ABSTRACT

Title of Dissertation: NEW MECHANISMS IN INFLATIONARY COSMOLOGY

Edward Joseph Broadberry
Doctor of Philosophy, 2026

Dissertation Directed by: Professor Raman Sundrum
Department of Physics

Cosmological Collider Physics (CCP), a mechanism in which heavy particles produced during inflation leave an observable footprint in primordial non-Gaussianities, carries the prospect of probing physics at scales far higher than any terrestrial collider. In this thesis supersymmetric (SUSY) orbifold grand unified theories (GUTs) are identified as a natural candidate predicting new particles which could be detected in CCP. The high scale of unification naively places SUSY orbifold GUTs out of reach, but there exists a "chemical potential" extension of CCP that increases the observable mass range.

A chemical potential mechanism is developed for spin-1 particles coupled to inflation, and allows for the possible detection of heavy gauge bosons from orbifold SUSY GUTs in future large scale structure surveys, and in 21-cm tomography. The mechanism requires a complex vector that is neutral under the standard model, eliminating the minimal $SU(5)$ and $SO(10)$ GUTs,

while leaving open the classic trinification scenario. The chemical potential mechanism is then extended to a pair of charged scalars appearing in a loop. The spectral decomposition is utilised to make the loop process analytically calculable, and is then applied to the detection of heavy Higgs triplets in the $SU(5)$ scenario.

In the final chapter a perturbative model of warm inflation is developed. Warm inflation is a variant of inflation in which primordial fluctuations are thermal, providing a foil to the standard lore that the origin of structure in our universe is due to quantum fluctuations. The chemical potential models, suitably modified, are the key to maintaining theoretical control in this framework, something which has long been a problem. The fluctuation-dissipation theorem, an important theorem in warm inflation, is modified by the chemical potential, and many of the results in the literature must be modified to take this change into account. The predictions of warm inflation are consistent with the CMB power spectrum, and primordial NGs are again the key to obtaining more information about the dynamics of inflation.

NEW MECHANISMS IN INFLATIONARY COSMOLOGY

by

Edward Joseph Broadberry

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2026

Advisory Committee:
Professor Raman Sundrum, Chair/Advisor
Professor Anson Hook
Professor Thomas Cohen
Professor Kaustubh Agashe
Professor Richard Wentworth

Dedication

посвящается Марии, моей очень замечательной и любимой жене.

Acknowledgments

In my time at the University of Maryland I have had the pleasure of working with many great physicists in the Maryland Centre for Fundamental Physics, MCFP. The person who had the most influence on my development as a theoretical physicist is my advisor Raman Sundrum. I will always be grateful for the many discussions we have had over the years, for his qualitative insights, and ability to get straight to the heart of a problem. I hope to carry these lessons with me in my future career. Beyond my professional development I want to thank him for always taking a genuine interest in my well-being.

I also want to specifically thank professor Anson Hook for his many invaluable physical insights over the course of our collaborations. I am also thankful to my other collaborators: Arushi Bodas, Zhaohui Xu, Sagnik Mondal, Sagar Airen, Lorenzo Ricci, Gustavo Marques-Tavares, and Saurav Das who have all had a distinct impact on my understanding of and approach to physics.

In addition to those mentioned above I want to thank the professors: Kaustubh Agashe, Theodore Jacobson, Thomas Cohen, and Zackaria Chacko, the HEP-th post docs I overlapped with: Jae Hyeok Chang, Manuel Buen-Abad, Abhishek Banerjee, and William DeRocco, and fellow HEP-th grad students: Abhish Dev, Clayton Ristow, Dawid Brzeminski, Deepak Sathyan, Batoul Banihashemi, Rodrigo Andrade e Silva, Ina Flood, Shourya Mukherjee, Rajrupa Mondal, and Kushan Panchel for making the MCFP a fun and collaborative environment, and for many useful discussions over the years.

Outside of my professional development I thank my housemates: Mike Weiner, Deepak Sathyan, Saurabh Kadam, Chung-Chun Hsieh, and Ernst Wilhelm Grunow VI for many fond memories and games of settlers of Catan. I also could not have got through these years without

the many good friends I have made at Maryland, of which there are too many to name, and the support of my friends and family back home.

Finally, this thesis is dedicated to my beautiful and brilliant wife Mariia Kharchenko. Thank you for bringing so much joy and love into my life, putting up with many late nights, being the guinea pig for all of my talks, and welcoming me into your family. I am so excited for what the future holds for us.

I would also like to thank the maryland centre for fundamental physics for their support, and the national science foundation for supporting this work with the grant PHY-2210361.

Table of Contents

Dedication ii

Acknowledgements iii

Table of Contents v

List of Tables ix

List of Figures x

List of Published Work xi

Chapter 1: Introduction to Cosmological Collider Physics 1
- 1.1 Why Inflation? . . . 1
- 1.2 Scalar Fields in an Expanding Spacetime . . . 6
- 1.3 The CMB Power Spectrum . . . 13
- 1.4 Non-Gaussianities . . . 17
- 1.5 Cosmological Collider Physics . . . 23

Chapter 2: Searching for GUTs in Cosmological Data 33
- 2.1 The Chemical Potential Mechanism . . . 33
- 2.2 Chemical Potential for Spin-1 . . . 36
- 2.3 Calculating the Bispectrum . . . 38
- 2.4 Results & Conclusions . . . 45
- 2.5 Orbifold Grand Unified Theories . . . 46
- 2.6 Coupling to Inflation . . . 50

Chapter 3: New Physics with Non-Zero Charge 54
- 3.1 Chemical Potential for Charged Scalars . . . 55
- 3.2 The Spectral Decomposition . . . 55
- 3.3 Extracting the Leading non-analytic Signal . . . 58
- 3.4 Results & Constraints . . . 62
- 3.5 Embedding in orbifold SUSY GUTs . . . 65

Chapter 4: A Calculable Model of Warm Inflation 70
- 4.1 The Framework of Warm Inflation . . . 71
- 4.2 Failure is a Dish Best Served Warm . . . 73
- 4.3 A Working Model . . . 77
- 4.4 Brownian Motion . . . 82
- 4.5 The Fluctuation-Dissipation Theorem with a Chemical Potential . . . 86
- 4.6 The Warm Inflation Power Spectrum . . . 90
- 4.7 A Working Warm Inflation Data Point . . . 94

Chapter 5: Conclusion 96

Appendix A: Grand Unification at the Cosmological Collider with Chemical Potential 99
A.1 Introduction . . . 100
A.2 Preliminaries . . . 105
A.2.1 Conventions . . . 106
A.2.2 Recap: chemical potential for spin-0 particles . . . 109
A.3 A Chemical Potential for Tree-Level Spin-1 Exchange . . . 112
A.3.1 Chemical potential in the rotated basis . . . 112
A.3.2 Expansion parameters . . . 113
A.3.3 Other lower dimension interactions . . . 114
A.3.4 Interactions . . . 115
A.4 Non-Gaussianity due to Massive Spin-1 Exchange . . . 116
A.4.1 The stationary phase approximation . . . 116
A.4.2 Other subdominant diagrams . . . 122
A.4.3 Apparent late-time divergences . . . 123
A.4.4 Correction to the power spectrum . . . 124
A.5 Signal Strength and Key Features . . . 126
A.5.1 Signal size . . . 126
A.5.2 Inferring the mass and the chemical potential . . . 127
A.5.3 Constraining the spin . . . 129
A.6 Embedding the Chemical Potential within Orbifold Unification . . . 131
A.6.1 Trinification . . . 131
A.6.2 Realising the chemical potential coupling . . . 132
A.7 Discussion . . . 136
A.8 Appendix . . . 137
A.8.1 Mode functions . . . 137
A.8.2 The full set of interactions . . . 139
A.8.3 Detailed stationary phase calculation . . . 141
A.8.4 Cancellation of late-time divergences . . . 145

Appendix B: Charged Loops at the Cosmological Collider with Chemical Potential 149
B.1 Introduction . . . 150
B.2 Preliminaries . . . 155
B.2.1 Primordial non-Gaussianity . . . 155
B.2.2 In-in formalism and mode functions . . . 158
B.3 Charged-pair Chemical Potential Model and Spectral Decomposition . . . 161
B.3.1 Chemical potential for a pair of charged scalars . . . 162
B.3.2 One-loop diagram under spectral representation . . . 164
B.3.3 Central result for the signal . . . 171
B.4 Systematic Computation: Tree-level Model Revisited . . . 176
B.4.1 Perturbative expansion under late-time-finite basis . . . 177
B.4.2 Contributions from B_{-+} . . . 182
B.4.3 Contributions from B_{++} . . . 185
B.5 Systematic Computation: One-loop Model . . . 189

B.5.1 Signal from the threshold limit . . . 189
B.5.2 Resonance backgrounds . . . 192
B.5.3 Beyond stationary phase approximation . . . 197
B.5.4 Combinatorial backgrounds from permuted diagrams . . . 203
B.5.5 Higher order corrections at one-loop . . . 207
B.6 Constraints and Signal Size . . . 213
B.6.1 Constraints on λ and beyond de Sitter . . . 213
B.6.2 Constraints on α . . . 215
B.6.3 Signal size . . . 217
B.7 Application: Colored Higgs in SUSY Orbifold GUTs . . . 219
B.7.1 SUSY inflation and the fifth dimension . . . 220
B.7.2 Realizing the chemical potential . . . 224
B.8 Discussion and Conclusions . . . 227
B.9 Appendix . . . 229
B.9.1 Integrals and Identities Involving Special Functions . . . 229
B.9.2 Asymptotic Expansions of $\widetilde{\mathbf{F}}^s_{i\mu}(p)$ . . . 231
B.9.3 Spectral Density in Non-relativistic Limit of de Sitter Spacetime . . . 233
B.9.4 Exact Expression of $B_{++,\mathrm{A}}$ . . . 236
B.9.5 Full Evaluation of the One-loop Model . . . 240
B.9.6 Corrections to the Power Spectrum . . . 248

Appendix C: Warm Inflation with Pseudo-scalar Couplings 252
C.1 Introduction . . . 253
C.2 A Simple Toy Model . . . 258
C.2.1 Response of the Thermal Bath . . . 260
C.2.2 Thermal Friction . . . 263
C.2.3 Two Point Function . . . 265
C.3 Other Examples of Chemical Potentials . . . 267
C.3.1 Chemical Potential Couplings . . . 268
C.3.2 Symmetry Breaking Couplings . . . 270
C.3.3 Topological Axion Couplings . . . 271
C.4 A warm ϕ^4 example . . . 279
C.4.1 The Background Solution . . . 279
C.4.2 Power Spectrum . . . 281
C.4.3 Thermal Potential . . . 282
C.4.4 The Data Point . . . 285
C.5 Conclusions and Further Work . . . 288
C.6 Appendix . . . 289
C.6.1 A Lightning Fast Review of the Thermal Field Double . . . 289
C.6.2 A lightning Fast Review of the In-in Formalism . . . 294
C.6.3 Parametric Resonance . . . 296
C.6.4 The Fluctuation Dissipation Theorem . . . 298
C.6.5 From Thermal Fluctuations to Inflaton Fluctuations . . . 301
C.6.6 Linear Response Theory and the Adiabatic Approximation . . . 308

Bibliography 311

List of Tables

3.1 The boundary conditions for various fields in the orbifold SUSY GUT, the $U(1)_R$ charges, and the mass spectrum of the KK modes. 68

B.1 The positions of the stationary phase points in the tree-level part of the loop integral. 183
B.2 The positions of poles in the upper half plane of the loop integrand. 201
B.3 Repeat of table 3.1. 222
B.4 A summary of the sources of poles in the upper half plane for the various diagrams contributing to the loop amplitude. 242
B.5 A list of exponential suppression factors for all the contributions to the loop amplitude. 243

List of Figures

1.1 The conformal diagram for the standard big bang cosmology. . . . 4
1.2 A diagram illustrating the general properties of a slow-roll potential. . . . 7
1.3 A diagram demonstrating how quantum fluctuations of the inflaton can lead to inflation ending earlier, or later in different parts of the universe. . . . 9
1.4 A demonstration of how a free field in de-Sitter acts like a time-dependent simple harmonic oscillator, with a demonstration of how the quadratic potential varies over time. . . . 10
1.5 The observed CMB power spectrum. . . . 16
1.6 The triangle formed by the momenta involved in the bispectrum. . . . 18
1.7 The contour that implements to $i\epsilon$ prescription to project onto the interacting vacuum. . . . 21
1.8 The tree level Feynman diagram for the bispectrum in the vanilla model. . . . 25
1.9 Demonstrating that cosmological collider physics is a quantum interference effect by pictorially representing the Feynman diagram expansion as a quantum superposition. . . . 30

2.1 The Feynman diagram for the tree level bispectrum in the spin-1 EFT. . . . 40
2.2 A plot of the maximum possible signal size within theoretical control. . . . 46
2.3 The 1-loop running of gauge couplings in both the SM and the MSSM 47
2.4 A sketch of the running couplings in orbifold SUSY GUTs 49
2.5 A diagram illustrating the orbifold setup. . . . 50

3.1 The loops diagrams contributing to the bispectrum at leading order. . . . 56
3.2 The loops diagrams in the spectral representation. . . . 57
3.3 The de-Sitter spectral density 58
3.4 The loop diagram in the spectral representation after all field redefinitions. . . . 59
3.5 The integrand for the leading contribution containing non-analyticities. . . . 60
3.6 The constraints on the size of the non-analytic signal due to the exchange of two heavy scalars in a loop. . . . 64
3.7 The parameter space for a pair of scalars that transform under a larger non-Abelian group. . . . 65
3.8 The extra-dimensional setup for the orbifold SUSY GUT, with the profile of the lowest lying KK modes shown. . . . 67

4.1 The full slow-roll potential including finite temperature corrections. . . . 80
4.2 Numerical fit to the power spectrum in warm inflation. . . . 93

A.1 Repeat of Figure 2.4 103
A.2 Repeat of figure 1.7 107
A.3 Feynman diagrams for the tree level contribution to the bispectrum due to exchange of a massive vector. . . . 117
A.4 The tree level Feynman diagrams contributing corrections to the power spectrum. 124
A.5 Repeat of figure 2.2 126

A.6 A plot of the non-analytic dependence of the bispectrum as a function of squeezing. 128
A.7 Repeat of figure 2.5 . . . 132

B.1 The momentum triangle for the bispectrum. . . . 158
B.2 The 1PI one-loop subdiagrams that contribute at leading order. . . . 165
B.3 The two types of topologies contributing to the bispectrum at one-loop. . . . 166
B.4 A demonstration of the spectral representation applied to a particular one-loop diagram. . . . 168
B.5 Repeat of figure 3.3 . . . 168
B.6 The integrand, and its magnitude, for the bispectrum in the spectral representation. 172
B.7 The effective tree-level diagrams after applying the spectral decomposition. . . . 178
B.8 The various effective tree-level diagrams contributing to the bispectrum, including the full multiplicity of Schwinger-Keldysh indices. . . . 180
B.9 A comparison of the magnitude of various contributions to the bispectrum. . . . 188
B.10 The one-loop oscillatory signal, and a comparison of various approximate approaches to its calculation. . . . 191
B.11 The oscillation frequency as a function of squeezing, showing that it is possible to separately determine the mass and chemical potential. . . . 192
B.12 The one-loop smooth background coming from the resonance limit . . . 196
B.13 Different contributions to the one-loop bispectrum for various choices of parameters. 198
B.14 The pole structure of the spectral integral in the upper half plane. . . . 201
B.15 The full one-loop bispectrum including all backgrounds for various choices of parameters. . . . 204
B.16 An example of applying a high-pass filter to the full one-loop bispectrum. . . . 206
B.17 A visualisation of the conditions for parametric resonance in terms of solutions to a polynomial equation. . . . 208
B.18 The numerical mode functions demonstrating the validity of our condition to avoid parametric resonance. . . . 211
B.19 Leading order corrections to the power spectrum. . . . 216
B.20 Repeat of giure 3.6 . . . 217
B.21 Repeat of figure 3.7 . . . 218
B.22 Repeat of figure 3.8 . . . 223
B.23 The maximum size of the Wilson coefficients for the F-term induced symmetry breaking operator. . . . 227
B.24 Contours used to define a special function needed for the loop computation. . . . 237
B.25 One-loop corrections to the power spectrum. . . . 250

C.1 Feynman diagrams for helicity violating scattering processes in models with an ALP coupled to an Abelian gauge theory with light fermions. . . . 272
C.2 The three-loop vacuum bubble which gives the leading-order contribution to the effective potential. . . . 283
C.3 Numerical fits for the power spectrum in warm inflation. . . . 305

Citations to Published Work

The main text of this dissertation is a summary of the following three papers I have either published, or added to ArXiv as preprints, with various collaborators over the course of my PhD:

- Chapter 2 & Appendix A - “Grand Unification at the Cosmological Collider with Chemical Potential" A. Bodas, E. Broadberry, R. Sundrum JHEP 01 (2025) 115.
- Chapter 3 & Appendix B - “Charged Loops at the Cosmological Collider with Chemical Potential", A. Bodas, E. Broadberry, R. Sundrum, Z. Xu, JHEP 01 (2026) 083
- Chapter 4 & Appendix C - “Warm Inflation with Pseudo-Scalar Couplings" - E. Broadberry, A. Hook, S. Mondal, https://arxiv.org/abs/2505.07943 (undergoing peer review)

It should be noted that all of the collaborators mentioned above had a role in the writing of the respective papers, which have been included in their original form as appendices to this thesis.

Chapter 1

Introduction to Cosmological Collider Physics

In the first chapter I will outline the reasons we believe the universe underwent a period of exponentially rapid expansion at early times. I will then show how this expectation naturally leads us to introduce a scalar field, called the inflaton, whose dynamics are responsible for the expansion. We will then see that inflation also explains the origin of structure in the universe as originating from tiny quantum fluctuations of the inflaton. In chapter 4 we will present the case for a competing hypothesis. In the second scenario the inflaton receives random thermal kicks, analogous to the fluctuations of a pollen grain in Brownian motion, which gives a "classical" origin for structure.

I will then argue that Non-Gaussianities (NGs) probe the interactions of the inflaton during inflation. We will see that NGs provide the tantalising prospect of discovering new particles with masses at energy scales far beyond the reach of any terrestrial collider, a program dubbed "cosmological collider physics" [1]. In the next two chapters we will build on this idea to see that precision cosmology may be utilised to detect particles from "Grand Unified Theories" (GUTs).

1.1 Why Inflation?

The idea of Inflation was first introduced to solve the horizon, flatness, and monopole problems [2]. I will briefly describe the horizon problem, but further details can be found in the following reviews [3, 4]. The horizon problem, which is ultimately a statement about initial conditions, has its root in the observation that on large scales our universe is homogeneous and isotropic. The most general spacetime metric with this property is the Friedmann-Lemaitre-

Robertson-Walker (FLRW) metric [5–8]. The FLRW metric, in the absence of spatial curvature, is given by

$$ds^2 = -dt^2 + a^2(t)d\vec{x}^2 = a^2(\eta)\left(-d\eta^2 + d\vec{x}^2\right), \tag{1.1}$$

where η is the so-called "conformal time". One can use the FLRW metric as an ansatz in Einstein's field equations to arrive at the Friedmann equation

$$H^2 = \frac{8\pi G}{3}\rho, \tag{1.2}$$

where $H = \frac{\dot{a}}{a}$ is the Hubble parameter, and ρ is the energy density. Consider the situation in which the energy density of the universe is that of a perfect fluid, with pressure $p = w\rho$. For matter $w = 0$, and for radiation $w = 1/3$. Naively one would expect that any sensible fluid would satisfy $w > 0$, i.e. if I have a box filled with fluid, the fluid will exert some pressure on the walls of the box which will oppose any effort to compress the fluid. For reasons that will become clear, a constituent satisfying $w > -1/3$ is said to satisfy the strong energy condition (SEC), and this was historically thought to be a condition satisfied by any fluid.

For a perfect fluid with constant w, one can use the covariant conservation of the energy-momentum tensor $\nabla_\mu T^{\mu\nu} = 0$, combined with the Friedmann equation to show that

$$\rho \propto a^{-3(1+w)}; \qquad \ddot{a} \propto -(1+3w). \tag{1.3}$$

Therefore, any fluid satisfying the SEC will lead to $\ddot{a} < 0$. We have a clear picture of the evolution of the universe. At some time, let's say $t = 0$, the scale factor $a(0) = 0$. The universe began in a hot big bang, and has been expanding ever since. The rate of expansion should be

decreasing over time, after all gravity will gravitate, won't it?

Let's think about the implications of this picture. As we evolve back in time the constituents of the universe are compressed into a smaller and smaller volume. Therefore, the expectation is that the average temperature of the "cosmic soup" will increase as we look back in time. When the universe was very young it must have been a high temperature relativistic plasma. As the universe expands it cools, and eventually the universe cools to temperatures $T \sim 0.25\,\text{eV}$[1]. Prior to this time all of the hydrogen in the universe was ionised, and the mean free path of a photon was very short. Then, in a process called "recombination", the free electrons combined with protons to form hydrogen. At this point the mean free path of photons drastically increased to larger than the size of the observable universe, and the photons emitted at this time have been travelling through the universe unobstructed until today. In 1965 these photons, forming the "cosmic microwave background" (CMB), were observed [9]. The CMB is a very good black body, and the CMB photons appear to be coming in all directions from what we call the "last scattering surface".

Why is this a problem? After all, just before the CMB was formed the photons were in thermal equilibrium with the rest of the constituents of the universe, so would you not expect the CMB photons to form a close to perfect black body? The issue here is that the temperature is the same in all directions. In figure 1.1 is a conformal diagram of the early universe, taken from [3]. The y-axis is conformal time, η. A quick glance at the metric in equation (1.1) tells us that light travels at 45° angles in the x-η plane. If we extrapolate back in time from the last scattering surface we see that different directions in the sky have not had enough conformal

[1]It might be surprising to the reader that this temperature is not closer to the ionisation energy of Hydrogen, 13.6 eV. However, at this time there happen to be many more photons in our universe than protons. Even at temperatures lower than 13.6 eV there are enough high energy photons at the tail of the Planck distribution to keep all of the hydrogen ionised.

