inflation. *JCAP*, 1806(06):034, 2018.
- [79] Ogan Özsoy, Susha Parameswaran, Gianmassimo Tasinato, and Ivonne Zavala. Mechanisms for Primordial Black Hole Production in String Theory. *JCAP*, 1807(07):005, 2018.
- [80] Christian T. Byrnes, Philippa S. Cole, and Subodh P. Patil. Steepest growth of the power spectrum and primordial black holes.
- [81] Vicente Atal and Cristiano Germani. The role of non-gaussianities in Primordial Black Hole formation.
- [82] Ryo Saito, Jun'ichi Yokoyama, and Ryo Nagata. Single-field inflation, anomalous enhancement of superhorizon fluctuations, and non-Gaussianity in primordial black hole formation. *JCAP*, 0806:024, 2008.
- [83] Kendrick M. Smith and Marilena LoVerde. Local stochastic non-Gaussianity and N-body simulations. *Journal of Cosmology and Astro-Particle Physics*, 2011:009, November 2011.
- [84] Christian T. Byrnes, Edmund J. Copeland, and Anne M. Green. Primordial black holes as a tool for constraining non-Gaussianity. *Phys. Rev.*, D86:043512, 2012.
- [85] Simone Ferraro and Kendrick M. Smith. Using large scale structure to measure f_{NL} , g_{NL} and τ_{NL} . *Phys. Rev.*, D91(4):043506, 2015.
- [86] Sam Young and Christian T. Byrnes. Signatures of non-gaussianity in the isocurvature modes of primordial black hole dark matter. *JCAP*, 1504(04):034, 2015.
- [87] Yuichiro Tada and Shuichiro Yokoyama. Primordial black holes as biased tracers. *Phys. Rev.*, D91(12):123534, 2015.
- [88] G. Franciolini, A. Kehagias, S. Matarrese, and A. Riotto. Primordial Black Holes from Inflation and non-Gaussianity. *JCAP*, 1803(03):016, 2018.
- [89] Christian Gross, Antonello Polosa, Alessandro Strumia, Alfredo Urbano, and Wei Xue. Dark Matter in the Standard Model? *Phys. Rev.*, D98(6):063005, 2018.
- [90] Jose R. Espinosa, Gian F. Giudice, Enrico Morgante, Antonio Riotto, Leonardo Senatore, Alessandro Strumia, and Nikolaos Tetradis. The cosmological Higgstory of the vacuum instability. *JHEP*, 09:174, 2015.
- [91] J. R. Espinosa, D. Racco, and A. Riotto. Cosmological Signature of the Standard Model Higgs Vacuum Instability: Primordial Black Holes as Dark Matter. *Phys. Rev. Lett.*, 120(12):121301, 2018.

- [92] J. R. Espinosa, D. Racco, and A. Riotto. Primordial Black Holes from Higgs Vacuum Instability: Avoiding Fine-tuning through an Ultraviolet Safe Mechanism. *Eur. Phys. J.*, C78(10):806, 2018.
- [93] Dario Buttazzo, Giuseppe Degrandi, Pier Paolo Giardino, Gian F. Giudice, Filippo Sala, Alberto Salvio, and Alessandro Strumia. Investigating the near-criticality of the Higgs boson. *JHEP*, 12:089, 2013.
- [94] ATLAS, CDF, CMS, and D0 Collaborations. First combination of Tevatron and LHC measurements of the top-quark mass. 2014.
- [95] Vardan Khachatryan et al. Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and 8 TeV. *Phys. Rev.*, D93(7):072004, 2016.
- [96] Morad Aaboud et al. Measurement of the top quark mass in the $t\bar{t} \rightarrow$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data. *Phys. Lett.*, B761:350–371, 2016.
- [97] Tevatron Electroweak Working Group and T. Aaltonen. Combination of CDF and D0 results on the mass of the top quark using up 9.7 fb^{-1} at the Tevatron. 2016.
- [98] Hiroko Niikura et al. Microlensing constraints on primordial black holes with the Subaru/HSC Andromeda observation.
- [99] David Wands, Karim A. Malik, David H. Lyth, and Andrew R. Liddle. A New approach to the evolution of cosmological perturbations on large scales. *Phys. Rev.*, D62:043527, 2000.
- [100] Joan Elias-Miro, Jose R. Espinosa, Gian F. Giudice, Hyun Min Lee, and Alessandro Strumia. Stabilization of the Electroweak Vacuum by a Scalar Threshold Effect. *JHEP*, 06:031, 2012.
- [101] Alexei A. Starobinsky. Stochastic de Sitter (inflationary) stage in the early universe. *Lect. Notes Phys.*, 246:107–126, 1986.
- [102] Alexei A. Starobinsky and Junichi Yokoyama. Equilibrium state of a selfinteracting scalar field in the De Sitter background. *Phys. Rev.*, D50:6357–6368, 1994.
- [103] Hayato Motohashi and Wayne Hu. Running from Features: Optimized Evaluation of Inflationary Power Spectra. *Phys. Rev.*, D92(4):043501, 2015.
- [104] José Ramón Espinosa, Davide Racco, and Antonio Riotto. A Cosmological Signature of the SM Higgs Instability: Gravitational Waves. *JCAP*, 1809(09):012, 2018.
- [105] David H. Lyth, Karim A. Malik, and Misao Sasaki. A General proof of the conservation of the curvature perturbation. *JCAP*, 0505:004, 2005.