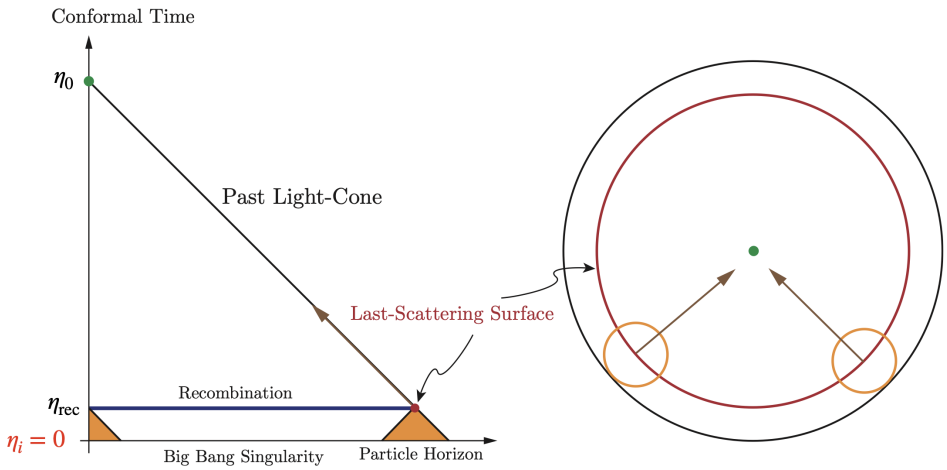


Figure 1.1: In the standard big bang cosmology the last scattering surface of the CMB consists of many regions that are out of causal contact. If all these parts of the universe have never been in causal contact, then why are they all the same temperature? Taken from [3].

time since the universe began to equilibrate. The puzzle of why different directions in the CMB are the same temperature, even though they have never been in causal constant, is called the horizon problem.

How do we solve the horizon problem? Two points during recombination are in causal contact if their past directed light-cones intersect. Given that light travels at 45° in conformal coordinates, the furthest distance between two points that are in causal contact at recombination is the conformal time that has elapsed since the beginning of the universe

$$\eta_{\rm rec} = \int_0^{t_{\rm rec}} \frac{dt'}{a(t')} = \int_0^{a_{\rm rec}} d\log a \left(\frac{1}{aH}\right). \tag{1.4}$$

This quantity is called the “comoving particle horizon". We have changed variables in order to express the comoving horizon in terms of the comoving Hubble radius $(aH)^{-1}$.

We can solve the horizon problem if more conformal time elapsed before recombination.

We know the values of the scale factor, and the Hubble parameter during recombination, but we could consider a different history for these quantities prior to the formation of the CMB. If we posit that there was a period of time in which the Hubble radius was decreasing, then there was a time at which the integrand in equation (1.4) was very large, and we can arrange for there to have been more conformal time before recombination, thus well separated parts of the CMB were in causal contact, and had time to equilibrate to the same temperature. How do we arrange for a shrinking Hubble radius?

$$\frac{d}{dt}(aH)^{-1} \propto -\ddot{a} \propto (1+3w). \tag{1.5}$$

We see that a shrinking Hubble radius requires the universe to accelerate, which implies the universe must consist of a fluid that violates the SEC!

The easiest way to violate the SEC is to add a so-called “cosmological constant", Λ, to Einstein's equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \rightarrow G_{\mu\nu} = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}. \tag{1.6}$$

We can consider the cosmological constant to be an effective fluid, and include it in the stress-energy tensor

$$T^{\Lambda}_{\mu\nu} = \frac{\Lambda}{8\pi G} g_{\mu\nu} \implies T^{\mu}_{\ \nu} = \frac{\Lambda}{8\pi G}\delta^{\mu}_{\nu}. \tag{1.7}$$

In an FLRW universe, by recalling $T^{\mu}_{\ \nu} = \mathrm{diag}(\rho, -p, -p, -p)$, we can identify the energy density and pressure of the cosmological constant

$$\rho_{\Lambda} = \frac{\Lambda}{8\pi G} = -p. \tag{1.8}$$

The cosmological constant can be treated as a fluid with $w = -1$. Notice that the energy density, ρ_Λ, is a constant, and unaffected by the scale factor. For this reason the cosmological constant is often referred to as “vacuum energy". From now on I will use vacuum energy and cosmological constant interchangeably.

Could inflation have been caused by a cosmological constant? The problem with utilising vacuum energy density to drive inflation is that it has no off-switch. We need the universe to expand rapidly, but then we need it to switch over to a radiation dominated universe. While a non-zero vacuum energy density is not relevant for the early universe, I note here that today our universe is accelerating in its expansion [10, 11], and all available data is consistent with a small vacuum energy density that is dominating the energy budget of the universe today [12]. We need something that acts like vacuum energy density, but can subsequently morph into something else.

The idea of slow-roll inflation is to have a potential energy dominated scalar field, called the inflaton, that acts like a slowly varying cosmological constant by rolling along an almost flat part of its potential, as illustrated in figure 1.2. At some point the shape of the potential changes, and the inflaton no longer acts like a cosmological constant. At this point the inflaton then dumps its energy into other degrees of freedom in a process called reheating.

1.2 Scalar Fields in an Expanding Spacetime

In this subsection we will study the dynamics of a scalar field, and determine the conditions required for slow-roll inflation to take place. We will then quantise the fluctuations of the inflaton, and use the free field theory of a scalar field in inflationary spacetime to illustrate the

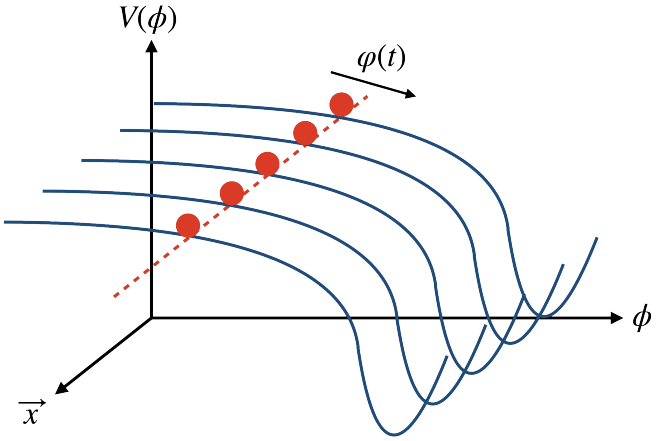


Figure 1.2: An illustration of the slow-roll potential. The inflaton field develops a homogeneous isotropic background in which it slowly-rolls down a slope, acting like cosmological constant. Eventually it reaches the edge of the slope and begins to oscillate about its minimum, ending inflation.

idea of cosmological particle production.

For a more thorough treatment of the classical dynamics of inflation consult [3]. The action of a scalar field coupled to gravity is given by

$$S = S_{\rm EH} + \int d^4x\sqrt{-g}\left(-\frac{1}{2}(\partial_\mu\phi)^2 - V(\phi)\right), \tag{1.9}$$

where $S_{\rm EH}$ is the Einstein-Hilbert action. We can derive the equations of motion from this action, with the ansatz of the FLRW metric, and a homogeneous isotropic background $\phi = \varphi(t)$

$$H^2 = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\varphi}^2 + V(\varphi)\right), \tag{1.10}$$

$$\ddot{\varphi} + 3H\dot{\varphi} + V'(\varphi) = 0, \tag{1.11}$$

where a prime indicates differentiation with respect to φ.

To have a period of rapid expansion, we would like the scalar field to act like a cosmological

constant, so we should be able to ignore the first term on the RHS of the first equation. We would also like this period of expansion to last for a long enough time to solve the horizon problem. We should therefore be able to ignore the first term in the second equation

$$\dot{\varphi}^2 \ll V(\varphi); \qquad \ddot{\varphi} \ll 3H\dot{\varphi}. \tag{1.12}$$

We can reformulate these two conditions in terms of the so-called "slow-roll parameters", which are conditions on the form of the potential

$$\epsilon_V = \frac{M_{\rm pl}^2}{2}\left(\frac{V'}{V}\right)^2 \ll 1, \tag{1.13}$$

$$\eta_V = M_{\rm pl}^2\frac{V''}{V}, \quad |\eta_V| \ll 1, \tag{1.14}$$

where $M_{\rm pl}^2 = 1/8\pi G$ is the Planck scale. If the slow-roll conditions are satisfied then the Hubble parameter is approximately constant, and the universe expands exponentially. The metric is that of de-Sitter space

$$ds^2 = -dt^2 + e^{2Ht}d\vec{x}^2 = \frac{1}{(H\eta)^2}(-d\eta^2 + d\vec{x}^2), \tag{1.15}$$

where $\eta \in (-\infty, 0)$. We can read-off from the equations of motion

$$H \approx \sqrt{\frac{V(\varphi)}{3M_{\rm pl}^2}}, \tag{1.16}$$

$$\dot{\varphi} \approx \frac{V'(\varphi)}{3H}. \tag{1.17}$$

The second equation tells us that the velocity of the inflaton is also approximately constant

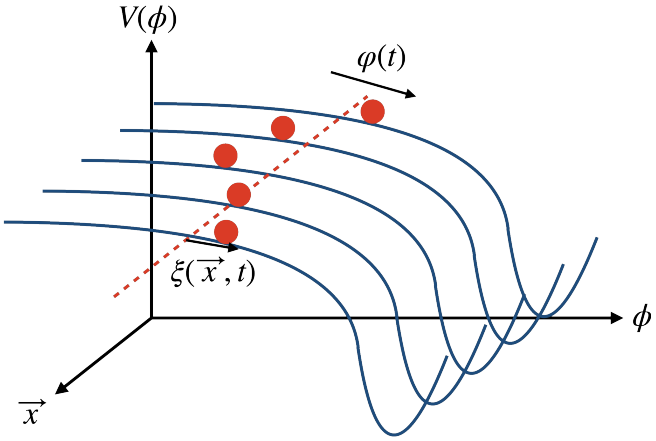


Figure 1.3: Due to quantum fluctuations of the inflaton field inflation ends slightly early or later in different directions in the sky. These deivations account for the anisotropies in the CMB, $\Delta T/T \sim 10^{-5}$.

during inflation.

We have introduced inflation to solve the horizon problem, but inflation automatically includes an origin for perturbations on top of a completely homogeneous and isotropic universe. The intuitive way to think about this, illustrated in figure 1.3, is to split the inflaton $\phi = \varphi(t) + \xi(\vec{x}, t)$, and say that quantum fluctuations of the inflaton cause the inflaton to be slightly ahead or behind of the homogeneous isotropic background. In different directions in the sky inflation either ended slightly earlier, or slightly later, depending on the sign of ξ in that direction.

I will briefly explain how to take the tiny fluctuations and turn them into observables, such as the CMB anisotropies, in the next subsection. For the rest of the section we will work in the idealised limit $\epsilon_V, \eta_V \to 0$. In this limit, ξ is a massless free field in de-Sitter space. There also exist many fields with a non-zero mass, so to study the quantum mechanics of the particles around during inflation we start by quantising a free massive scalar field in de-Sitter space with action

$$S_\chi = \frac{1}{2}\int d^3x dt \frac{1}{(H\eta)^2}\left[(\partial_\eta \chi)^2 - (\partial_i \chi)^2 - \frac{m^2}{(H\eta)^2}\chi^2\right]. \tag{1.18}$$

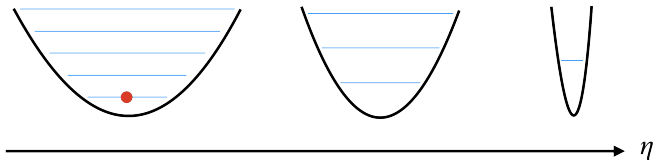


Figure 1.4: At each instant in the time, each k mode is a simple harmonic oscillator with a different frequency. The Bunch-Davies initial condition is to place the universe in the ground state of each oscillator at early times, $\eta \to -\infty$. As time evolves the state corresponding to the ground state in the infinite past has some overlap with excited states at some future time, and therefore the expansion of the universe has led to particle production.

A more thorough treatment of the particle production in curved spacetime can be found in [13]. Due to spatial translation invariance, it is useful to Fourier transform, and we will also find it useful to rescale by conformal time

$$\chi(\vec{x},\eta) = \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{x}} (-\hat{\chi}_{\vec{k}}\eta). \tag{1.19}$$

In terms of these variables we find that the action for a free scalar in de-Sitter space is equivalent to an infinite number of harmonic oscillators with a time-dependent frequency

$$S_\chi = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3} d\eta \left\{ (\partial_\eta \hat{\chi}_{\vec{k}})^2 - \omega_k^2(\eta)\hat{\chi}_{\vec{k}}^2 \right\}; \qquad \omega_k^2(\eta) = k^2 + \frac{m^2 - 2H^2}{H^2\eta^2}. \tag{1.20}$$

For each k mode, the situation is illustrated in figure 1.4. At each instant of time there is a simple harmonic oscillator with a different frequency. We quantise the field by writing the Heisenberg operator

$$\hat{\chi}_{\vec{k}} = \bar{f}_k(\eta) a_{-\vec{k}} + f_k(\eta) a^\dagger_{\vec{k}}, \tag{1.21}$$

where $f_k(\eta)$ satisfies the classical equation of motion, and the creation/annihilation operators satisfy the usual commutation relations $[a_{\vec{k}}, a^\dagger_{\vec{k}'}] = (2\pi)^3\delta(\vec{k} - \vec{k}')$. However, since there are two

independent solutions, which choice do we make for $f_k(\eta)$? The choice, called the Bunch-Davies vacuum [14], is to place the universe in the ground state in the infinite past

$$|\psi\rangle = |0\rangle; \qquad a_{\vec{k}}|0\rangle = 0, \tag{1.22}$$

where the annihilation operator comes with the negative energy solution. The Bunch-Davies condition corresponds to the choice of "mode function"

$$\bar{f}_k(\eta) = \frac{\sqrt{\pi}}{2} e^{-\pi\mu/2} e^{i\pi/4} \sqrt{-\eta} H^{(1)}_{i\mu}(-k\eta) \to \frac{1}{\sqrt{2k}} e^{-ik\eta}, \quad \eta \to -\infty, \tag{1.23}$$

where $\mu = \sqrt{m^2/H^2 - 9/4}$.

The mode function above is true at all times, but the positive and negative frequency modes mix into each other over time. What happens to this mode function at late times, $k\eta \to 0$? If we assume that the particle is quite massive, $\mu \gg 1$

$$\bar{f}_k(\eta) \to \sqrt{\frac{-\eta}{2\mu}} \left\{ \left(\frac{-k\eta}{2}\right)^{i\mu} + e^{-\pi\mu} \left(\frac{-k\eta}{2}\right)^{-i\mu} \right\}. \tag{1.24}$$

To interpret this solution we use the WKB approximation for the mode functions to write the mode function in terms of the creation and annihilation operators for the instantaneous SHO at late times

$$\hat{\chi}_{\vec{k}} \approx \frac{b_{-\vec{k}}}{\sqrt{2\omega_k(\eta)}} e^{-i\int \omega_k(\eta')d\eta'} + \frac{b^\dagger_{\vec{k}}}{\sqrt{2\omega_k(\eta)}} e^{i\int \omega_k(\eta')d\eta'} \tag{1.25}$$

$$= \sqrt{\frac{-\eta}{2\mu}} \left\{ b_{-\vec{k}}(-\eta)^{i\mu} + b^\dagger_{\vec{k}}(-\eta)^{-i\mu} \right\}. \tag{1.26}$$

By matching the two late time limits of the mode function to the definition of the creation and annihilation operators at late times we find that there is a “Bogoliubov transformation" relating the two types of creation and annihilation operators

$$a_{\vec{k}} = \alpha\, b_{\vec{k}} + \beta\, b^{\dagger}_{-\vec{k}}; \qquad |\alpha| \approx 1, \quad |\beta| \approx e^{-\pi\mu}. \tag{1.27}$$

We interpret this as telling us that the expansion of spacetime has led to particle creation. I can ask the question, what is the expectation value for the number of particles with momentum $\vec{k}$ at late times?

$$\langle n_k \rangle = \langle \psi | b^{\dagger}_{\vec{k}} b_{\vec{k}} | \psi \rangle = e^{-2\pi\mu}. \tag{1.28}$$

This is often interpreted as analogous to a Boltzmann suppression, with the Hawking temperature of de-Sitter identified with $T = H/2\pi$ [15], where the Taylor expansion of the Bose-Einstein distribution function $(e^{2\pi\mu} - 1)^{-1} \approx e^{-2\pi\mu}$, for $\mu \gg 1$. However, we should keep in mind that the state $|\psi\rangle$ is a pure state, and not a thermal ensemble, so this is really an analogy, and not a rigorous statement about the state of the universe.

We will return to the topic of particle production later, but now I want to take the massless limit of everything we have just done, to see what we can learn about the statistics of inflaton fluctuations during inflation. In the massless limit we can write the inflaton field

$$\xi_{\vec{k}}(\eta) = c_{-\vec{k}}\, \bar{u}_k(\eta) + c^{\dagger}_{\vec{k}} u_k(\eta); \qquad u_k(\eta) = \frac{H}{\sqrt{2k^3}}(1 - ik\eta)e^{ik\eta}. \tag{1.29}$$

Notice that at late times the mode functions asymptote to a constant, $u_k(\eta) \to H/\sqrt{2k^3}$. It is

then a simple matter to use Wick's theorem to calculate the statistics of inflaton fluctuations (in the limit $\epsilon_V, \eta_V \to 0$)

$$\langle \xi_{\vec{k}} \xi_{\vec{k}'} \rangle = \frac{H^2}{2k^3}(2\pi)^3 \delta(\vec{k} + \vec{k}'), \tag{1.30}$$

$$\langle \xi_{\vec{k}_1} \cdots \xi_{\vec{k}_n} \rangle = \begin{cases} 0, & n \text{ is odd} \\ \langle \xi_{\vec{k}_1} \xi_{\vec{k}_2} \rangle \cdots \langle \xi_{\vec{k}_{n-1}} \xi_{\vec{k}_n} \rangle + \text{ permutations}, & n \text{ is even} \end{cases} \tag{1.31}$$

The fact that odd-point correlation functions are zero, and even correlation functions can be factorised in terms of the two point function is the statement that the statistics of the inflaton are Gaussian. This is a consequence of the inflaton acting as a free field, and we will see later that Non-Gaussianities (NGs), i.e the failure of the statistics to be Gaussian, are a direct probe of the interactions of the inflaton during inflation. We will discuss how to include interactions later, but first we should discuss how to relate what we have just done to cosmological observables.

1.3 The CMB Power Spectrum

The best test of the ideas of inflation comes from the CMB power spectrum. In the CMB power spectrum we are measuring the deviations of the temperature in the CMB from the average value in different directions, which hardly seems related to the two point correlator $\langle \xi_{\vec{k}} \xi_{-\vec{k}} \rangle$ in a universe that exponentially expands forever. However, the calculation of the two-point function in the previous subsection can be directly related to observations of the anisotropies with the following prescription:

1. The fluctuations in the inflaton lead to inhomogeneities and anisotropies in our universe.

In principle we should include perturbations to the metric, as well as inflaton fluctuations, and see how inflaton fluctuations ultimately relate to the observable deviations in the metric. However, in GR things are complicated by the fact that a change of coordinates can mix these two quantities into one another! There is a coordinate system independent way to characterise this deviation called the gauge invariant scalar curvature perturbation, ζ. During inflation there is a coordinate system in which

$$\zeta_{\vec{k}} = -H\frac{\xi_{\vec{k}}}{\dot{\varphi}}. \tag{1.32}$$

2. The next thing we have to deal with is the fact that the universe was never exactly in de-Sitter space. Hubble was not constant during inflation, but slowly varying. It turns out we can easily restore the effect of $\epsilon_V, \eta_V \neq 0$ by noticing that the mode function asymptotes to a constant soon after $-k\eta = 1$. We define the crossing time $\eta_X = -1/k$, and we simply replace all variables with their values evaluated at the crossing time

$$\langle \zeta_{\vec{k}} \zeta_{-\vec{k}} \rangle' \equiv P_\zeta(k) = \frac{H_X^4}{\dot{\varphi}_X^2}\frac{1}{2k^3}. \tag{1.33}$$

We can recast the crossing time into the more general form $k = a_X H_X$. From now on a prime after angle brackets indicates factoring out $(2\pi)^3\delta(\vec{k}_{\rm tot})$.

3. The next piece of information we need is that it is no coincidence that the mode function is no longer changing when $k < aH$. It turns out it is a general property of adiabatic[2]

[2]It turns out there are two types of scalar perturbations to a homogeneous isotropic universe: adiabatic, and isocurvature perturbations. Inflation leads to strictly adiabatic perturbations because all of the perturbations come from the same source. The currently available data supports perturbations in our universe being adiabatic in nature, a result often cited as a hint that inflation is on the right track.

perturbations that the gauge-invariant scalar curvature perturbation satisfies [3]

$$\dot{\zeta}_{\vec{k}} \propto \frac{k^2}{(aH)^2}. \tag{1.34}$$

We say that a mode with $k < aH$ is "outside the horizon" and frozen.

The modes that we observe in the CMB exited the horizon during inflation. This is a crucial point because it means that the unknown physics of reheating has no effect on the perturbations which were generated during inflation. The CMB power spectrum is therefore a probe of the very earliest moments in the beginning of the universe.

4. Finally, the CMB modes re-entered the horizon before recombination, and the initial condition (1.33), or "primordial" contribution to the two-point function, is modified by "transfer functions" in a known way [4]. One can go back and forth between the primoridal perturbations and the observed temperature anisotropies of the CMB, shown in figure 1.5. In addition to the evolution of initial conditions after modes have re-entered the horizon, we also choose to represent the power spectrum by expanding in spherical harmonics, with multipole moment, l. This is a convenient choice because we are observing the temperature anisotropies as a function of angle in the sky.

While I haven't given a derivation of why the above prescription works, I hope that I have made it plausible, and the interested reader can find rigorous justifications of all of these steps in [3,4].

Experimentalists place bounds on the value of

$$\Delta_s^2(k) \equiv \frac{k^3}{2\pi^2} P_\zeta(k), \tag{1.35}$$

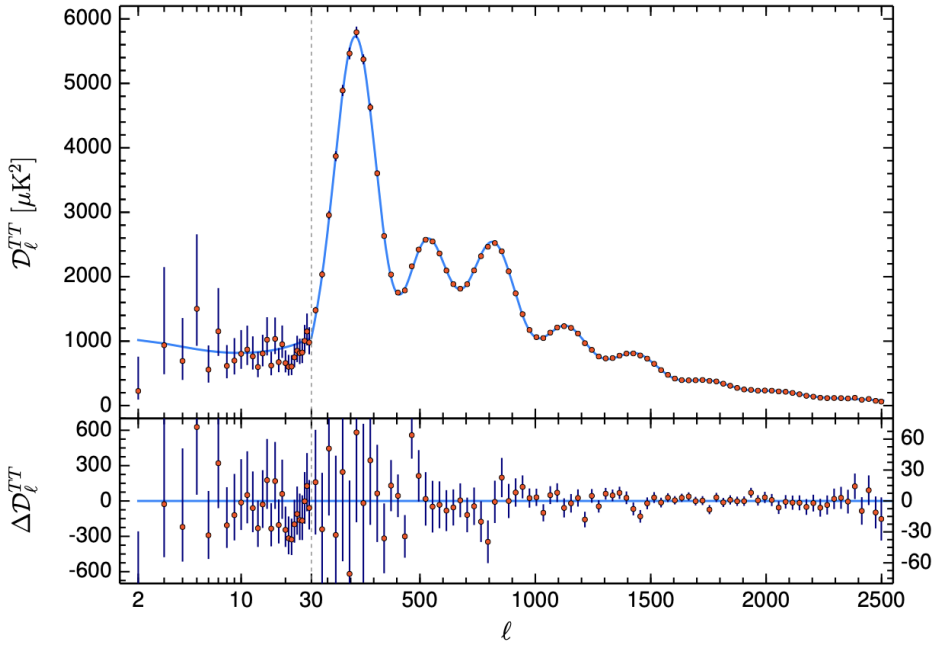


Figure 1.5: The CMB power spectrum, taken from [16].

defined to factor out the $1/k^3$ scaling, which is a consequence of the scale-invariance of exact de-Sitter space. This scaling is therefore broken by the fact that Hubble is not exactly constant during inflation, and the quantity in equation (1.35) does depend on k through the k-dependence of the crossing time. The experimentalists place a bound on the "scalar spectral index", or "spectral tilt"

$$n_s - 1 \equiv \frac{d \log \Delta_s^2}{d \log k} = 2\eta_V - 6\epsilon_V. \tag{1.36}$$

The current best fit value depends on which datasets you combine [17], but they all agree on a value around $n_s \approx 0.965$. This suggests that the slow-roll parameters are $\epsilon_V, \eta_V \sim \mathcal{O}(10^{-2})$.

As well as the spectral tilt, there is a precise measurement for the amplitude of the power

spectrum

$$A_s = \Delta_s^2(k_0) \approx 2.1 \times 10^{-9}, \tag{1.37}$$

for a reference wavenumber $k_0 = 0.05\ \mathrm{Mpc}^{-1}$. Combining this measurement with equation (1.33) gives a value for the ratio of the Hubble scale to the slow-roll velocity

$$\frac{\dot{\varphi}}{H^2} = \frac{1}{\sqrt{4\pi^2 A_s}} \implies \dot{\varphi} \approx (60H)^2, \tag{1.38}$$

a result we will make great use of in the next two chapters.

In summary we have seen that one can calculate the statistics of the inflaton fluctuations by treating it as a free quantum field in de-Sitter space. There is a well-defined prescription to relate that calculation to the observed CMB power spectrum. Finally we saw that experimentalists have very precisely measured the power spectrum, and this has led to precise measurements of many of the most important quantities that characterise inflation.

1.4 Non-Gaussianities

The CMB power spectrum is a probe of the two-point function of the inflaton, and the data we have is consistent with it being Gaussian, as predicted by inflation. The simplest non-Gaussianity one could consider is a non-zero three-point function, or "bispectrum". It is conventional to normalise the three-point function relative to the power spectrum, and define

$$F(k_1, k_2, k_3) = \frac{5}{6} \frac{\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle'}{P_\zeta(k_1)P_\zeta(k_2) + P_\zeta(k_1)P_\zeta(k_3) + P_\zeta(k_2)P_\zeta(k_3)}. \tag{1.39}$$

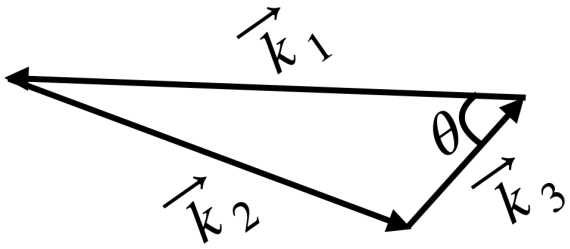


Figure 1.6: Taken from [18]/appendix A. The triangle formed by the momenta involved in the bispectrum. The angle, θ, and the ratio of the lengths k_1/k_3 completely characterises the shape of the triangle.

Due to momentum conservation the three momenta must form a triangle, as illustrated in figure 1.6. The scale-invariance of de-Sitter space implies that, at zeroth order in slow-roll parameters, the bispectrum only depends on the shape of the triangle. An overall rescaling, or a rotation do not change the result. This crucial property allows several independent measurements of the same quantum expectation value by combining measurements of different triangles with the same shape.

There are two conventions we will use for characterising the shape of the triangle. Firstly, the shape is entirely characterised by the ratio of two of the sides of the triangle, $p = 2k_1/k_3$, and the angle between them, $\cos\theta$. An equivalent way to characterise the triangle is to define the two variables

$$p \equiv \frac{k_1 + k_2}{k_3}, \qquad \chi \equiv \frac{k_1 - k_2}{k_3}. \tag{1.40}$$

The two definitions of p are equivalent in the squeezed limit, $k_1 \sim k_2 \gg k_3$, and in this limit $\chi \approx \cos\theta$, which can be easily derived from the cosine rule. It is conventionally assumed, without loss of generality, that the length of k_3 is less than or equal to the length of the other two sides.

Placing experimental bounds on the three-point function usually requires fitting to some

kind of template for how the signal varies as a function of the shape of the triangle. For a non-exhaustive list of recent papers consult [19–23]. The convention is to state bounds as an upper bound on the amplitude of the equilateral triangle configuration

$$f_{\rm NL} \equiv F(k,k,k) < \mathcal{O}(10), \tag{1.41}$$

where the specific value is shape dependent.

We should mention that the proof that inflaton statistics are Gaussian relied on the idea that the inflaton was a free massless field in de-Sitter space. However, if we restore $\epsilon_V, \eta_V \neq 0$ then there are inflaton self-interactions with a strength proportional to the slow-roll parameters. The size of the bispectrum in single-field slow-roll inflation was first computed in [24], and then this calculation was refined in [25]. The prediction is found to be

$$f_{\rm NL} \approx 0.1(n_s - 1). \tag{1.42}$$

This sets a lower bound on the expectation for the bispectrum of $|f_{\rm NL}| \sim \mathcal{O}(10^{-3})$, often referred to as the "gravitational floor".

There are three types of experiments that probe Non-Gaussianities. Firstly, we can measure three-point correlators in the CMB [21, 23], from which the current bounds are taken (see table 1). Secondly, large scale structure surveys have the potential to probe $f_{\rm NL} \sim \mathcal{O}(1)$ due to the extra availability of modes in a 3D map relative to the 2D map of the CMB [26–29]. More futuristically, 21-cm tomography has the potential to probe $f_{\rm NL} \sim \mathcal{O}(10^{-2})$.

If the inflaton interacted with other fields during inflation, then the NGs may be much

larger than this. From now on we work in the limit of exact de-Sitter space, and note that there are corrections to all of our results, but they are all proportional to the slow-roll parameters, and should therefore be small. How do we include interactions in the formalism of the previous subsection? To begin with, in the Schrodinger picture, the "state" of the universe at conformal time η is given by

$$|\psi\rangle = Te^{-i\int_{-\infty}^{\eta} d\eta' H(\eta')}|\Omega\rangle, \tag{1.43}$$

where we recall that because every k-mode redshifts in time, as $\eta \to -\infty$ each mode becomes so short wavelength that its dynamics are no longer sensitive to Hubble. We therefore say that $|\Omega\rangle$ is the Minkowski vacuum state in the limit of early times.

We now use the same trick as in flat space QFT to project from the interacting vacuum onto the free vacuum of the previous subsection. We pick some very early time, $\eta_* \ll \eta$, and consider the following state

$$Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta} d\eta' H(\eta')}|0\rangle = Te^{-i\int_{\eta_*}^{\eta} d\eta' H(\eta')} Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta_*} d\eta' H(\eta')}|0\rangle, \tag{1.44}$$

where $|0\rangle$ is the free Bunch-Davies vacuum at early times. Since all modes are very UV at early times, I can always pick an early enough η_*, that the time-dependent interacting de-Sitter Hamiltonian acts like the time-independent interacting Minkowski Hamiltonian. The contour we are integrating over in the η' plane is illustrated in figure 1.7. We write the second factor in equation (1.44) in terms of a complete set of eigenstates of the, now time-independent,

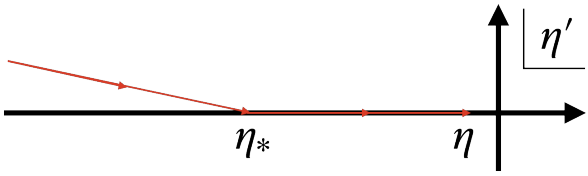


Figure 1.7: A demonstration of the contour we integrate along to implement the $i\epsilon$ prescription.

Hamiltonian at early times

$$\begin{aligned}\lim_{T\to\infty} Te^{-i\int_{-T(1-i\epsilon)}^{\eta_*} d\eta' H(\eta')}|0\rangle &= \lim_{T\to\infty}\sum_n e^{-iE_n(\eta_*+T(1-i\epsilon))}|n\rangle\langle n|0\rangle \\ &= \lim_{T\to\infty} Te^{-i\int_{-T}^{\eta_*} d\eta' H(\eta')}|\Omega\rangle\left(\langle\Omega|0\rangle e^{-\epsilon E_0 T}\right), \end{aligned} \tag{1.45}$$

where in the second step we recognised that only the ground state survives in the limit $T\to\infty$. We then repackaged the phase on the first line in terms of the Hamiltonian acting on the ground state.

Suppose we wanted to calculate the expectation value of some operator, continuing to stick in the Schrodinger picture of the interacting theory for now. We can utilise the result above to rewrite the expectation value

$$\langle\psi|\mathcal{O}_s(\vec{x}_1,\cdots\vec{x}_n)|\psi\rangle = \frac{\langle 0|\bar{T}e^{i\int_{-\infty(1+i\epsilon)}^{\eta} d\eta' H(\eta')}\mathcal{O}_s(\vec{x}_1,\cdots,\vec{x}_n)Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta} d\eta' H(\eta')}|0\rangle}{\lim_{T\to\infty}|\langle\Omega|0\rangle|^2 e^{-2\epsilon E_0 T}}, \tag{1.46}$$

where on the right we have applied equation (1.45) to (1.43), and on the left we have simply taken the Hermitian conjugate. Note that $\bar{T}$ indicates the anti-time ordering operator, because hermitian conjugation inverts the order of all operators.

To simplify this equation, as in flat space QFT, we simply divide by $1=\langle\Omega|\Omega\rangle$. Going through a similar exercise to the one above it is easy to show that we can write the norm of the

interacting vacuum in terms of the same operators as shown above, and we arrive at

$$\langle\psi|\mathcal{O}_s(\vec{x}_1,\cdots\vec{x}_n)|\psi\rangle = \frac{\langle 0|\bar{T}e^{i\int_{-\infty(1+i\epsilon)}^{\eta}d\eta' H(\eta')}\mathcal{O}_s(\vec{x}_1,\cdots,\vec{x}_n)Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta}d\eta' H(\eta')}|0\rangle}{\langle 0|\bar{T}e^{i\int_{-\infty(1+i\epsilon)}^{\eta}d\eta' H(\eta')}Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta}d\eta' H(\eta')}|0\rangle}. \tag{1.47}$$

Now it looks like we have made progress, but I should remind you that the Hamiltonian in the exponential is the full interacting Hamiltonian, so we don't know how it acts on the free vacuum. In addition to this, the Schrödinger picture operator will generically be built out of Schrödinger picture fields in the interacting theory, and we don't know how they act on the free vacuum either.

The final step is to pass to the interaction picture, but this is complicated relative to the same step in flat space QFT because even the free Hamiltonian has explicit time-dependence. We define the time evolution operators in both the full and free theory as time ordered exponentials

$$i\frac{d}{d\eta}U(\eta;\eta_0) = H(\eta)U(\eta;\eta_0) \implies U(\eta;\eta_0) = Te^{-i\int_{\eta_0}^{\eta}d\eta' H(\eta')}, \tag{1.48}$$

$$i\frac{d}{d\eta}U_0(\eta;\eta_0) = H_0(\eta)U_0(\eta;\eta_0) \implies U_0(\eta;\eta_0) = Te^{-i\int_{\eta_0}^{\eta}d\eta' H_0(\eta')}. \tag{1.49}$$

For ease of notation, if I give the time-evolution without η_0 it should be understood that $\eta_0 = -\infty(1\pm i\epsilon)$ with one sign for U, and the other for $U^\dagger$. Now consider the numerator of equation (1.47). The operators inside the expectation value are of the form $U^\dagger\mathcal{O}_sU$. We pass to the interaction picture by giving the interaction picture operators the time-dependence of the free theory (i.e these are the Heisenberg operators of the free theory, and we know how they act on the free vacuum)

$$\mathcal{O}_I(\vec{x}_1,\cdots,\vec{x}_n,\eta) = U_0^\dagger(\eta)\mathcal{O}_s(\vec{x}_1,\cdots,\vec{x}_n)U_0(\eta). \tag{1.50}$$

There is a third time evolution operator, for the interaction picture, which we implicitly define by $U(\eta;\eta_0) = U_0(\eta;\eta_0)U_I(\eta;\eta_0)$. We can solve for the form of U_I by taking the time derivative of the above equation to show

$$i\frac{d}{d\eta}U_I(\eta;\eta_0) = H_I(\eta)U_I(\eta;\eta_0) \implies U_I(\eta;\eta_0) = Te^{-i\int_{\eta_0}^{\eta} d\eta' H_I(\eta')}, \tag{1.51}$$

with H_I the interaction picture interaction Hamiltonian. Plugging all of these results into equation (1.47) gives the well known "master formula" for in-in correlation functions which we will use many times in the next few chapters. Other derivations of this formula are given in [3,30]

$$\langle\Omega|\mathcal{O}_H(\vec{x}_1,\cdots\vec{x}_n,\eta)|\Omega\rangle = \frac{\langle 0|\bar{T}e^{i\int_{-\infty(1+i\epsilon)}^{\eta} d\eta' H_I(\eta')}\mathcal{O}_I(\vec{x}_1,\cdots,\vec{x}_n)Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta} d\eta' H_I(\eta')}|0\rangle}{\langle 0|\bar{T}e^{i\int_{-\infty(1+i\epsilon)}^{\eta} d\eta' H_I(\eta')}Te^{-i\int_{-\infty(1-i\epsilon)}^{\eta} d\eta' H_I(\eta')}|0\rangle}. \tag{1.52}$$

On the left hand side I moved from the Schrodinger to the Heisenberg pictures so that it is clear that the interpretation is that of the quantum expectation value of some observable, at the time η. We will generally take $\eta \to 0$ as a proxy for the end of inflation. The right-hand side is the expectation of a product of interaction picture fields acting on the free vacuum, for which we can use Wick's theorem to derive Feynman rules as we do in flat space QFT. The role of the denominator is to cancel vacuum bubbles, so we simply use the numerator, and omit those diagrams from our expansion.

1.5 Cosmological Collider Physics

Cosmological collider physics is the idea that a measurement of primordial NGs can be used to discover new particles. The original papers on calculating the bispectrum due to interactions

the inflaton has with other fields [1, 31–33] culminated in calculating what I will refer to in this thesis as the "vanilla" cosmological collider signal. At the end of this section we will explain the pros and cons of the vanilla cosmological collider mechanism, and in the next two chapters we will utilise the "chemical potential" mechanism to overcome many of the problems.

To calculate a cosmological collider signal we need to first write down a simple model for how the inflaton could interact with another field. The simplest Lagrangian we could write is[3]

$$\mathcal{L} = \sqrt{-g}\left(-\frac{1}{2}(\partial_\mu\phi)^2 - V(\phi) - \frac{1}{2}(\partial_\mu\sigma)^2 - \frac{1}{2}M_\sigma^2\sigma^2 - \frac{c}{\Lambda}(\partial_\mu\phi)^2\sigma\right), \tag{1.53}$$

where the UV cutoff $\Lambda > \sqrt{\dot{\varphi}} \approx 60H$, the largest dynamical source of energy during inflation. From now on we will work in units where $H = 1$, restoring with dimensional analysis where it is needed. We expand the inflaton about its background $\phi = \varphi + \xi$, and assume the potential, $V(\phi)$, is chosen to satisfy the conditions for slow-roll inflation. We then neglect the backreaction on the metric from $c \neq 0$, and work in the de-Sitter limit $\epsilon_V, \eta_V \to 0$. The former statement implies some bound on the size of c which we won't care about since we are only interested in the qualitative features of the final answer in this section, and the latter statement is justified by the observations we made in the last two sections.

After expanding about the background fields, we treat ξ as a massless field in de-Sitter

[3]It is important that we choose only derivative interactions for the inflaton, or at least interactions that preserve a shift symmetry for ϕ, so that the slow-roll nature of the potential is not corrected by the interactions. If we had not followed this rule we would be punished with IR divergences in our calculation of the bispectrum.

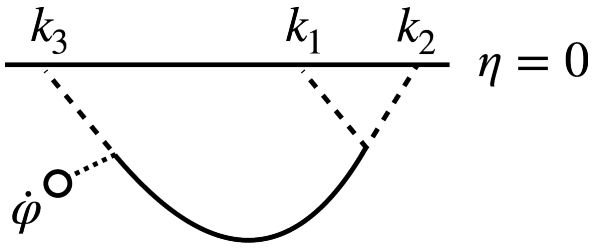


Figure 1.8: The leading Feynman diagram for the bispectrum. This diagram contains four diagrams, with the 4 possible combinations of whether each vertex originated from the time-ordered, or anti time-ordered exponential.

space, and the relevant interaction terms are

$$\mathcal{L}_{\text{Mix}} = \frac{c}{\Lambda}\frac{2}{\eta^3}\dot{\varphi}\sigma\xi', \tag{1.54}$$

$$\mathcal{L}_3 = -\frac{c}{\Lambda}\frac{1}{\eta^2}\sigma(\xi'^2 - (\partial_i\xi)^2), \tag{1.55}$$

where a prime indicates differentiation with respect to conformal time. As in the case of flat-space QFT we use equation (1.52) by Taylor expanding the exponentials, and keeping only the leading terms in the perturbative coupling, c[4].

After Taylor expanding the exponentials we have two types of vertices. Firstly, vertices that originated from the Taylor expansion of the time-ordered exponential are called (+)-type vertices. Vertices that originated from the anti time-ordered exponential are called (−)-type vertices. At tree level there are 4 diagrams that contribution to the bispectrum, illustrated in figure 1.8. Each diagram contains one two-point vertex, and one three-point vertex. We choose to label each diagram as $I_{s_1 s_2}$, where $s_1, s_2 = \pm$ label the two-point, three-point vertices respectively.

[4]Much like in flat space QFT, the derivation we used in the previous section implicitly assumed that the interaction terms do not modify the canonical momentum. However, since we are restricting ourselves to derivative interactions this assumption does not hold. Luckily, as in flat space QFT, we don't need to be careful about this subtlety, and the slickest way to see this is in the path integral formalism [30].

The bispectrum simplifies, and is dominated by a single term

$$\begin{aligned}\langle\xi_{\vec{k}_1}\xi_{\vec{k}_2}\xi_{\vec{k}_3}\rangle' &= I_{++}+I_{+-}+I_{-+}+I_{--}\\ &= I_{++}+I_{+-}+c.c.\\ &\approx I_{++}+c.c.\end{aligned} \tag{1.56}$$

the reason for which we will see in a moment. The dominant term can be expressed in terms of the master formula

$$I_{++}=\langle 0|\xi_{\vec{k}_1}\xi_{\vec{k}_2}\xi_{\vec{k}_3}(-i)^2T\left\{\int_{-\infty_-}^{0}d\eta_1 H_{\rm Mix}(\eta_1)\int_{-\infty_-}^{0}d\eta_2 H_3(\eta_2)\right\}|0\rangle, \tag{1.57}$$

where $\infty_\pm=\infty(1\pm i\epsilon)$. When quantising the massive scalar we rescaled by a factor of $(-\eta)$, so that the theory was equivalent to a harmonic oscillator with a time-dependent frequency. It's easiest now to jump straight to the properly normalised wavefunctions and write

$$\sigma_{\vec{k}}(\eta)=\bar{f}_k(\eta)a_{-\vec{k}}+f_k(\eta)a^\dagger_{\vec{k}};\quad \bar{f}_k(\eta)=\sqrt{\frac{\pi}{4k^3}}e^{-\pi\mu/2}e^{i\pi/4}(-k\eta)^{3/2}H^{(1)}_{i\mu}(-k\eta). \tag{1.58}$$

The quantised form of the inflaton is still given by equation (1.29). In the late-time limit, with the approximation $m\gg H\implies\mu\approx m$, the mode function is

$$\bar{f}_k(\eta)\to\frac{1}{\sqrt{2k^3}}\frac{(-k\eta)^{3/2}}{\sqrt{m}}\left\{\left(\frac{-k\eta}{2}\right)^{im}+e^{-\pi m}\left(\frac{-k\eta}{2}\right)^{-im}\right\}. \tag{1.59}$$

To understand the qualitative features of the integral we will find that it is more useful to directly use this form of the mode function.

Expanding (1.57) with Wick's theorem leads to two types of propagator

- The "bulk to boundary" propagator for the inflaton

$$\langle 0|\xi_{\vec{k}_i}\xi(\eta)|0\rangle = \frac{1}{2k^3}(1-ik\eta)e^{ik\eta}. \tag{1.60}$$

- The "bulk to bulk" propagator for the heavy field

$$\begin{aligned} G_{++}(k,\eta_1,\eta_2) &= \langle 0|T\sigma(\eta_1)\sigma(\eta_2)|0\rangle \\ &= \theta(\eta_1-\eta_2)\bar{f}_k(\eta_1)f_k(\eta_2)+\theta(\eta_2-\eta_1)\bar{f}_k(\eta_2)f_k(\eta_1). \end{aligned} \tag{1.61}$$

In cosmological correlators we are left with integrals over time, even after applying all of the Feynman rules, because there is no energy conservation. However, it turns out that the leading behaviour of the integrals over conformal time can be extracted by performing a stationary phase approximation, with the stationary point having the interpretation as a time during which energy is conserved. Only the second term in (1.61) will allow for such an interpretation, so we will only keep this term in what follows to extract the leading behaviour of the integral.

We apply these results to the quantum expectation value above, and in line with figure 1.8 we assume that $\xi_{\vec{k}_3}$ contracts with the ξ in the two-point vertex.[5] The step function will lead to a nasty nested integral, however the stationary points will be well separated in time, so we can effectively assume there is no time-ordering.

To write out the final result, I will make the substitutions, $x_{1,2} = -k_3\eta_{1,2}$, and work in the

[5]There are two other diagrams where the momentum going through the solid line is k_1, or k_2. These diagrams do not lead to the signal we would search for in cosmological collider physics, so we ignore them in this section. They will play an important role in chapter 3, where they can contribute a significant background.

squeezed limit, $p = 2k_1/k_3 = 2k_2/k_3$, and $\vec{k}_1 \cdot \vec{k}_2 = -k_1 k_2$.

$$I_{++} \approx -\frac{c^2}{\Lambda^2}\dot{\varphi}\frac{1}{4mk_1^3k_2^3}\left(\frac{p}{2}\right)^2 \int_0^\infty dx_1 x^{-1/2} e^{-ix_1}\left\{e^{-\pi m}\left(\frac{x_1}{2}\right)^{im} + \left(\frac{x_1}{2}\right)^{-im}\right\} \times \int_0^\infty dx_2 x_2^{-1/2} e^{-ipx_2}\left[1 + ipx_2 - \frac{p^2x_2^2}{2}\right]\left\{\left(\frac{x_1}{2}\right)^{im} + e^{-\pi m}\left(\frac{x_1}{2}\right)^{-im}\right\}. \quad (1.62)$$

Let us look at each integral in turn. The first integral over x_1 has two terms of the stationary form with phase

$$e^{ig(x_1)} = e^{i(-x_1 \pm m\log x_1)}. \quad (1.63)$$

Only the positive sign, i.e the term with the exponential suppression, has a stationary phase at $x_* = m$, or at conformal time $\eta_* = -m/k_3$. How does this have the interpretation of energy conservation at the x_1 vertex? If you look at the vertex a heavy particle, which we are assuming to be non-relativistic, is changing into a massless inflaton with momentum k_3. The energy of the heavy particle is dominated by its rest mass, and the energy of the inflation redshifts over time, $E = -k_3\eta$. The stationary phase time, η_*, corresponds to the moment energy would be conserved in this interaction[6].

In the second integral, the phase is now

$$e^{ih(x_2)} = e^{i(-px_2 \pm m\log x_2)}. \quad (1.64)$$

The same analysis of the previous case holds, except now the term that can satisfy energy

[6]The attentive reader has probably realised that we have cheated in this derivation. We have cheated by assuming the heavy particle is so non-relativistic that we can drop its momentum entirely. If we included the non-zero momentum in the mode function we would find that there is no stationary phase approximation to the x_1-integral. However, if one calculates the integral exactly, and then takes the squeezed limit, the final result agrees with what we derived in this section [33], so we allow ourselves this one cheat to move through the calculation at a faster pace.

conservation does not have the exponential suppression! The stationary phase time is $\eta_* = -m/(k_1+k_2)$, or $x_* = m/p$, which has the interpretation as an energy conserving decay of the heavy particle in the two inflatons. In the squeezed limit, $p \gg 1$, the times are well separated, justifying our decision to ignore the step function. Notice that $x_2 \ll x_1$, which is equivalent to the three-point vertex occuring later, justifying our decision to drop the other term in (1.61).

We can perform the stationary phase integral, approximating each integral as a Gaussian, with the help of the general formula

$$\int F(x)e^{iG(x)} \approx \sqrt{\frac{2\pi}{|G''(x_*)|}}F(x_*)e^{iG(x_*)}, \tag{1.65}$$

to extract how the bispectrum behaves. We drop an overall phase, but the second integral provides a phase that depends on the shape of the triangle $h(x_*) \supset -m\log p$. The final result is

$$I_{++} \approx -\frac{c^2}{\Lambda^2}\frac{m\pi}{4k_1^3k_2^3}\dot{\varphi}e^{-\pi m/H}p^{3/2-im/H} \implies F \approx -\frac{10m\pi}{3}\frac{c^2}{\Lambda^2}\dot{\varphi}^2e^{-\pi m/H}p^{-3/2-im/H} + c.c., \tag{1.66}$$

where we have restored factors of Hubble in exponents. How do we interpret this answer? Firstly, the exponential suppression, $e^{-\pi m/H}$, is the "Boltzmann suppression" we mentioned when studying the free field theory, except in this case we have the square root of the Boltzmann factor. Secondly there is a factor $p^{-3/2-im/H}$, a non-analyticity due to the propagation of an on-shell heavy particle that cannot be faked by any local interactions.

The interpretation of the calculation above is that this is the $e^{\pm imt}$ phase in the wavefunction of the heavy particle, with the times determined by the stationary phase points. By varying the ratio of the sides of the triangle, we are pinning how long the heavy particle propagated, and

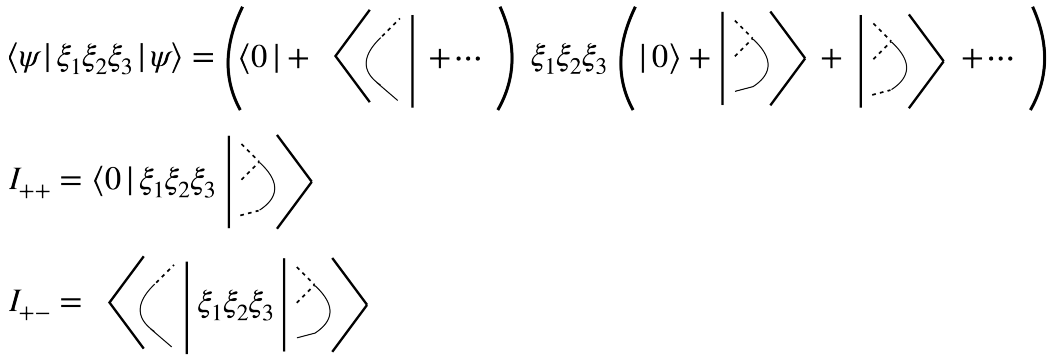


Figure 1.9: We can express the in-in expectation value as the a matrix element in a state that is in a quantum superposition. Each term in the superposition is represented by a Feynman diagram. We see that the two types of diagram we considered can be expressed as a wavefunction overlap.

measuring the amplitude. In addition, the a^{-3} dilution of a massive particle as it propagates is responsible for the factor of $p^{-3/2}$.

The fact that we are sensitive to the phase of the wavefunction suggests this is a quantum interference effect. This also explains why we have the square root of the exponential suppression factor, and allows us to quickly argue that we were correct to ignore the I_{+-} diagram. The argument, summarised pictorially in figure 1.9, goes as follows: The state of the universe is a quantum superposition, which is obtained by Taylor expanding the exponential in equation (1.52), and acting on the free vacuum. The bispectrum is the quantum expectation value of the operator $\xi_1\xi_2\xi_3$ in this state.

The diagram we calculated I_{++} is the overlap of the free vacuum with a state. We can interpret the state as the expansion of the universe pair producing the heavy particle, which propagates for a while, and then decays into inflatons. Recall from our treatment of the free theory that the probability of creating a massive particle obtained a Boltzmann suppression factor, $e^{-2\pi m/H}$, which was the square of the Bogoliubov coefficient (1.27). The amplitude to create a heavy particle is therefore the square root of this, so the overlap of a state which has pair

created the heavy particle once, and the vacuum, will only contain one factor of the Bogoliubov coefficient.

Now consider the other diagram I_{+-}, which consists of the overlap of two different states. In each state the heavy particle is pair produced, and one decays into inflaton(s) while the other survives until the end of inflation. This diagram therefore requires us to pay two factors of the Bogoliubov coefficient, and is exponentially suppressed relative to I_{++}.

The final aspect of equation (1.66) I want to point out is that it is completely independent of $\cos\theta$, the other variable which characterises the momentum triangle. This is because the exchanged particle is a scalar. It turns out that if the exchanged particle has even spin, j, the signal is proportional to $P_j(\cos\theta)$, and if the signal has odd spin things are a bit more complicated, as we will see in the next chapter. It is therefore plausible that we could infer the spin directly from the angular dependence of a cosmological collider signal.

The takeaway from this section is that a particle created during inflation can imprint itself in the NGs of the inflaton. These NGs then evolve into NGs of the CMB or other cosmological observables such as large scale structure. The exchanged particle leads to a distinct oscillatory non-analytic signal that cannot be faked by any local interaction. The signal is a quantum interference effect, and an observation of such a signal could allow us to discover a new particle, and obtain its mass! The current bound on the Hubble scale during inflation is [17]

$$H \lesssim 5 \times 10^{13}\,\text{GeV}. \tag{1.67}$$

Cosmological collider physics is therefore an opportunity to discover particles with masses many orders of magnitude above the reach of any terrestrial collider humans could ever even compre-

hend building.

Based on the calculation we just did there is some reason for pessimism. For masses $m \gg H$, the signal is exponentially suppressed, and probably experimentally undetectable. For masses $m \gtrsim H$, the oscillatory signal becomes very low frequency, and the $p^{-3/2}$ kicks in before a full oscillation can be observed. The mechanism therefore requires a coincidence, a particle must exist with a mass very close to the Hubble scale during inflation. We will see in the next two chapters that the chemical potential mechanism can alleviate this concern, and we identify well motivated candidate theories where new particles are expected with masses around this scale, namely grand unified theories (GUTs).

Chapter 2

Searching for GUTs in Cosmological Data

This chapter is based on the paper, "Grand Unification at the Cosmological Collider with Chemical Potetial" [18], which can be found in its original form in Appendix A. In this chapter I will introduce the idea of utilising the "chemical potential" mechanism to overcome many of the aforementioned problems with cosmological collider physics in its original incarnation. After briefly describing how to include a chemical potential for a massive scalar we will extend the mechanism to producing massive spin-1 particles. We will conclude that it is possible to have observably large sizes for the bispectrum over a wide range of parameters.

We will see that the upper bound on the chemical potential is $\lambda < 60H < 3 \times 10^{15}$ GeV. It is natural to consider whether we could search for the heavy gauge bosons predicted by grand unified theories (GUTs), given that they are expected to have masses around this scale. After a brief review of the evidence for GUTs, and a quick overview of the virtues of their extra-dimensional "orbifold" incarnations, we will discuss how to couple them to inflation.

2.1 The Chemical Potential Mechanism

We all know from statistical mechanics that there is one way to overcome Boltzmann suppression, introduce a chemical potential

$$e^{-E/k_B T} \to e^{-(E-\lambda)/k_B T}. \tag{2.1}$$

It turns out that it is possible to introduce couplings to the inflaton that act like a chemical potential[1] for the heavy particles we would like to search for in cosmological collider physics

[1] I will remind the reader here that the analogy with statistical mechanics is really an analogy. The state of the universe is still a pure state, not a mixed state resembling the grand canonical ensemble.

[18,34–36]. The chemical potential allows for unsuppressed particle production of particles with masses, $m < \lambda$. Not only does this increase the upper limit of energies probed by inflation, but it also widens the window of possible masses that are observable. How big can λ be? We have to identify the largest dynamical energy source during inflation, which as we have already noted is $\sqrt{\dot{\varphi}} \approx 60H$.

In [35] the kinetic energy of the inflaton was utilised to provide a chemical potential for a massive complex scalar with the following Lagrangian:

$$\mathcal{L}_\chi \supset -|\nabla_\mu \chi|^2 - m^2|\chi|^2 + \frac{1}{\Lambda}\nabla_\mu \phi J^\mu + \alpha\chi + h.c. \tag{2.2}$$

where J^μ is the Noether current of χ number conservation that can be derived from a Lagrangian with a $U(1)$ rephasing symmetry for χ. For earlier attempts to utilise chemical potentials in cosmological collider physics see [34, 37, 38]. The first interaction term is a chemical potential for χ number, which can be seen by plugging in the inflaton slow-roll background

$$\frac{1}{\Lambda}\nabla_\mu \phi J^\mu \to \frac{\dot{\varphi}}{\Lambda} J^0 \implies \lambda = \frac{\dot{\varphi}}{\Lambda}. \tag{2.3}$$

The second interaction is a linear $U(1)$ breaking term. If we don't include this term then the coupling between the inflaton and the heavy field is redundant and can be removed by the field redefinition $\chi \mapsto e^{-i\phi/\Lambda}\chi$. However, in the presence of the linear breaking term, this field redefinition does not completely remove the coupling. The Lagrangian after performing this

field redefinition is in the form that is most convenient for calculations

$$\mathcal{L}_\chi = -|\nabla_\mu \chi|^2 - m^2|\chi|^2 + \alpha\chi e^{-i\phi/\Lambda}. \tag{2.4}$$

This form of the Lagrangian appears to break the rule I stated in chapter 1, that we should include only derivative interactions of the inflaton. However, there is still a shift symmetry

$$\phi \mapsto \phi + c; \qquad \chi \mapsto e^{ic/\Lambda}\chi, \tag{2.5}$$

which protects the slow-roll nature of the potential.[2]

The prediction for the bispectrum in this model has some qualitatively new features compared to the vanilla signal [35]

$$F \approx |f_{\text{oscil}}(\mu, \lambda)| p^{-3/2+i(\mu-\lambda)} + c.c; \qquad \mu < \lambda. \tag{2.6}$$

Firstly, there is no exponential suppression. The detectable mass window has therefore been widened from a small region around Hubble, to potentially more than an order of magnitude in Hubble units.

The second interesting feature is that the frequency is now $\mu - \lambda$, rather than just μ. This can also improve the prospects for detectability as the oscillations may be much higher frequency, and it is more likely that a full oscillation will occur before the $p^{-3/2}$ suppression kills the signal. Finally, I note that the authors found that the frequency of oscillation is modified in

[2]However, now that the interactions are not manifestly shift-symmetric, the integrals do appear to contain IR divergences. The existence of the shift symmetry guarantees the IR divergences will cancel, but only in the sum over all Feynman diagrams.

the unsqueezed limit, $p \approx 2$, to $\lambda p/(1+p)$. It is therefore possible to infer the values of both μ, and λ from the signal. We will find the same result holds for spin-1, so we defer the derivation of this feature for later.

2.2 Chemical Potential for Spin-1

In this section we extend the chemical potential mechanism of the previous section to spin-1 massive particles. The lesson we takeaway from the scalar chemical potential is that we should work in the rotated basis with the following rules:

1. We require a complex vector, A_μ, so that it has a conserved charge which can acquire a chemical potential.

2. We must break the $U(1)$ symmetry of the complex vector, whilst preserving a shift symmetry for the inflaton

$$A_\mu \mapsto e^{i\alpha/\Lambda} A_\mu; \qquad \phi \mapsto \phi + \alpha. \tag{2.7}$$

   The way to enforce this symmetry is to include a factor of $e^{-i\phi/\Lambda}$ for every factor of A_μ that does not include A^*_μ.

The lowest dimension choice of interaction is given by

$$\mathcal{L} = -\frac{1}{2}|F_{\mu\nu}|^2 - m^2|A_\mu|^2 + c\Lambda\nabla_\mu\phi A^\mu e^{-i\phi/\Lambda} + h.c. \tag{2.8}$$

However, this term is redundant. To see this note that $\nabla_\mu \phi e^{-i\phi/\Lambda} = i\Lambda \nabla_\mu (e^{-i\phi/\Lambda})$, so after integrating by parts we obtain $\nabla_\mu A^\mu = 0$. The next lowest dimension operator is given by

$$\mathcal{L}_{\rm int} = \frac{c}{\Lambda^3} \nabla_\mu \phi A^\mu (\nabla_\nu \phi)^2 e^{-i\phi/\Lambda} + h.c. \tag{2.9}$$

Even though this term is dimension-7, we will see that it still has the potential to produce observably large NGs.

Interactions with massive spin-1 particles have an extra complication relative to the scalar case. The interactions of longitudinal modes becomes strong in the UV, leading to the breakdown of the EFT. To isolate this effect one can represent the longitudinal mode as the derivative of a scalar [39],

$$A_\mu \mapsto \frac{1}{m} \nabla_\mu \pi, \tag{2.10}$$

showing that interactions of the longitudinal mode are enhanced by a factor of E/m, for processes with characteristic scale, E. Requiring the EFT of the vector coupled to the inflaton does not get strongly coupled before the UV cutoff, Λ, constrains the size of the interaction

$$c < \frac{m}{\Lambda}. \tag{2.11}$$

If this constraint is not satisfied[3], then higher order terms in c will not be suppressed relative to the leading order calculation of the bispectrum, and the EFT will have completely broken down.

To quantise a spin-1 field in de-Sitter space we focus on the longitudinal mode, as it is the

[3]This constraint may carry some extra phase factors. However, they will always work in our favour, and will not make much difference to our final results.

only mode that appears in the bispectrum calculation. For a more detailed account of quantising vectors in de-Sitter space consult the appendix of [33]. We expand the longitudinal mode

$$A_\mu^L = \bar{\sigma}_\mu(k,\eta) a_{-\vec{k}} + \sigma_\mu(k,\eta) b_{\vec{k}}^\dagger, \tag{2.12}$$

where we have included two types of creation/annihilation operators for the complex field. The mode function can be expanded

$$\bar{\sigma}_\mu(k,\eta) = \left(\bar{\sigma}_0(k,\eta), \bar{\sigma}_1(k,\eta)\hat{k}\right). \tag{2.13}$$

The longitudinal mode is the only mode with a non-zero A_0, and the longitudinal mode satisfies $\vec{k}\cdot\vec{A}_{\vec{k}} = k|\vec{A}_{\vec{k}}|$, where as the transverse modes satisfy $\vec{k}\cdot\vec{A}_{\vec{k}} = 0$. The mode functions can be explicitly calculated, and are expressible in terms of the mode function of the scalar field (1.58)

$$\bar{\sigma}_0(k,\eta) = \frac{k}{m}\bar{f}_k(\eta); \qquad \bar{\sigma}_1(k,\eta) = -\frac{i}{k}\left(\bar{\sigma}_0' - \frac{2}{\eta}\bar{\sigma}_0\right). \tag{2.14}$$

The only other difference is that for vector fields the mass parameter is modified relative to the scalar case, $\mu = \sqrt{m^2/H^2 - 1/4}$.

2.3 Calculating the Bispectrum

To calculate the bispectrum we expand around the inflaton background and obtain the same two types of vertices as in the vanilla case.

$$\mathcal{L}_{\rm Mix} + \bar{\mathcal{L}}_{\rm Mix} = \frac{c}{\Lambda}(-\eta)^{i\lambda}\left\{i\eta\lambda^3 A_0\xi + 3\eta^2\lambda^2 A_0\delta\xi' - \lambda^2\eta^2 A_i\partial_i\xi\right\} + h.c.\,, \tag{2.15}$$

$$\begin{aligned}\mathcal{L}_3 + \bar{\mathcal{L}}_3 = \frac{c}{\Lambda^2}(-\eta)^{i\lambda}\Bigg\{&\frac{\lambda^3\eta}{2}A_0\xi^2 - 3i\lambda^2\eta^2 A_0\xi\xi' - 3\lambda\eta^3 A_0(\xi')^2 \\ &+\lambda\eta^3 A_0(\partial_i\xi)^2 + i\lambda^2\eta^2 A_i\xi\ \partial_i\xi + 2\lambda\eta^3 A_i\xi'\ \partial_i\xi\Bigg\} + h.c.\end{aligned} \tag{2.16}$$

There are two important features of these interactions that we should bare in mind. Firstly, the factor

$$(-\eta)^{i\lambda} = e^{i\lambda\log(-\eta)} = e^{-i\lambda t}, \tag{2.17}$$

should be thought of as a large energy source, injecting (or removing) λ at each vertex. Mechanically it will provide a stationary point in the integrals like equation (1.62) without any Boltzmann suppression. The chief consideration in determining the leading diagram is now whether energy is injected at each vertex. We will see that now I_{-+} is the only diagram in which energy is injected at both vertices. The second important feature is that equation (2.15) only involves either A_0, or $A_i\partial_i\xi$. This form of interaction implies that only the longitudinal mode of the vector is exchanged, because the spatial component is proportional to $\vec{k}\cdot\vec{A}_{\vec{k}}$, which is only non-zero for the longitudinal mode.

The Feynman diagrams are shown in figure 2.1. The tree level bispectrum is, as before, the sum of 4 diagrams that split into

$$\langle\xi_{\vec{k}_1}\xi_{\vec{k}_2}\xi_{\vec{k}_3}\rangle' = I_{-+} + I_{++} + c.c., \tag{2.18}$$

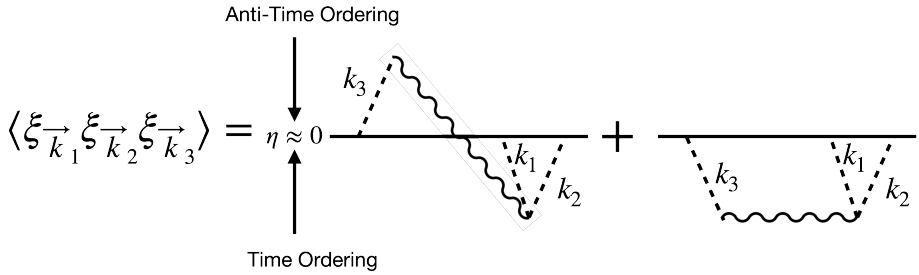


Figure 2.1: Feynman diagrams for the tree level contributions to the bispectrum, due to the exchange of a massive vector. We introduce a new notation in which vertices above the $\eta \to 0$ line are taken from the anti time-ordered exponential, and vertices below the line are taken from the time-ordered side. The two diagrams shown are therefore I_{-+}, and I_{++}.

where the diagram I_{-+} is now the dominant diagram because it is the only diagram that has energy injected in both vertices.

The dominant diagram, I_{-+}, can be found by expanding (1.52), and taking one interaction from each exponential. There is also the question of whether we take $(-\eta)^{\pm i\lambda}$ for each vertex. The dominant diagram is

$$I_{-+} = \langle 0| \int d^4x_- \sqrt{-g}(-i)\bar{\mathcal{L}}_{\rm Mix}\xi_{\vec{k}_3}\xi_{\vec{k}_1}\xi_{\vec{k}_2} \int d^4x_+ \sqrt{-g}(+i)\mathcal{L}_3|0\rangle, \tag{2.19}$$

which corresponds to an injection of energy at each vertex. The term $\mathcal{L}_{\rm Mix}\bar{\mathcal{L}}_3$ will lead to energy removed at each vertex. This contribution will be suppressed by a very large exponential factor.

In this diagram there is no time-ordering, and the integrals break up nicely into a product of two integrals

$$I_{-+} = \frac{|c|^2}{\Lambda^3} I^-_{k_3} I^+_{k_{12}} + c.c. \tag{2.20}$$

where the individual integrals can be expressed in terms of the mode functions for the complex

vector

$$I^-_{k_3} = \frac{1}{\sqrt{2k_3^3}} \int_{-\infty}^{0} \frac{d\eta}{\eta^4}(-\eta)^{-i\lambda} \left\{(-i\eta\lambda^3 \bar{u}_{k_3} + 3\eta^2\lambda^2 \bar{u}'_{k_3})\bar{\sigma}_0(k_3,\eta) - i\lambda^2\eta^2 k_3 \bar{u}_{k_3}\bar{\sigma}_1(k_3,\eta)\right\}, \quad (2.21)$$

$$I^+_{k_{12}} = \frac{1}{\sqrt{4k_1^3k_2^3}} \int_{-\infty}^{0} \frac{d\eta}{\eta^4}(-\eta)^{i\lambda} \left\{ \left(\left(\frac{\lambda^3\eta}{2} + \lambda\eta^3 k_1 k_2 \right) u_{k_1} u_{k_2} - 3i\lambda^2\eta^2 u_{k_1} u'_{k_2} - 3\lambda\eta^3 u'_{k_1} u'_{k_2} \right) \sigma_0 \right.$$
$$\left. + \vec{k}_2 \cdot \hat{k}_3 \left(-\lambda^2\eta^2 u_{k_1} u_{k_2} + 2i\lambda\eta^3 u'_{k_1} u_{k_2}\right) \sigma_1(k_3,\eta) + (k_1 \leftrightarrow k_2) \right\}. \quad (2.22)$$

In the second integral we note that the external inflaton fields $\xi_{\vec{k}_1}\xi_{\vec{k}_2}$ have two choices of contractions with the inflaton fields in the three-point vertex. In the scalar case, the two contractions gave identical integrals, so I multiplied by a factor of 2, but in this case the contributions are not identical, so I have included $k_1 \leftrightarrow k_2$ to account for the other choice.

In the squeezed limit, the top line of equation (2.22) is symmetric in the exchange of k_1, and k_2, but the bottom line is almost antisymmetric. If we expand out the dot product $\vec{k}_1 \cdot \vec{k}_3 = -k_1 k_3 \cos\theta$ whereas $\vec{k}_2 \cdot \vec{k}_3 = k_2 k_3 \cos\theta$. These terms almost cancel, except in the squeezed limit $k_1 \neq k_2$, and we can use the cosine rule to find

$$k_2 = k_1 \left(1 - \frac{k_3}{k_1}\cos\theta + O\left(\frac{k_3}{k_1}\right)^2\right) \implies k_1 - k_2 = k_3 \cos\theta. \quad (2.23)$$

This causes the term multiplying σ_1 to be proportional to $\cos^2\theta$. If this cancellation did not exist, then the term muliplying σ_1 would dominate over the σ_0 term, and the bispectrum would be proportional to $\cos\theta = P_1(\cos\theta)$. However, the cancellation leads to a more complicated angular dependence.

To solve the integral we will need the WKB approximation for the vector mode functions, but we require a more detailed form. However, unlike in (1.62), we will not need the exponentially

suppressed piece. The more detailed WKB approximation for the mode functions are given by equations (A.80), and (A.81). We make the same substitution, $x = -k_3\eta$, and find

$$I^-_{k_3} = \frac{\lambda^2}{2\sqrt{2}mk_3^{3/2-i\lambda}} \int_0^\infty dx \frac{C(x)}{x^{3/2}(x^2+\mu^2)^{1/4}} e^{ig(x)}, \tag{2.24}$$

where the phase is given by

$$g(x) = -\lambda \ln x + x + \int^x \sqrt{1+\mu^2/y^2}dy, \tag{2.25}$$

and $C(x)$ can be found in equation (A.91). We can interpret the phase in the integrand as the statement that this vertex is dominated by the time at which energy is conserved, just as in the scalar case.

To see this interpretation explicitly, note that in terms of conformal time $dt = d\eta/\eta$, so we can write each term in the phase as

$$g(x) = \int^x \frac{dy}{y}\left(y + \sqrt{y^2+\mu^2} - \lambda\right) = \int^\eta \frac{d\eta'}{\eta'}\left((-k_3\eta') + \sqrt{(-k_3\eta')^2+\mu^2} - \lambda\right). \tag{2.26}$$

Each term in the final expression can be interpreted as an energy. The first term is the energy of the massless inflaton, with momentum k_3. The second is the energy of a massive particle, if we think of $\mu \approx m$. The final term is to be interpreted as an energy injection by the chemical potential, λ. The stationary phase approximation is saying that λ, the energy injected at the vertex, is the amount required to create an inflaton and a massive particle, each with momentum k_3.

A similar interpretation can be given for the second integral, which after the substitution

$x = -k_3\eta$, is given by

$$I^{+}_{k_{12}} = \frac{\lambda}{2\sqrt{2}mk_1^3k_2^3k_3^{-3/2+i\lambda}}\int_0^\infty dx \frac{A(x)+B(x)\cos^2\theta}{x^{3/2}(x^2+\mu^2)^{1/4}}e^{ih(x)}, \tag{2.27}$$

where the phase is given by

$$h(x) = \lambda \ln x - px - \int^x \sqrt{1+\mu^2/x'^2}dx', \tag{2.28}$$

and the other functions can be found in the appendix (A.98), (A.99).

Both of these integrals can be solved by a stationary phase approximation, and the details can be found in appendix A. For our purposes we work out the stationary phase time in the second integral

$$\begin{aligned} h'(x) &= \frac{\lambda}{x} - p - \sqrt{1-\mu^2/x^2} = 0 \\ &\implies x_* = \frac{p\lambda}{p^2-1} - \frac{\sqrt{\lambda^2-\mu^2+\mu^2p^2}}{p^2-1}. \end{aligned} \tag{2.29}$$

We will be interested in two limits of this result. Firstly, in the squeezed limit, $p \gg 1$, we find $x_* \approx \frac{\lambda-\mu}{p}$. The stationary phase only exists for $\lambda > \mu$, which is what we expected. In this limit the phase

$$e^{ih(x_*)} \approx p^{i(\mu-\lambda)}. \tag{2.30}$$

However, if $\lambda > 2\mu$, there exists a regime in which $p < \lambda/\mu$, and there is a limit

$$e^{ih(x_*)} \approx (1+p)^{i\lambda}. \tag{2.31}$$

The existence of the two regimes implies that a measurement of how the non-analytic behaviour varies as a function of squeezing can be used to independently measure both the mass and chemical potential [35]

$$F \propto \begin{cases} (1+p)^{i\lambda}, & p < \lambda/\mu, \\ p^{i(\mu-\lambda)}, & p > \lambda/\mu. \end{cases} \tag{2.32}$$

Now let us focus on the squeezed limit. The rest of the calculation is performed in detail in appendix A, and we quote the result here

$$F(p, \cos\theta) = \frac{1}{2} f_{\rm oscil}(\mu, \lambda, \cos\theta) p^{-5/2+i(\mu-\lambda)} + c.c. \tag{2.33}$$

$$f_{\rm oscil}(\mu, \lambda, \cos\theta) = \frac{40\pi|c^2|}{3\sqrt{2}m^2}\dot{\varphi}\frac{\lambda^3}{\Lambda^3}\sqrt{\frac{\lambda}{\mu}}\frac{C(\mu,\lambda)\left[A(\mu,\lambda)+B(\mu,\lambda)\cos^2\theta\right]}{(\lambda-\mu)(\lambda^2-\mu^2)}. \tag{2.34}$$

The key features to pay attention to in this result are the extra dilution factor, $F \propto p^{-5/2}$, which is a result of the cancellation between diagrams we mentioned above, the new phase $\mu-\lambda$, which we also anticipated, and comes from the new stationary phase time due to the energy injection from the chemical potential, and finally the angular structure $F \propto A + B\cos^2\theta$.

The angular dependence can be expressed in terms of Legendre polynomials

$$F \propto A(\mu,\lambda)P_0(\cos\theta) + B(\mu,\lambda)\cos\theta P_1(\cos\theta). \tag{2.35}$$

This is consistent with the general expectation for odd spins. The general result in the “vanilla" case can be found in (C.26) of [33]. The chemical potential result differs only in the order one factors multiplying each Legendre polynomial.

2.4 Results & Conclusions

As we mentioned in chapter 1, it is conventional to express constraints in terms of the equilateral triangle configuration, $f_{\rm NL}$. To figure out the largest possible theoretical value of $f_{\rm NL}$, we need the constraints on the coupling, $|c|$. There are two constraints on $|c|$:

- The EFT for an interacting spin-1 field breaks down unless

$$c < m/\Lambda \tag{2.36}$$

- There is another constraint on the size of the cuopling from ensuring that contributions to the power spectrum are small. For details consult appendix A

$$\frac{|c|^2\Lambda^2}{m^2} < 0.1\frac{\Lambda^4}{\lambda^4}. \tag{2.37}$$

In figure 2.2 we show the largest possible size of $f_{\rm NL}$, including constraints on the size of the coupling. In this graph we restrict to the region $\mu < \lambda/2$, because as $\mu \to \lambda$, our answer blows up. This is because the stationary phase approximation breaks down in that limit. A full analytic calculation of the integrals is doable, but left for future work.

The key takeaway is that for a wide range of parameters we are able to obtain

$$f_{\rm NL} \sim \mathcal{O}(1 - 10^3), \tag{2.38}$$

within theoretical control. Much of this plot will already be ruled out. As we said the precise constraint will require fitting to a specific template, but $f_{\rm NL} < \mathcal{O}(10 - 50)$ is a reasonable

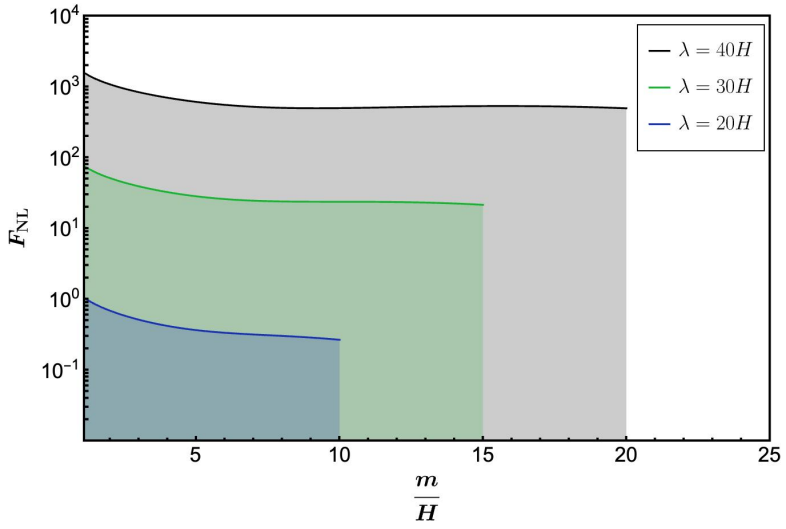


Figure 2.2: Taken (and slightly adapted) from [18]. A plot of the maximum possible signal size $f_{\rm NL}$ within theoretical control for various values of λ. We chose $\cos\theta = 0.5$, and restricted to $\mu < \lambda/2$, as the stationary phase approximation breaks down when $\mu \to \lambda$.

expectation. We therefore see that it is possible to saturate the current bounds for the larger values of the chemical potential, and for small values of λ signal sizes are possible within the reach of future LSS, or 21-cm experiments.

2.5 Orbifold Grand Unified Theories

In the previous few sections we have studied the EFT of a massive spin-1 particle coupled to the inflaton, and we claimed that it is possible to obtain observably large non-analytic NGs. This probably seems to the reader to be rather speculative. Why should I think a spin-1 particle with its mass in the right range even exists, let alone that it couples in the right way to the inflaton? The existence of heavy gauge bosons with masses in the right range is predicted by grand unified theories (GUTs) [40–42].

In GUTs the three gauge groups in the standard model, $G_{\rm SM} = SU(3) \times SU(2) \times U(1)$ emerge at low energies due to the spontaneous breaking of a simple gauge group, $G_{\rm GUT}$. At tree

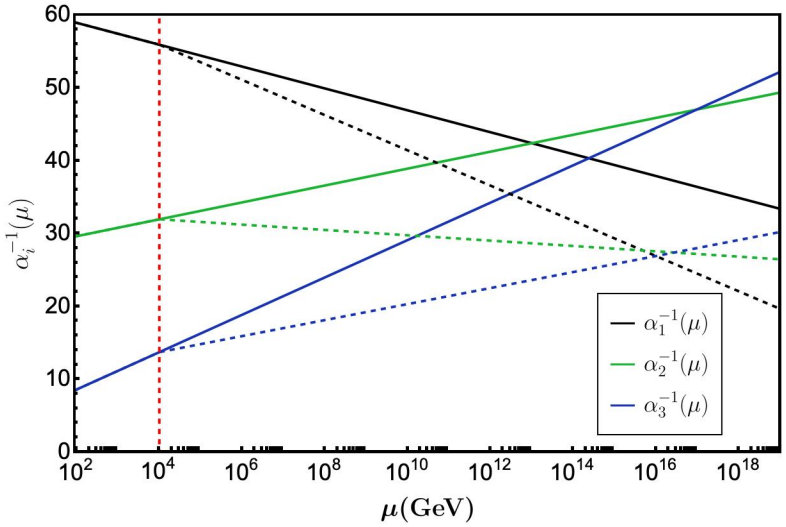


Figure 2.3: The 1-loop running couplings in the standard model (solid lines), and the MSSM (dashed lines). I have ignored the top mass threshold, and included the top quark in the entire running. I have also assumed that all the superpartners have the same mass, $m_s = 10$ TeV, shown by the red dashed line on the plot. A more refined treatment has no impact on the conclusions we will draw from this plot. To obtain expressions for the 1-loop beta function of gauge couplings consult chapter 16 of [43].

level GUTs predict that the SM gauge couplings should unify at the scale of symmetry breaking, usually called the "GUT scale". At loop level the exact unification can be slightly modified. The fact that the standard model gauge couplings approximately unify at high energies, as shown in figure 2.3, is often cited as circumstantial evidence for this idea. One other appealing aspect of grand unification is the fact that the standard model matter fields fit nicely into representations of the GUT gauge groups. GUTs then have the potential to explain the seemingly random quantum numbers of the particles in the standard model.

There are a few problems with the minimal models of grand unification. Firstly, GUTs predict proton decay at a rate that is in tension with current bounds [44, 45]. Another issue with GUTs is the doublet-triplet splitting problem (the nomenclature derives from the minimal $SU(5)$ model). The Higgs field must survive to low energies to cause electroweak symmetry

breaking at $\nu \sim \mathcal{O}(100)$ GeV, but if the Higgs fits into an irreducible representation of the GUT gauge group, then there are extra components. The extra components (the triple in $SU(5)$) can mediate proton decay, and proton decay bounds imply the triplet must have a mass of order the GUT scale. Explaining why the SM Higgs doublet is very light, whilst the triplet is very heavy is called the doublet-triplet splitting problem.

Both of these problems are solved in extra dimensional "orbifold" GUTs [46–48]. In orbifold GUTs the gauge symmetry is broken $G_{\rm GUT} \to G_{\rm SM}$ by the boundary conditions on a brane in an extra dimension, rather than by the Higgs mechanism. The vanilla cosmological collider prospects of orbifold GUTs were investigated in [49]. They found that the signal due to the exchange of the heavy gauge-bosons is highly suppressed, and the leading signal was a small signal due to the exchange of the KK graviton. We are interested in figuring out if one can utilise the chemical potential to boost the signal for the gauge bosons.

There are a few other issues with GUTs in their original incarnation. Firstly, a glance at figure 2.3 tells us that the approximate unification at $M_{\rm GUT} \sim 10^{14}$ GeV is not very exact. In addition to this, minimal GUTs have a large hierarchy problems for the Higgs mass. It requires extreme fine-tuning to have a Higgs mass that is so much smaller than the GUT scale. However, in the minimal supersymmetric standard model (MSSM), the gauge coupling unification, shown by the dashed lines in figure 2.3, is close to being exact. Supersymmetry can also solve the hierarchy problem, as quadratic divergences to the Higgs mass cancel between the standard model particles and their superpartners. The fact that we have not seen the superpartners at colliders implies supersymmetry must be broken at some scale, m_S. We don't know what this scale is, and I have chosen for illustrative purposes $m_s = 10$ TeV, which is safe from collider bounds, but moving m_s has little effect on the apparent unification, or the fact that the

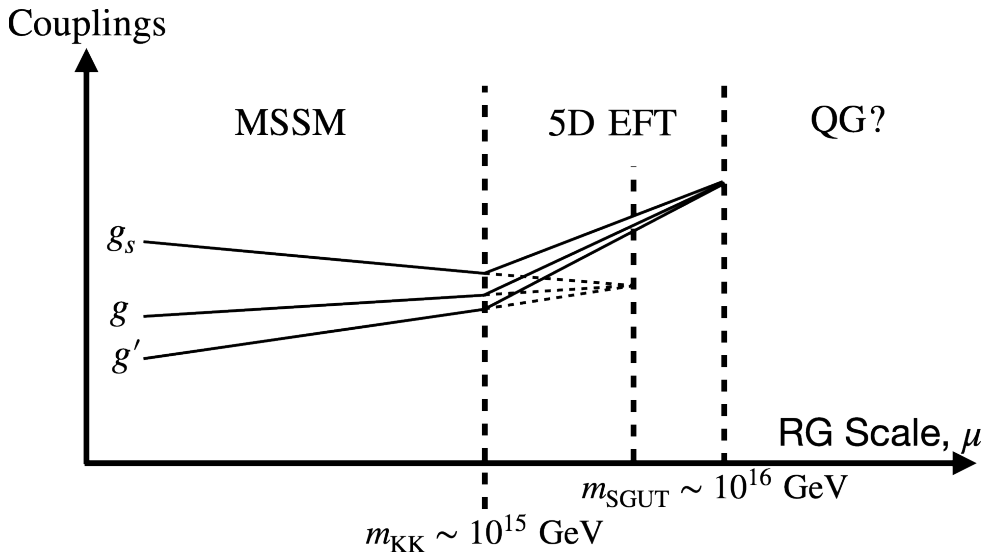


Figure 2.4: Adapted from ref. [48], and taken from [18]/appendix A. In orbifold GUTs the KK modes kick in at the lower scale of $m_{\rm KK} \sim 10^{15}$ GeV. Above this scale the gauge couplings become dimensionful and there is power law running until the couplings unify at some higher scale.

unification is at $M_{\rm SGUT} \sim 10^{16}$ GeV.

From the point of view of cosmological collider physics the SUSY GUT scale is well above the maximum value of the Hubble scale. In addition to this, it is even above the maximum value of the chemical potential. However, if we consider the KK scale (i.e the mass scale of new particles) in orbifold SUSY GUTs it is moderately lower than the scale one would infer from figure 2.3 at $m_{\rm KK} \sim 10^{15}$ GeV [48]. The running of the gauge couplings in this scenario is illustrated in figure 2.4.

We have seen that orbifold (non-)SUSY GUTs are a well motivated scenario for physics beyond the standard model with spin-1 particles expected to exist with masses around the scale to be probed by cosmological collider physics. The final thing to note is that the chemical potential mechanism requires a complex vector, that is uncharged under the standard model! This does not exist in the $SU(5)$, and $SO(10)$ options. However, the classic “trinification"

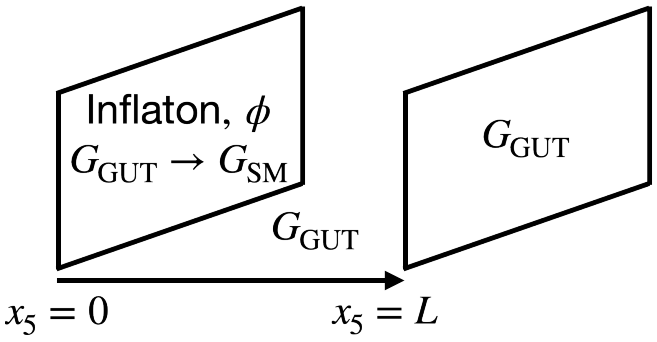


Figure 2.5: Taken from [18]/appendix A. setup for orbifold grand unification. We consider the inflaton to be a 4D field localised to a brane at $x_5 = 0$ in an extra dimension. The extra dimension is an interval of length, L. In the "bulk", and on the brane at $x_5 = L$, the full Trinification gauge symmetry is respected. However, we break the gauge symmetry down to the standard model on the brane at $x_5 = 0$ with a judicious choice of boundary conditions.

scenario [50] with gauge group $G_{\rm GUT} = SU(3)_C \times SU(3)_L \times SU(3)_R/\mathbb{Z}_3$ provides a candidate.

2.6 Coupling to Inflation

The extra dimensional setup for the orbifold GUT is shown in figure 2.5. The symmetry breaking of trinification down to the standard model in this context has been studied in [51,52]. We combine that setup with the inflationary setup in [49], in which the inflaton is considered to be a 4D field localised to a brane in the extra dimension. Outside of the brane at $x_5 = 0$, the full gauge invariance of the GUT group is respected, but on the brane at $x_5 = 0$ we impose the following boundary conditions on the gauge fields

$$\nabla_5 A_\mu^{\rm SM}|_{x_5=0} = 0, \tag{2.39}$$

$$A_\mu^{\not{\rm SM}}|_{x_5=0} = 0. \tag{2.40}$$

This ensures that only standard model gauge invariance is respected on this brane. By imposing Neumann boundary conditions on all generators at the other end of the interval

$$\nabla_5 A_\mu|_{x_5=L} = 0, \tag{2.41}$$

we ensure that after KK decomposition there is only a zero mode for the standard model gauge bosons.

Let us suppose that Y_μ is the candidate gauge boson which is neutral under the SM. We can add a brane localised interaction with the inflaton

$$S_{\text{brane}} \supset \frac{c_1}{\Lambda_{5D}^{4.5}} \int d^4x \nabla_\mu \phi (\nabla_\rho \phi)^2 e^{-i\phi/\Lambda} \nabla_5 Y_\mu + h.c. \tag{2.42}$$

where $\Lambda_{5D} = 16\pi^2/3g_5^2$ is the UV cutoff of the 5D EFT. We can map this coupling onto our spin-1 EFT by performing the KK decomposition

$$Y_\mu(x, x_5) = \sqrt{\frac{2}{L}} \sum_{n=1}^{\infty} \sin\left(\frac{(n-1/2)\pi x_5}{L}\right) Y_\mu^{(n)}(x). \tag{2.43}$$

For the first KK mode, we can plug this into equation (2.42), and identify

$$c = \frac{c_1 \pi}{\sqrt{2}} \left(\frac{\Lambda}{\Lambda_{5D}}\right)^3 (\Lambda_{5\text{D}} L)^{-3/2}. \tag{2.44}$$

It turns out that for any reasonable value of the chemical potential, the strongest constraint on the strength of the coupling was always equation (2.36), $c < m/\Lambda$. After KK decomposition the

masses of the KK modes are given by

$$m_n = \left(n - \frac{1}{2}\right) m_{KK}; \qquad m_{KK} = \frac{\pi}{L}, \tag{2.45}$$

so the first KK mode has $m_1 = \pi/2L$. We can recast the coupling

$$c = \frac{2}{\sqrt{\pi}} c_1 \left(\frac{m_1}{\Lambda}\right)^{3/2} \left(\frac{\Lambda}{\Lambda_{5D}}\right)^{9/2}. \tag{2.46}$$

If we assume c_1 is an order one number, and the cutoff of the extra dimensional EFT is around the same scale as the cutoff of the inflationary EFT then we find that the size of the coupling is smaller than the constraint (2.36) by a modest suppression of $\sqrt{m_1/\Lambda}$, and we can almost reach the maximum coupling in figure 2.2. The full model building of a supersymmetric implementation of the coupling was left for future work.

In this chapter we have seen that the chemical potential mechanism can lead to observably large signals due to the tree-level exchange of a spin-1 particle. As a target we considered the heavy gauge bosons predicted by orbifold GUTs. Compared to a previous study [49], in which the authors found that only the KK graviton had an observably large signal, the signal size for gauge bosons with a chemical potential is large enough to be detected in future LSS surveys, and 21-cm tomography experiments. In addition to this the chemical potential mechanism brings the supersymmetric orbifold GUTs into play.

The only drawback to this mechanism is the requirement of a SM neutral complex gauge boson, which prevented us from considering some of the classical grand unified gauge groups like $SU(5)$, and $SO(10)$. The requirement of neutrality is a consequence of the fact that the inflaton

is neutral, and it is a general property that searching for charged new physics will require us to calculate a loop diagram. In the next chapter we will take on this challenge, and consider a target for cosmological collider physics within the classic $SU(5)$ scenario for grand unification.

Chapter 3

New Physics with Non-Zero Charge

This chapter is based on the paper "Charged Loops at the Cosmological Collider with Chemical Potential" [36], which can be found in its original form in Appendix B. In this paper we apply the chemical potential mechanism to a pair of charged scalars, which at leading order appear in a loop diagram. Loops are very hard to calculate in a time-dependent spacetime, so we need a good reason to go to the effort of computing something so arduous. For previous papers on calculating loops in de-Sitter space consult [34, 35, 37, 38, 53–67].

The reason to care about loops is the same reason we were forced to consider the trinification gauge group in the previous chapter. Any heavy particle that appears in a tree-level exchange must be neutral, because the inflaton is neutral. It is a robust possibility that new particles could carry some conserved quantum number, and would therefore have to appear in loops, such that charge is conserved at each vertex. In this chapter we will write down a model of a charged scalar for which a mathematical trick can be utilised to tame the leading order loop diagram that contributes to the bispectrum.

We will then consider as an application the detection of the Higgs triplets within supersymmetric orbifold GUTs. We will see that the loop signal is far smaller than the tree level exchange, and it is possible to obtain signals $f_{\rm NL} \sim \mathcal{O}(10^{-2})$. This size of signal is still above the gravitational floor, and it is possible that 21-cm tomography could probe signals of this size. We will also find that in addition to the oscillatory signal there are a number of sources of analytic background, which are far larger than the signal.

3.1 Chemical Potential for Charged Scalars

Consider two complex scalars, χ_1, χ_2. We want to consider scalars with some non-zero conserved charge, so let us assume that the two scalars are oppositely charged under a $U(1)$ symmetry

$$\chi_1 \mapsto e^{i\theta}\chi_1; \qquad \chi_2 \mapsto e^{-i\theta}\chi_2. \tag{3.1}$$

To include a chemical potential for the pair of scalars we follows the rules set forth in chapter 2, and [18, 35]. We assume a shift symmetry for the inflaton that is mixed up with a rephasing symmetry for the scalars

$$\phi \mapsto \phi + c; \qquad \chi_{1,2} \mapsto e^{-ic/2\Lambda}\chi_{1,2}, \tag{3.2}$$

where the factor of $1/2$ in the exponential is conventional.

The leading order Lagrangian that satisfies all of the symmetry constraints is given by

$$\mathcal{L} = \sum_{i=1,2} \left(-|\nabla_\mu \chi_i|^2 - m_i^2|\chi_i|^2\right) + \alpha\chi_1\chi_2 e^{i\phi/\Lambda} + h.c. \tag{3.3}$$

In particular, there is no tadpole for either χ_1, or χ_2 as it would break the $U(1)$ symmetry under which the scalars are oppositely charged.

3.2 The Spectral Decomposition

To calculate the bispectrum we apply the in-in master formula (1.52). We define some general notation that will help with the book-keeping required for a loop computation, and originates from [30]. We introduce the variable, “Schwinger-Keldysh indices", labelling a Feynman

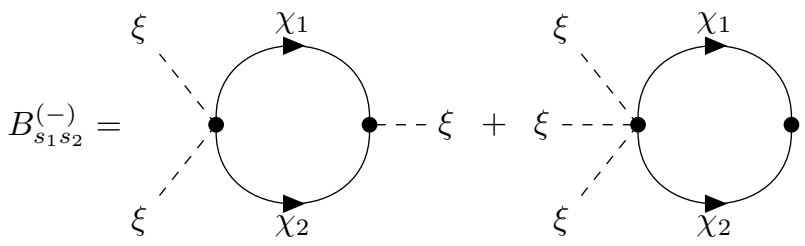


Figure 3.1: Feynman diagrams for the one-loop contribution to the bispectrum, due to the exchange of a pair of massive scalars. For each diagram there are 4 possible combinations of the pair (s_1, s_2). The third index labels the direction of the flow of χ_i-number. There are an additional two diagrams, $B^{(+)}_{s_1 s_2}$, with the arrows reversed.

diagram contributing to the bispectrum, $B^{(s_3)}_{s_1 s_2}$. Firstly, $s_{1,2} = \pm$, label the two vertices. A plus, as before, indicates that the vertex came from Taylor expanding the time-ordered exponential, and a minus indicates an origin in the anti time-ordered term. We also introduce a third index, $s_3 = \pm$, which further labels the direction of chi-number flow in the diagram. For example, if $s_3 = -$, then the left-hand vertex includes a factor of $\chi_1\chi_2$, and the right-handed vertex includes a factor of $\bar{\chi}_1\bar{\chi}_2$.

At leading order in α, there are two topologies of diagram that contribute, with an example illustrated in figure 3.1. We will see in the next section that with a judicious choice of field redefinition we can recast the contribution from the second kind of diagram into the first kind, leaving us with a single diagram to calculate (up to the multiplicity of varying the indices s_i).

These diagrams can both be simplified by use of the de-Sitter spectral density

$$\langle \mathcal{O}(x_1)\bar{\mathcal{O}}(x_2)\rangle_{s_1 s_2} = \int_0^\infty \rho^{dS}_{\mathcal{O}}(\mu)\langle \chi(\mu, x_1)\bar{\chi}(\mu, x_2)\rangle_{s_1 s_2}, \tag{3.4}$$

where $\chi(\mu, x)$ is a scalar field with mass parameter μ, and the notation $\langle\cdots\rangle_{s_1 s_2}$ is a way of labelling the different possible two-point functions. For example if $s_1 = s_2 = +$, then it is the

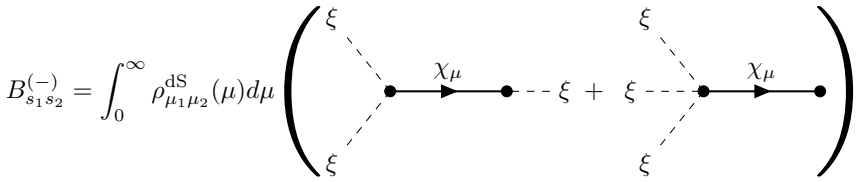


Figure 3.2: After applying the spectral density the loops diagrams turn into the tree-level exchange of a fictitious scalar, except we have to integrate over all possible masses for the exchanged particle, weighted by the spectral density.

time-ordered two points function etc. We will utilise the spectral density with $\mathcal{O} = \chi_1\chi_2$ to express the two diagrams in terms of tree-level exchanges, illustrated in figure 3.2. To make use of this "spectral representation" we have to know the weighting factor. The spectral density in de-Sitter space has been derived in numerous sources [57, 59, 68–70], and is given by

$$\rho^{\rm dS}_{\mu_1\mu_2}(\mu) = \frac{1}{8\pi^4}\frac{\mu\sinh\pi\mu}{\pi}\frac{\prod_{\epsilon,\epsilon_1,\epsilon_2=\pm}\Gamma\left[\frac{3}{4}+\frac{i}{2}\epsilon(\mu+\epsilon_1\mu_1+\epsilon_2\mu_2)\right]}{\Gamma\left(\frac{3}{2}-i\mu\right)\Gamma\left(\frac{3}{2}+i\mu\right)}. \tag{3.5}$$

A plot of the spectral density is given in figure 3.3 where we have also included the spectral density in Minkowski space, and an approximate form of the de-Sitter spectral density, valid near threshold, $\mu \approx \mu_1 + \mu_2$. In the limit where the two scalars combine to a pair with a very large invariant mass, the spectral density should be insensitive to the curvature of spacetime, and we recover the Minkowski expression. In Minkowski space the invariant mass of the pair cannot be less than the sum of their masses, but in de-Sitter space there is an exponentially suppressed tail for masses below threshold.

To obtain the threshold limit, we set $\mu = \mu_1 + \mu_2 + \epsilon$. All gamma functions except two can be replaced by the Stirling approximation $|\Gamma(x+iy)| \approx \sqrt{2\pi}|y|^{x-1/2}e^{-\pi|y|/2}$ [71], and we obtain

$$\rho^{\rm dS}_{\mu_1\mu_2}(\mu) \approx \frac{e^{\pi\epsilon/2}}{4\pi^3}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}}\Gamma\left(\frac{3}{4}+i\frac{\epsilon}{2}\right)\Gamma\left(\frac{3}{4}-i\frac{\epsilon}{2}\right), \tag{3.6}$$

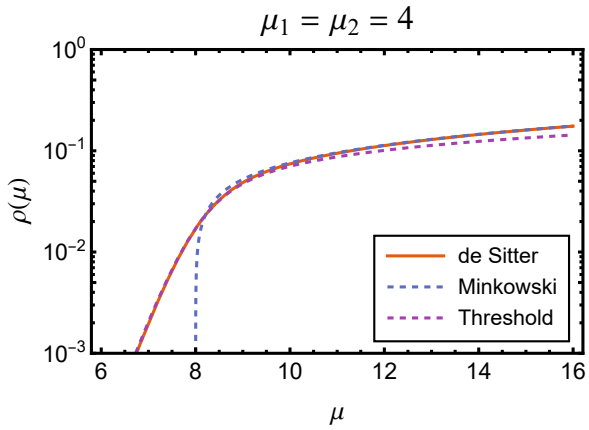


Figure 3.3: Taken from [36]/appendix B. The de-Sitter spectral density, compared with the Minkowski spectral density, and the near-threshold approximation, for $\mu_1 = \mu_2 = 4$.

where $\mu_{12} = \mu_1 + \mu_2$. It is also possible to derive this form for the spectral density directly from non-relativistic quantum mechanics in de-Sitter space. This derivation is given in appendix B.

3.3 Extracting the Leading non-analytic Signal

In figure 3.2 we make it clear that the spectral decomposition has morphed a loop diagram into a tree-level exchange with an extra integral. One question we can ask is whether the tree-level amplitude could be derived from a Lagrangian involving the fictitious scalar, χ, with mass parameter, μ? It turns out that the tree-level amplitudes can be derived from the same Lagrangian we introduced at the beginning of chapter 2, and was studied in [35]

$$\mathcal{L}_\chi = -|\nabla_\mu \chi|^2 - m^2|\chi|^2 + \alpha\chi e^{i\phi/\Lambda} + h.c. \tag{3.7}$$

where $\mu = \sqrt{m^2 - 9/4}$.

Since field redefinitions have no effect on amplitudes, we can perform a field redefinition of the fictitious scalar, χ, in the "fake" Lagrangian, without changing the amplitude. In $\mathcal{L}_\chi$, the

$$B^{(-)}_{s_1 s_2} = \int_0^\infty \rho^{\rm dS}_{\mu_1 \mu_2}(\mu) d\mu$$

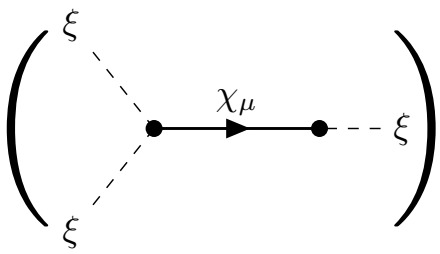


Figure 3.4: After a field redefinition of the fictitious scalar we are left with a single tree-level exchange diagram contributing to the bispectrum at leading order.

slow-roll background gives a vev, $\chi_0(\eta)$, to the scalar. If we shift by that vev $\chi \mapsto \chi_0 + \chi$, then the new Lagrangian no longer includes the second topology of diagram in figure 3.2. The field redefinition has left us with a single diagram, shown in figure 3.4

$$\langle \xi_{\vec{k}_1} \xi_{\vec{k}_2} \xi_{\vec{k}_3} \rangle' = \sum_{s_1, s_2, s_3} B^{(s_3)}_{s_1 s_2} \tag{3.8}$$

To perform the integrals over the invariant mass of the pair, μ, we need to calculate the tree-level amplitude exactly.

The full calculation is performed in appendix B, but to extract the non-analytic piece, i.e the signal in cosmological collider physics, we only need the stationary phase approximation to the tree-level amplitude. As in chapter 2, we find that the leading diagram comes from $B^{(-)}_{-+}$, and the amplitude is given by [35]

$$B^{(-)}_{-+} + B^{(+)}_{+-} = \frac{\pi}{16 k_1^3 k_2^3} \frac{|\alpha|^2}{\Lambda^3} \frac{\lambda^{1/2} \mu^{-1/2}}{\lambda^2 - \mu^2} p^{3/2 - i(\lambda - \mu)} e^{i \vartheta_*(\lambda, \mu)} + \text{c.c.}, \tag{3.9}$$

where the phase $\vartheta_*(\lambda, \mu) = \lambda \log 2\lambda + \mu \log 2\mu - (\lambda + \mu) \log(\lambda + \mu)$. Relative to the previous chapter we are unable to drop an overall $\mu-$dependent phase, as this may be important in the

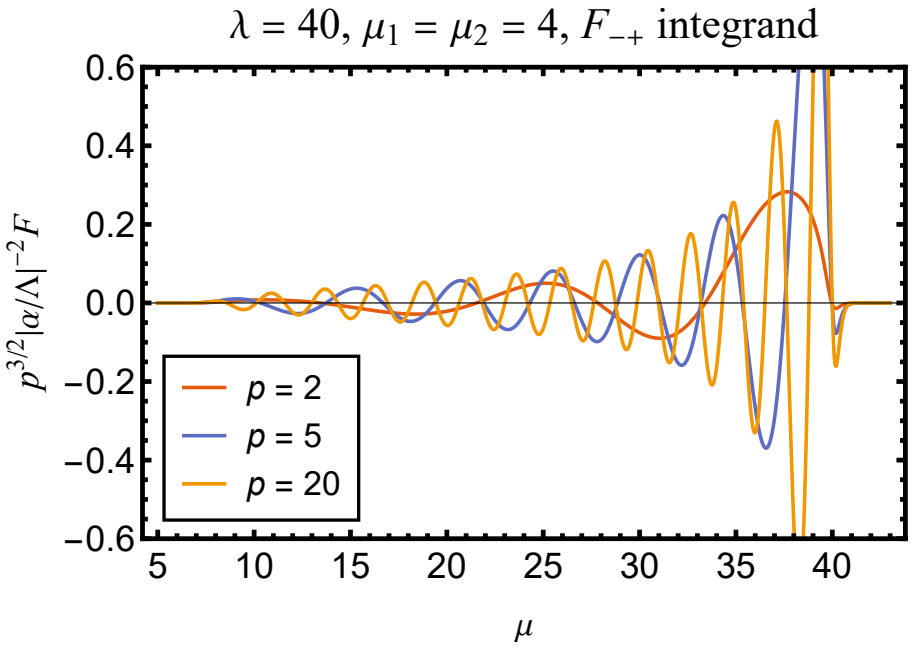


Figure 3.5: Taken from [36]/Appendix B. The integrand for the contribution to $B^{(-)}_{-+}$. The integrand oscillates rapidly before exponentially falling off at either end, when $\mu = \mu_1 + \mu_2$, or λ. In the absence of a stationary phase, integrals of this form are dominated by the endpoints.

integral over μ.

Recall that equation (3.9) is only valid if $\mu < \lambda$. It turns out that in the exact amplitude there is a peak as $\mu \to \lambda$, followed by a quick exponential decline for $\mu > \lambda$. The integrand for a few choices of parameter is illustrated in figure 3.5. The integrand is highly oscillatory in μ, except at $\mu = \mu_1 + \mu_2$, or λ. At the endpoints the integrand falls off exponentially. Integrals like this are dominated by the regions around the endpoints.

Both of these endpoints correspond to physical thresholds in Minkowski space. It is impossible to combine two particle of mass μ_1, μ_2 and obtain an invariant mass $\mu < \mu_1 + \mu_2$. There is not enough energy to produce particles with mass greater than the chemical potential, $\mu > \lambda$. In de-Sitter space there is an exponentially suppressed leak a little bit beyond either of these thresholds, but it makes sense physically that these are the two regions of interest. The

non-analytic piece of the signal is proportional to $p^{i(\mu-\lambda)}$, so the contribution corresponding to $\mu \approx \mu_1 + \mu_2$ will provide the leading non-analytic signal with phase $p^{i(\mu_1+\mu_2-\lambda)}$. The other piece, $\mu \approx \lambda$, will contribute an analytic background.

We will now extract the analytic contribution from the threshold region. We are free to use the stationary phase approximation (3.9) for the tree-level amplitude, and the threshold limit (3.6) for the spectral density. We then modify the integrand in such a way that it is unmodified in the region $\mu \approx \mu_1 + \mu_2 \equiv \mu_{12}$, but the contributions from other regions are removed

$$B^{(-)}_{-+} \approx B^{(-)}_{-+,\text{Tree}} \int_{-\infty}^{\infty} \rho^{\text{dS}}_{\mu_1\mu_2}(\mu_{12}+\epsilon)\beta^{i\epsilon} d\epsilon, \tag{3.10}$$

where $\beta = 2\mu_{12}p/(\lambda+\mu_{12})$, is chosen to match the phase of the integrand as $\mu \to \mu_{12}$. This form of the integral ensures that for $\epsilon \neq 0$ the rapid oscillations cancel out.

The integral can be computed analytically, the details of which can be found in appendix B. We ultimately normalise the bispectrum as in (1.39). To analyse the qualitative features we should compare the loop and tree-level chemical potential signals side-by-side

$$\begin{aligned} F_{\text{tree}} &= \frac{5\pi}{6}\frac{|\alpha|^2}{\Lambda^2}\sqrt{\frac{\lambda}{\mu}}\frac{\lambda}{\lambda^2-\mu^2}p^{-3/2+i(\mu-\lambda)} + c.c. \\ F_{\text{loop}} &= \frac{5}{24\sqrt{2\pi}}\frac{|\alpha|^2}{\Lambda^2}\sqrt{\frac{\lambda}{\mu_{12}}}\frac{\lambda}{\lambda^2-\mu_{12}^2}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}}\left(\frac{\lambda+\mu_{12}}{\mu_{12}}\right)^{3/2} p^{-3+i(\mu_{12}-\lambda)} + c.c. \end{aligned} \tag{3.11}$$

The loop-level result differs from the tree-level result by the additional of a factor of the square root of the reduced mass, and an additional factor involving the chemical potential. There is also an increased decay, $F_{\text{loop}} \propto p^{-3}$, relative to the tree-level result reflecting the dilution of two

massive particles as they propagate in an expanding spacetime. Finding the correct parametric dependence on these parameters is an achievement for a loop level process, and cannot be reproduced by estimations of the loop integrals [34, 35, 37, 38, 53, 54].

In addition to the non-analytic signal (3.11), there are many sources of analytic (or at least non-oscillatory) background:

- As we mentioned, there is a piece of the $B_{-+}^{(-)}$ integral that comes from the $\mu \approx \lambda$ region. It turns out that this contribution is around ten times as large as the non-analytic piece (3.11). The bispectrum therefore resembles a high frequency oscillation on top of a large smooth background.

- There is also a comparably large background coming from the B_{++} diagrams.

- In addition to the two pieces mentioned above there is a "combinatoric background" from the $B_{-+}^{(-)}$ diagram, but with the values of the momentum on the external legs permuted. These diagrams also lead to a large analytic background.

We leave the treatment of calculating these backgrounds to appendix B, and content ourselves with a summary of the main constraints derived from these backgrounds in the next section.

3.4 Results & Constraints

There are several generic constraints on the possible signal size due to the exchange of a loop of scalars. Firstly, the background is much larger than the signal. For a generic point in parameter space the background is around 10 times as large as the signal. The existing constraints on $f_{\rm NL} < \mathcal{O}(10)$ therefore imply a rather stringent constraint on the size of the

signal, which depends on the signal to noise ratio[1].

There are two theoretical constraints that emerge from ensuring that only the leading order contribution in α is important. If we set the inflaton to its slow-roll background in equation (3.3) the interaction is a time-dependent mass mixing term. We worked in the perturbative limit, in which we did not include the effect of this mass mixing in the propagator for the heavy scalars. Resumming the mass mixing into the propagator has two effects:

- It can lead to parametric resonance, increasing the amplitude of the propagator by a factor $\mathcal{A}_r$. We must ensure that these corrections are small, $\log \mathcal{A}_r \ll 1$, which leads to a constraint

$$|\alpha| \ll \sqrt{\frac{\prod_{\epsilon_1,\epsilon_2}(\lambda + \epsilon_1\mu_1 + \epsilon_2\mu_2)}{2\pi\lambda}}, \tag{3.12}$$

  where $\epsilon_i = \pm 1$.

- The mass mixing also corrects the frequency of the propagator, which would lead to a different non-analytic signal. To ignore these higher order effect, $\delta\omega \ll \omega$ leads to

$$|\alpha| \ll \lambda\mu. \tag{3.13}$$

A derivation of both of these conditions is given in [36] and appendix B.

Finally, the inflaton interactions can correct the amplitude of the power spectrum (1.33). There are a few contributions to the power spectrum, all calculated in appendix B. We must ensure $\delta\mathcal{P} \ll \mathcal{P}$. The four constraints that I have just mentioned are shown for a few choices of

[1]It seems quite likely that we are being unnecessarily strict with this constraint. It turns out that the constraint $f_{\rm NL} < \mathcal{O}(10)$ is far more shape dependent than we thought, and in a recent paper [23], 1.7σ evidence was found for a cosmological collider signal with $f_{\rm NL} \sim 200$. It is therefore likely that we could have significantly relaxed this constraint.

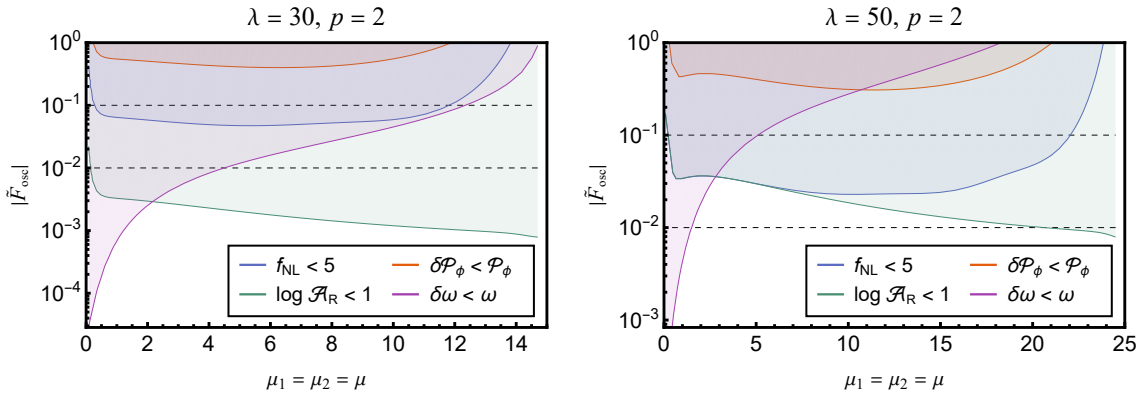


Figure 3.6: Taken from [36]/appendix B. The maximum size of the non-analytic signal in the equilateral triangle configuration. We show the four constraints mentioned in this section with the specific values: (i) $f_{\rm NL} < 5$, (ii) $\delta\mathcal{P} < \mathcal{P}$, (iii) $\log \mathcal{A}_r < 1$, (iv) $\delta\omega < \omega$.

parameters in figure 3.6. It appears that for almost the entire parameter space the most stringent constraint comes from parametric resonance. This constraint is not really a theoretical constraint on the validity of the model, it is only an issue with our ability to calculate.

If we had included the mass mixing in the propagator then the isometries of de-Sitter space would have been partially broken, and we could no longer use the spectral density to calculate the loop. It is clearly desirable to try to push into the green region, and we leave that calculation to future work. If we stay outside the parametric resonance regime then it appears that the maximum signal size varies between $f_{\rm NL} \sim \mathcal{O}(10^{-3} - 10^{-2})$, which is on the edge of what is possible to observe.

We can do slightly better than figure 3.6 by considering complex scalars, χ_1, χ_2 with multiplicity N_χ, originating from representations of a larger non-Abelian group, rather than being charged under $U(1)_V$. In this scenario the size of the bispectrum is multiplied by N_χ, but the parametric resonance constraint on α is unmodified. Figure 3.7 shows how including a non-zero multiplicity can open up new regions of parameter space observable by future experiments.

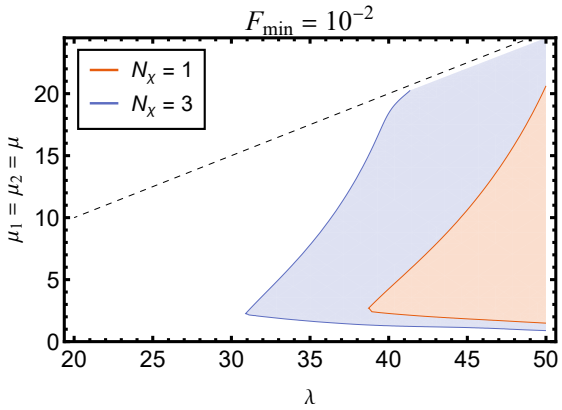


Figure 3.7: Taken from [36]/appendix B. A comparison of the observable parameter space assuming $F_{\text{Min}} = 10^{-2}$ between a single pair of complex scalar, and two sets of complex scalar with multiplicity, $N_\chi = 3$. We picked $N_\chi = 3$ because that is the case in our application in the next section.

It may seem like we introduced the multiplicity for the sole purpose of being observable, but a non-zero multiplicity is actually the norm, and we will see in the next section that our intended application to grand unification will naturally come with $N_\chi = 3$.

3.5 Embedding in orbifold SUSY GUTs

In this section we will discuss how to apply the loop bispectrum calculation to search for particles from orbifold SUSY GUTs. Firstly, note that de-Sitter space is incompatible with SUSY, so we must break SUSY during inflation. However, since we expect the SUSY breaking scale, $m_{\text{SUSY}} \gtrsim$ TeV [72], we must restore SUSY at the end of inflation. To achieve this we follow the construction in [73, 74]. We consider the inflaton to be part of a chiral superfield

$$\Phi = \frac{1}{\sqrt{2}}(\sigma + i\phi) + \cdots , \tag{3.14}$$

where we have dropped the fermionic superpartner of the inflaton, and the F-term. We identify ϕ as the inflaton, and σ is often called the "sinflaton". If we assume that Φ always appears in the Kahler potential in the combination, $\Phi + \bar{\Phi}$, then we can avoid a potential for ϕ.

To include the slow-roll potential we add another chiral superfield, S, which appears in the superpotential as

$$\mathcal{L} \supset -\int d^2\theta S f(-i\sqrt{2}\Phi) + c.c. \supset -F_S f(\phi) + c.c. \tag{3.15}$$

where F_S is the auxiliary field associated with S. After integrating out F_S we identify the slow-roll potential

$$V(\phi) = |f(\phi)|^2 = 3M_{\rm pl}^2 H^2, \tag{3.16}$$

where in the final step we have used the approximate background equations during slow-roll, equation (1.10). In appendix B it is shown that one can make S, σ heavy, and that it is appropriate to ignore supergravity corrections.

We couple the inflaton to an $SU(5)$ orbifold SUSY GUT, studied in [48, 75–77], with the setup illustrated in figure 3.8. SUSY requires two chiral superfields H_u, H_d to represent the Higgs field of the standard model for two reasons. Firstly, only holomorphic terms can appear in the superpotential, so giving masses to both the up and down type quarks requires two separate Higgs fields. Secondly, to ensure anomaly cancellation for standard model gauge invariance. Embedding this setup in an $SU(5)$ GUT implies these two Higgs fields are part of an $SU(5)$

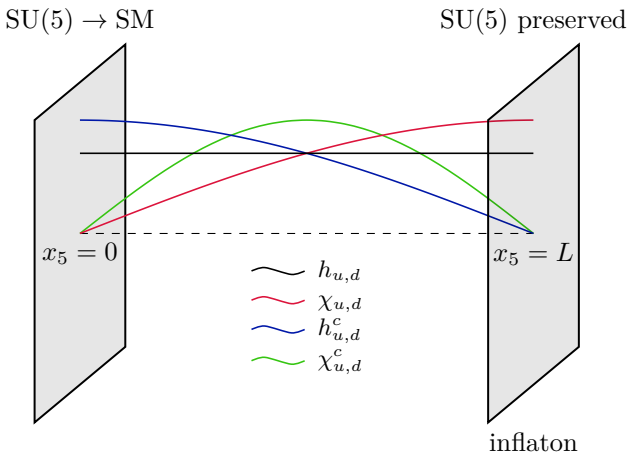


Figure 3.8: Taken from [36]/appendix B. The extra dimensional setup, with the profiles of lightest modes of the various Higgs fields. At $x_5 = 0$ the boundary conditions break the $SU(5)$ gauge symmetry down to the standard model in same way as chapter 2. The boundary conditions at $x_5 = L$ preserve the full gauge symmetry, but break the $\mathcal{N} = 2$ SUSY down to $\mathcal{N} = 1$. The inflaton is chosen to live on the brane at $x_5 = L$ to make it easy to couple to the KK modes of the coloured Higgs triplets.

multiplet

$$H_u = \begin{pmatrix} \chi_u \\ \\ h_u \end{pmatrix}; \qquad H_d = \begin{pmatrix} \chi_d \\ \\ h_d \end{pmatrix}, \tag{3.17}$$

where H_u is a fundamental of $SU(5)$, and H_d is an anti-fundamental. The top entry $\chi_{u,d}$ are the coloured Higgs triplets which motivated us to consider an orbifold setup in chapter 2.

In 5D there is an additional complication. $\mathcal{N} = 1$ supersymmetry in 5D descends down to $\mathcal{N} = 2$ SUSY in 4D. This implies that there are additional chiral superfields, $h^c_{u,d}$, $\chi^c_{u,d}$. We use the superscript, c, because the additional superfield must be oppositely charged under gauge symmetries relative to its original.

We chose boundary conditions at $x_5 = 0$ to break the $SU(5)$ gauge symmetry down to the standard model in exactly the same way as in chapter 2. We impose Dirichlet boundary conditions for these additional superfields, h^c, χ^c, at $x_5 = L$ to break the $\mathcal{N} = 2$ SUSY down

to $\mathcal{N}=1$. We assume that the inflaton lives on the brane at $x_5 = L$ to facilitate coupling to the coloured Higgs triplet fields. The various profiles for the lightest modes of each Higgs field are shown in figure 3.8. The boundary conditions and KK mass spectrum for each Higgs field is shown in table 3.1. We have also assumed the existence of a $U(1)_R$ symmetry which is crucial for ensuring the orbifold setup is not ruled out by proton decay bounds [75–77].

Particles	Boundary conditions	KK spectrum	$\mathrm{U}(1)_R$
h_u, h_d	$(+,+)$	nM_{C}	0
χ_u, χ_d	$(-,+)$	$(n+1/2)M_{\mathrm{C}}$	0
h_u^c, h_d^c	$(+,-)$	$(n+1/2)M_{\mathrm{C}}$	2
χ_u^c, χ_d^c	$(-,-)$	$(n+1)M_{\mathrm{C}}$	2

Table 3.1: Taken from [36]/appendix B. A list of the various boundary conditions, mass spectrum, and $U(1)_R$ charges of the Higgs fields in the orbifold setup. A positive sign indicates imposing Neumann BCs on the brane, and a negative sign indicates Dirichlet BCs.

Finally we must realise the chemical potential couplings between the inflaton and the coloured Higgs triplets. For a discussion of how to arrive at this term consult appendix B, we will content ourselves with writing out the final answer

$$\mathcal{L} \supset \int d^4\theta \left[\frac{\kappa_1}{\Lambda_5^3} \chi_u \chi_d \bar{S} S \exp\left(\frac{\sqrt{2}\Phi}{\Lambda}\right) + c.c. \right] \supset \frac{2\kappa_1}{\pi} \frac{m_{\mathrm{KK}}}{\Lambda_5^3} |F_s|^2 \chi_u^{(0)} \chi_d^{(0)} e^{i\phi/\Lambda} + c.c. \tag{3.18}$$

We can read off the mapping onto the EFT we used in the rest of paper, identifying the coupling

$$\alpha = \frac{6\kappa_1}{\pi} \frac{m_{\mathrm{KK}} M_{\mathrm{pl}}^2 H^2}{\Lambda_5^3}, \tag{3.19}$$

where we have used equation (3.16). The analysis of the previous section should now apply with

the identification, $\chi_1 = \chi_u^{(0)}$, $\chi_2 = \chi_d^{(0)}$. The best fit value for the KK scale is $m_{\rm KK} \sim 10^{15}$ GeV, and the light KK modes of the Higgs triplets can be read off from table 3.1 $m_\chi \sim 5 \times 10^{14}$ GeV, which is within the range of the chemical potential mechanism.

In this chapter, motivated by the results of chapter 2, we have analysed the prospects of searching for cosmological collider signals due to a loop process involving the exchange of a pair of scalars that have been given a chemical potential. The general motivation for this was the acknowledgement that a robust possibility for new physics is that it may be charged under some conserved quantum number, which will require the new physics to appear in loops. We applied this idea to searching for the Higgs triplet particles in supersymmetric orbifold GUTs, and found that detecting the signal in large scale structure surveys is challenging, but there is a reasonable prospect for detection in 21-cm tomography experiments. However, we also noted that it may be possible to obtain larger signals within a region of parameter space where our ability to calculate breaks down. In the next chapter we will switch gears, and consider a radical alternative scenario for how the early universe played out.

Chapter 4

A Calculable Model of Warm Inflation

In the previous two chapters we have considered the prospects of searching for heavy new particles produced during inflation. All of this was computed within the framework of slow-roll inflation, assuming a very high Hubble scale, $H \lesssim 5 \times 10^{13}$ GeV. Within the slow-roll inflationary framework the quantum fluctuations of the inflaton provide the initial conditions for the perturbations on top of a homogeneous isotropic spacetime, which then grow into the large scale structure of our universe.

It is a rather profound statement to say that all of the structure in the universe originated from quantum mechanical fluctuations. Is it possible to test this? An ironclad test of the quantum mechanical nature of perturbations, something like a Bell inequality test, is unlikely as we do not seem to have access to data about non-commuting observables during inflation. The best we can do is to ask if it is possible to explain the data another way.

This chapter is dedicated to answering that question by modelling a "classical" origin for perturbations. In "warm inflation" the physical picture is very similar to "cold inflation", which is what we will use to refer to the class of inflationary models we have studied so far. There is still an inflaton that develops a slow-roll background. However, the universe is no longer assumed to be empty, but filled with a high temperature gas of radiation that interacts with the inflaton. The interactions of the radiation with the inflaton provide random kicks, analogous to random kicks in Brownian motion, which give an origin for perturbations as thermal fluctuations.

The idea of warm inflation is many decades old [78–81], but it has been difficult to construct a model within which to reliably calculate the predictions of warm inflation. In this chapter we will study a simple perturbative model, and in the context of this model we will understand the key features of warm inflation.

4.1 The Framework of Warm Inflation

This section follows the treatment in [82, 83]. The idea of warm inflation is to include a radiation bath that is present during inflation. The immediate issue with this idea is that the energy density of radiation dilutes, $\rho_r \sim a^{-4}$. The exponential expansion of the universe during inflation therefore dilutes away the radiation extremely quickly. The way to get around this issue is to couple the radiation to the inflaton in such a way that energy is constantly transferred from the inflaton into the radiation, heating it up as the universe tries to cool it down. The background equations of motion during warm inflation are:

$$H^2 = \frac{1}{3M_{\rm pl}^2}\left(V(\phi) + \frac{1}{2}\dot{\phi}^2 + \rho_r\right), \tag{4.1}$$

$$\ddot{\phi} + (3H + \gamma)\dot{\phi} + V'(\phi) = 0, \tag{4.2}$$

$$\dot{\rho}_r + 4H\rho_r = \gamma\dot{\phi}^2, \tag{4.3}$$

where ρ_r is the radiation energy density.

There are three modifications of the background equations relative to (1.10). Firstly, we have included the energy density of the radiation in the Friedmann equation, in addition to the energy density of the inflaton field. Secondly, we have modelled the interaction of the inflaton with the radiation as providing an additional source of friction in the inflaton equation of motion. Finally, there is a third equation which governs how the radiation evolves. The first two terms in equation (4.3) are equivalent to the statement that radiation dilutes as a^{-4} in the absence of interactions. The right hand side is modelling the energy injection into the thermal bath of the energy that is taken out of the inflaton by the friction term, γ. If γ is large enough then we may

find a quasi-steady state solution with $\dot{\rho}_r \approx 0$.

Slow-roll inflation is now able to proceed in almost the same way as in the previous chapters. For the universe to expand exponentially, Hubble must be approximately constant, so we assume that the energy density of the universe is vacuum dominated. In order that inflation lasts long enough to solve the Horizon problem we also assume that the second derivative can be dropped in the second equation

$$V(\phi) \gg \dot{\phi}^2, \rho_r; \qquad \ddot{\phi} \ll 3H(1+Q)\dot{\phi}, \tag{4.4}$$

where we have defined the ratio of the thermal friction to Hubble friction

$$Q \equiv \frac{\gamma}{3H}. \tag{4.5}$$

It is conventional to call warm inflation with $Q < 1$ the "weak regime", and warm inflation with $Q > 1$ the "strong regime". Relative to chapter 1 the slow-roll parameters are modified

$$\epsilon_V = \frac{M_{\rm pl}^2}{2(1+Q)}\left(\frac{V'}{V}\right)^2, \tag{4.6}$$

$$\eta_V = \frac{M_{\rm pl}^2}{1+Q}\frac{V''}{V}. \tag{4.7}$$

The interesting thing to note about the modified slow-roll parameters is that in the strong regime, if $Q \gg 1$, the parameters gain an extra suppression. It may therefore be possible to slow-roll along steeper potentials than in the cold inflationary scenario, which may be desirable from a UV perspective.

We have summarised, at the macroscopic level, what background equations we would like to derive from a model of warm inflation. The desires seem benign enough, so why is warm inflation so hard to model build? The key issue is that at finite temperature there is a correction to the slow-roll potential of the inflaton

$$V(\phi) \to V(\phi) + V_{\rm Therm}(\phi, T). \tag{4.8}$$

The finite temperature correction to the potential generically modifies the shape of the potential so drastically that it cannot support slow-roll [84]. One way to get around this is to try to cancel thermal contributions to the potential [85, 86].

However, we will take the approach of considering psuedo-scalar couplings [83, 87–91]. In fact, we will see that the models of inflation we have already studied in chapters 2, and 3 provide the blueprint for how to couple the inflaton to the radiation without generating a large thermal back-reaction that spoils the whole idea.

4.2 Failure is a Dish Best Served Warm

In this section we will shoot for the stars and try to build a model of warm inflation with no finite temperature correction to the slow-roll potential. This will not work. However, we will learn many important lessons about how warm inflation works, and we will land on a cloud, because a slight modification of this model will turn out to work perfectly. We begin by recalling that the chemical potential couplings of our cosmological collider physics models lead to both unsuppressed particle production, and protected the slow-roll potential of the inflaton. This

seems like a good place to begin

$$\mathcal{L}_{\text{int}} = \frac{1}{f}\partial_\mu \phi J^\mu. \tag{4.9}$$

To make calculations easy we will assume the radiation sector is composed of a complex scalar field, χ.

Recall from chapter 2 that the chemical potential coupling has no effect unless the $U(1)$ of the complex scalar is explicitly broken, and that calculations are easiest in a rotated basis $\chi \mapsto e^{i\phi/f}\chi$. We introduce a quartic $U(1)$ breaking interaction

$$\mathcal{L}_{\text{int}} = -\frac{\lambda_\chi}{4!}|\chi|^4 - \frac{\alpha}{4!}\chi^4 e^{4i\phi/f} + h.c. \tag{4.10}$$

where α is a perturbative quartic coupling, $\alpha < 16\pi^2$, and λ_χ is a self-interaction term that keeps the radiation sector in thermal equilibrium.

We are looking for a quasi steady state solution to the equations of motion. We therefore begin by assuming that the χ-fields are in equilibrium with temperature, T, and chemical potential, μ. We also assume a slow-roll background for the inflaton $\dot{\varphi} \approx$ const. Our next step is to calculate the energy injection from the inflaton into the radiation sector, and check that it can maintain the radiation bath as the universe expands. To do this we look at equation (4.3)

$$\dot{\rho}_r + 4H\rho_r = \gamma\dot{\phi}^2, \tag{4.11}$$

and recognise this as a Boltzmann equation, where the right hand side is the collision term[1].

The collision term consists of any scattering processes that net injects or removes en-

[1]One can also calculate the friction coefficient, γ, directly from equation (4.2). This calculation is performed in appendix C, and leads to the same result as the Boltzmann equation calculation.

ergy from the radiation sector. The leading contribution comes from the two-to-two scattering processes

$$\chi(p_1) + \chi(p_2) \rightleftharpoons \bar{\chi}(p_3) + \bar{\chi}(p_4). \tag{4.12}$$

After expanding around the slow-roll background in (4.10) the sole impact of the phase is to modify the energy conservation delta function

$$\delta(E_1 + E_2 - E_3 - E_4) \mapsto \delta\left(E_1 + E_2 - E_3 - E_4 + 4\dot{\varphi}/f\right). \tag{4.13}$$

During the scattering processes, if two particles scatter into two anti-particles, $4\dot{\varphi}/f$ energy is injected into the radiation sector. However, in the reverse process energy is removed.

For a derivation of the collision term due to two-to-two scattering consult [4]. We can apply the general result to our model and obtain

$$\begin{aligned}\dot{\rho}_r = 4\frac{\dot{\varphi}}{f}\alpha^2 \int d\Pi_{12} \int d\Pi_{34}(2\pi)^4\delta^3(\vec{p}_{\rm tot})\delta(E_1 + E_2 - E_3 - E_4 + 4\dot{\varphi}/f) \\ \times(\bar{f}_3\bar{f}_4(1+f_1)(1+f_2) - f_1 f_2(1+\bar{f}_3)(1+\bar{f}_4)).\end{aligned} \tag{4.14}$$

The distribution functions are given by the Bose-Einstein distribution functions

$$f_i = \frac{1}{e^{\beta(\omega_i - \mu)} - 1} \implies \frac{1+f_i}{f_i} = e^{\beta(\omega_i - \mu)}, \tag{4.15}$$

$$\bar{f}_i = \frac{1}{e^{\beta(\omega_i + \mu)} - 1} \implies \frac{1+\bar{f}_i}{\bar{f}_i} = e^{\beta(\omega_i + \mu)}. \tag{4.16}$$

Note that we have given the particles and antiparticles opposite signed chemical potentials because we have introduced an applied chemical potential for χ-number, $\mathcal{L}_{\rm int} \supset \frac{\dot{\varphi}}{f} J^0$.

To solve the phase-space integral we must Taylor expand $\dot{\varphi} \ll fT$. We should conclude that for the interaction to act like an effective friction term in the inflaton equation of motion warm inflation is always a Taylor expansion in $\dot{\varphi}$. The phase-space integral can be solved numerically, and we find

$$\dot{\rho}_r = 4\frac{\dot{\varphi}}{f} c_1 \alpha^2 T^3 \left(4\frac{\dot{\varphi}}{f} - 4\mu \right), \tag{4.17}$$

where c_1 is given by $c_1 \approx (0.0589 - 0.0632 \log(m/T) - 0.143 e^{-m/T})/(2\pi)^4$. We can therefore infer that the friction coefficient is given by

$$\gamma = \frac{16 c_1 \alpha^2 T^3}{f^2} \left(1 - \frac{\mu f}{\dot{\varphi}} \right). \tag{4.18}$$

If we could ignore the second term in the brackets, then we need only plug this into the equations of motion and fit to the data. We would successfully have a model of warm inflation, and we could have completely sidestepped the theoretical issue of the finite temperature correction to the slow-roll potential! However, we know from the form of the interaction that $\mu = \dot{\varphi}/f$. If we plug this in $\gamma = 0$. It has all gone horribly wrong, and we are left wondering why bad things happen to mediocre people.

Thankfully we can understand at a fairly intuitive level what went wrong here. Let us assume for a second that we can turn off the slow-roll background, and begin with a bath of χ-particles in thermal equilibrium. At this point the χ-number violating scattering processes (4.12) are happening efficiently, but the rate in one direction is equal to the rate in the other direction, and there is no net energy injection. In this equilibrium there are as many particles as antiparticles, $\mu_{\rm dist} = 0$, where $\mu_{\rm dist}$ is the chemical potential that appears in the distribution

functions (4.15), (4.16). This is the statement that in equilibrium the chemical potential for a non-conserved charge is always $\mu_{\rm tot} = 0$, a statement that will remain true even when we turn on the slow-roll background.

Suppose we now turn on the slow-roll background, $\dot{\varphi} \neq 0$. We are now applying an external chemical potential $\mu_{\rm applied} = -\dot{\varphi}/f$ to χ-number. Due to the energy injection the rates in equation (4.12) receive some phase space suppression/enhancement such that the rate to scatter into particles is now greater than the rate to scatter into antiparticles. The number of particles now increases relative to antiparticles, and the only way the distributions can adjust to do this is to adjust $\mu_{\rm dist} \neq 0$. Eventually a new equilibrium is reached where the rates in equation (4.12) balance out. In this new equilibrium $\mu_{\rm tot} = \mu_{\rm dist} + \mu_{\rm applied} = 0 \implies \mu_{\rm dist} = \dot{\varphi}/f$. In this new equilibrium there is no net energy injection since the two rates have balanced out.

4.3 A Working Model

The insight for why the pure chemical potential model failed also tells us how to modify the model. Let us define

$$\Delta n_\chi = n_\chi - n_{\bar{\chi}}. \tag{4.19}$$

In chemical equilibrium the rates in (4.12) balance out, implying $\Delta \dot{n}_\chi = 0$. In the chemical potential model (4.10) the rate of energy injection is proportional to the rate of χ-number violation. We must break this proportionality to ensure there is still energy injection in chemical equilibrium. We try the following interaction terms

$$\mathcal{L}_{\rm int} = -\frac{\lambda_\chi}{4!}|\chi|^4 - \frac{\alpha_1}{4!}\chi^4 e^{4i\phi/f} - \frac{\alpha_2}{4!}\chi^4 e^{5i\phi/f} + h.c. \tag{4.20}$$

In this model the first interaction injects/removes $4\dot{\varphi}/f$ energy in χ-number violating interactions, and the second term injects/remove $5\dot{\varphi}/f$ energy.

It is now impossible to completely remove the interactions and swap them for a chemical potential term with a field redefinition. If we redefine $\chi \mapsto \chi^{-i\phi/f}$, or $\chi \mapsto \chi e^{-5i\phi/4f}$, then we can remove one of the interactions in favour of a chemical potential $\mu = -\dot{\varphi}/f$, or $-5\dot{\varphi}/4f$. It is therefore unclear what the value of the μ_{applied} is in the presence of both interactions. We can find the value of $\mu \equiv \mu_{\text{dist}} = -\mu_{\text{applied}}$ by calculating the Boltzmann equation for Δn_χ. We then find the friction term from the Boltzmann equation for ρ_r. There are two terms in each expression, one for each interaction, and we can pull them directly from equation (4.17)

$$\Delta\dot{n}_\chi = 4c_1\alpha_1^2T^3\left(4\frac{\dot{\varphi}}{f} - 4\mu\right) + 4c_1\alpha_2^2T^3\left(5\frac{\dot{\varphi}}{f} - 4\mu\right), \tag{4.21}$$

$$\dot{\rho}_r = 4\frac{\dot{\varphi}}{f}c_1\alpha_1^2T^3\left(4\frac{\dot{\varphi}}{f} - 4\mu\right) + 5\frac{\dot{\phi}}{f}c_1\alpha_2^2T^3\left(5\frac{\dot{\varphi}}{f} - 4\mu\right). \tag{4.22}$$

The chemical potential can be calculated from equation (4.21) by setting $\Delta\dot{n}_\chi = 0$

$$\mu = \frac{\alpha_1^2 + \frac{5}{4}\alpha_2^2}{\alpha_1^2 + \alpha_2^2}\frac{\dot{\varphi}}{f}, \tag{4.23}$$

which is an answer we probably could have guessed from the fact that if $\alpha_2 \to 0$, $\mu = \dot{\varphi}/f$, and if $\alpha_1 \to 0$, $\mu = 5\dot{\varphi}/4f$. The friction coefficient can now be inferred from equation (4.22) by plugging in the result for the chemical potential

$$\gamma = \frac{\alpha_1^2\alpha_2^2}{\alpha_1^2 + \alpha_2^2}\frac{c_1T^3}{f^2}. \tag{4.24}$$

The final result exhibits the phenomenon of "collective symmetry breaking". If either $\alpha_i \to 0$, the model still has a shift symmetry for the inflaton, and the friction coefficient $\gamma \to 0$. Whilst we have not given a general proof, these results are suggestive that having a non-zero thermal friction coefficient, γ, somehow requires the breaking of shift symmetry, and there is no way to sidestep the issue of the finite temperature potential in warm inflation.

It turns out that we can guess the form of the finite temperature potential up to an overall coefficient. Firstly, the two interactions in (4.20) have collectively broken the shift symmetry for ϕ. However, it has been broken down to a discreet shift

$$\phi \mapsto \phi + 2\pi f, \tag{4.25}$$

so we should expect the potential to be a trigonometric function of ϕ/f. By the collective symmetry breaking argument the potential should be proportional to $\alpha_1\alpha_2$. Finally, the complex scalar must be light, $m_\chi \ll T$, to act like radiation. The dimensions therefore can only be made up by the temperature, T. The full calculation of the finite temperature potential is given in appendix C, with the help of some results in [92, 93]

$$V_{\text{Therm}}(\phi, T) = \frac{1.72}{4!}\frac{\alpha_1\alpha_2}{16\pi^2}T^4 \cos\left(\frac{\phi}{f}\right). \tag{4.26}$$

The final result is consistent with the points made above, and all the work of the calculation is to obtain the factor of $1.72/(4! \times 16\pi^2)$.

The effect of this potential on slow-roll is illustrated in figure 4.1. Slow-roll inflation now occurs with a set of ultra high frequency wiggles on top of a slow-roll background. As long as

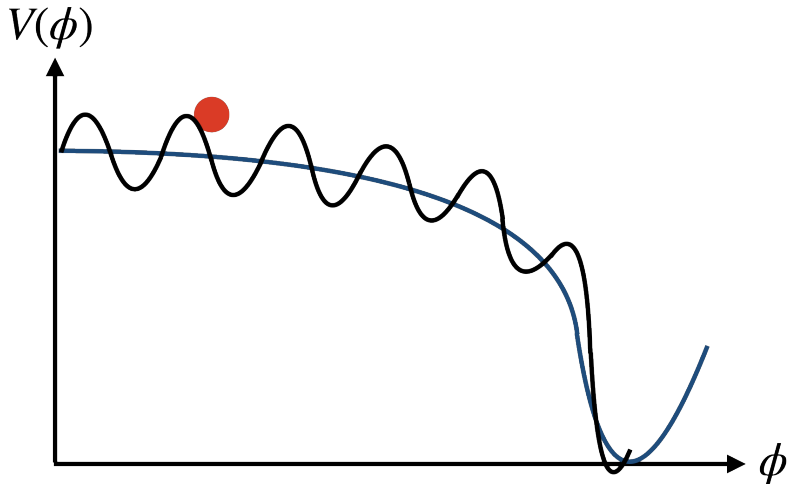


Figure 4.1: In the presence of the radiation slow-roll inflation occurs on a potential with high frequency wiggles on top of the slow-roll potential.

the inflaton has enough kinetic energy to easily roll over the wiggles, then they should average out to zero and the finite temperature potential can be ignored [94]

$$\frac{1}{2}\dot{\phi}^2 \gg \frac{1.72}{4!}\frac{\alpha_1\alpha_2}{16\pi^2}T^4. \tag{4.27}$$

We must impose this constraint on any set of parameters when we ultimately fit to data.

The presence of the high frequency wiggles leads to one further constraint. The high frequency potential can lead to parametric resonance of non-zero k-modes. The following derivation follows closely the treatment in [95]. If I expand around the background field $\phi = \varphi + \xi$, the equation of motion for a non-zero k mode is given by

$$\ddot{\xi}_k + \Gamma\dot{\xi}_k + \left(\frac{k^2}{a_0^2}e^{-2Ht} - A\cos\omega t\right)\xi_k = 0. \tag{4.28}$$

To zoom in on the effect of parametric resonance we have dropped both the slow-roll potential, and some extra terms that originate from interactions with the radiation sector that we will

need in the next section. The various constants are given by

$$\Gamma = 3H + \gamma; \qquad \omega = \frac{\dot{\varphi}}{f}; \qquad A = \frac{1.72}{4!}\frac{\alpha_1\alpha_2}{16\pi^2}\frac{T^4}{f^2}. \tag{4.29}$$

At the time, t_*, defined by

$$\frac{k}{a_0}e^{-Ht_*} \approx \frac{\omega}{2}, \tag{4.30}$$

the mode ξ_k, becomes unstable. To see this notice that the final term in (4.28) will have a component that acts as a periodic forcing term on resonance. During the instability the mode grows exponentially

$$\xi_k(t) = \xi_k(t_*)\exp\left(\frac{A}{2\omega}(t - t_*)\right), \tag{4.31}$$

provided we can ignore the friction term during the instability, $\Gamma/\omega \ll A/\omega^2$.

We can estimate the size of the exponential growth by finding the length of time the mode is unstable, which is given by

$$\Delta t = \frac{1}{H}\frac{2A}{\omega^2}. \tag{4.32}$$

The condition $A \ll \omega^2$ is the condition that the inflaton has enough kinetic energy to roll over the wiggles in the potential. The instability is therefore short relative to a Hubble time, further justifying our choice to ignore the friction term during the instability. We can plug the time period into the exponential growth factor to estimate that each k-mode grows by a factor

$$\xi_k \to \exp\left(\frac{1}{H}\frac{A^2}{\omega^3}\right)\xi_k. \tag{4.33}$$

If each k-mode is produced exponentially due to parametric resonance, then it will back-react

on both the geometry and the slow-roll background. If we don't want to worry about this then we must ensure the exponential growth is small,

$$A^2 \ll H\omega^3, \tag{4.34}$$

which is the final constraint that we can derive from the background equations. Roughly speaking, this constraint seems to limit us to the weak regime of warm inflation.

In this section we have found a model for warm inflation that works, calculated the thermal friction, and then derived several constraints that derive from the thermal back-reaction on the slow-roll potential. At the background level we need only find some choices of parameters that satisfy all the constraints and we have a working model of warm inflation. However, to determine if warm inflation can explain the data in our universe we have to fit the precise data measurements we have of A_s, and $n_s - 1$. This requires us to calculate the predictions for the power spectrum in warm inflation, which we will build up to over the next few sections.

4.4 Brownian Motion

In the next section we will consider fluctuations, $\phi = \varphi(t) + \xi(\vec{x}, t)$, and calculate the two-point function of the gauge-invariant scalar perturbation, ζ, due to thermal fluctuations. Before calculating the thermal fluctuations in warm inflation it is useful to review Brownian motion. Brownian motion is the random motion of particles in a fluid (originally pollen grains in water).

Let us work in 1D for simplicity, and call the position of a pollen grain, $x(t)$. As the pollen grain moves around in the fluid it gets kicked in a random direction every time it interact with a

fluid molecule. The pollen grain is more likely to be knocked back in the direction it is travelling than in any other direction. We separate out the force due to the random kicks into a net drag term, and a random force that averages to zero

$$m\ddot{x}(t) + \Gamma\dot{x}(t) = \Upsilon(t). \tag{4.35}$$

The second term is supposed to represent the random force applied to the pollen grain at the time, t. Since it is random, $\Upsilon(t)$[2] is NOT a function! In fact, it is impossible for us to predict the trajectory of the pollen grain[3]. Everything we can say about $\Upsilon(t)$ is statistical in nature. Firstly, we have separated out the net drag, Γ, so we know

$$\langle \Upsilon(t) \rangle = 0, \tag{4.36}$$

where the angle brackets indicate some kind of average, and we will be sloppy about exactly what that average means.

The next thing to note is that the timescale for collisions with the atomic molecules is very small relative to the timescales at which a human eye can observe the random motion of the pollen grain. It is also reasonable to assume that the direction a pollen grain gets knocked it at one time is independent of the direction it was knocked in at any previous time. The random force, $\Upsilon(t)$, therefore reflects the average effect of a large number of independent and identically distributed random kicks! We can therefore use the central limit theorem to assume that the statistics of $\Upsilon(t)$ are Gaussian! This is very encouraging because, as we noted in chapter 1,

[2]Sorry but I am running out of Greek letters.

[3]Note that this is classical uncertainty, not quantum mechanical, and is due to our ignorance about the positions of all the molecules in the fluid.

we know experimentally that the statistics of inflaton fluctuations are Gaussian to a very high degree! All of the information about the random force is therefore contained in the two-point function

$$\langle \Upsilon(t)\Upsilon(t')\rangle = A\delta(t-t'), \tag{4.37}$$

where A is some undetermined constant.

Remarkably, we can solve for the pollen's trajectory using Green's function techniques, assuming $\xi(t)$ is a source. In this case the Green's function is exactly solvable and we can write (setting m=1 for simplicity)

$$v(t) = v_0 e^{-\Gamma t} + e^{-\Gamma t}\int_0^t e^{\Gamma t'}\Upsilon(t'), \tag{4.38}$$

where $v(t) = \dot{x}(t)$. Formally speaking we have solved for the pollen's trajectory. We can calculate statistics about the trajectory by inheriting the statistical properties of the random force term, which I will refer to as the noise term from now on. Let us consider the two-point function of the velocity, in the limit that $t \to \infty$,

$$\langle v^2(t)\rangle \to \frac{A}{2\Gamma}, \tag{4.39}$$

where we have used the two-point function of the noise, and then computed the integral over time.

Let us take a minute to interpret this result. Both of the constants, Γ, and A are unknown constants that I have included in my model of Brownian motion. However, if we were to try to derive both of these constants from first principles we would find that they both emerge from

the same physical origin, the random kicks that the pollen grain gets from the fluid molecules, and we would therefore expect some relationship between the two coefficients. In fact, such a relationship exists, and is called the “fluctuation-dissipation" theorem! We can derive the fluctuation-dissipation theorem by imagining we placed an ensemble of pollen grains in the water, and let them evolve for some time.

If we interpret equation (4.39) as an ensemble average then we should expect that after a while the pollen grains will move in thermal equilibrium with the water, and should therefore satisfy the equipartition theorem

$$\frac{1}{2}\langle v^2(t)\rangle = \frac{1}{2}T \implies A = 2\Gamma\, T. \tag{4.40}$$

This equation is the statement of the fluctuation-dissipation theorem that is used in the warm inflation literature [96, 97]. We will find that a fluctuation-dissipation theorem does exist in warm inflation, but we will discover that the particular relation of equation (4.40), used in all the warm inflation literature, is modified by the chemical potential, μ.

There is one final lesson we can take from the Brownian motion example. In this case one could ask the question, do the pollen grains thermalise? If they thermalise do I somehow need to modify my equations to take this into account? The result we just derived says that in term of calculating statistics of the pollen grains trajectory it is immaterial whether the pollen grains thermalise. I simply solve the stochastic differential equation, and if the pollen grains thermalise (i.e t is very large), then it will be reflected in my results. When calculating the thermal fluctuations of the inflaton, we will derive an equation that is similar to the equation of motion for the pollen grain. Given what we have just noted, we are free to blindly calculate

the statistics of Υ by solving the equation of motion, and we will get the right result.

4.5 The Fluctuation-Dissipation Theorem with a Chemical Potential

To calculate the prediction for the CMB power spectrum during inflation we must consider fluctuations on top of the isotropic homogeneous background. We must first identify the equivalent of the noise term. To find this we write down the equation of motion for the inflaton, including the interaction (4.20)

$$\ddot{\phi} + 3H\dot{\phi} - \frac{\nabla^2}{a^2}\phi + V'(\phi) = \mathcal{L}'_{\rm int}(\phi). \tag{4.41}$$

In the previous section we used the Boltzmann equation to calculate the thermal average of the right hand side

$$\langle \mathcal{L}'_{\rm int}(\phi) \rangle = -\gamma\dot{\phi}. \tag{4.42}$$

In analogy with Brownian motion we assume that the interaction between the inflaton and the radiation will cause random kicks to the inflaton field, ϕ.

As in Brownian motion there is a net drag/frictional effect, summarised by equation (4.42). On top of the net friction term, there are random kicks that average to zero. We can find the random kicks by subtracting the frictional effect from the interaction term

$$\Upsilon(\vec{x}, t) \equiv \mathcal{L}'_{\rm int} - \langle \mathcal{L}'_{\rm int}(\phi) \rangle. \tag{4.43}$$

It is now possible to calculate the amplitude of the fluctuations $\langle \Upsilon(x)\Upsilon(y) \rangle = A\delta^{(4)}(x-y)$ by calculating a thermal expectation value, or by using the fluctuation-dissipation theorem. The

thermal expectation value is calculated in detail in appendix C, and gives the result

$$A = \frac{2c_1 T^4}{f^2}(16\alpha_1^2 + 25\alpha_2^2) = 2\gamma T \frac{(16\alpha_1^2 + 25\alpha_2^2)(\alpha_1^2 + \alpha_2^2)}{\alpha_1^2 \alpha_2^2}. \tag{4.44}$$

Notice that this is modified relative to $2\gamma T$ by a dimensionless (but rather large) model-dependent coefficient.

We will derive this result from the fluctuation-dissipation theorem with a chemical potential, but we should first make sense of it from a physical perspective. Consider taking the limit $\alpha_2 \to 0$ in equation (4.44). In this limit we get back to the pure chemical potential model. In this limit the friction, $\gamma \to 0$, but the fluctuations are still non-zero (and in fact they are unmodified by the chemical potential at leading order). The reason for this is because $\gamma \to 0$ is a consequence of the balancing of the two rates in (4.12). Whilst the two rates are balanced there are still many scattering processes occurring per unit time, and therefore the fluctuations are still large! It is easy to see from this special case that $A/\gamma \to \infty$ because the dissipation is cancelled entirely by the chemical potential.

Based on the discussion above, it is clear that the coefficients, γ, A can be phrased in terms of the thermal expectation value of the operators in equation (4.20)

$$\mathcal{L}_{\text{int}} \supset \mathcal{O}_1(x) + \mathcal{O}_2(x) + h.c. \qquad \mathcal{L}'_{\text{int}} = \frac{4i}{f}\mathcal{O}_1(x) + \frac{5i}{f}\mathcal{O}_2(x) + h.c. \tag{4.45}$$

where $\mathcal{O}_1 = -\frac{\alpha_1}{4!}\chi^4 e^{4i\dot{\varphi}/f}$, etc. We can write the thermal friction term

$$\begin{aligned}\gamma\dot{\varphi} \equiv \langle \mathcal{L}'_{\rm int}\rangle &= \frac{4}{f}\int_{-\infty}^{t_x} dt_y \int d^3y \langle[\mathcal{O}_1(x), \mathcal{O}_1^\dagger(y)]\rangle \\ &\quad + \frac{5}{f}\int_{-\infty}^{t_x} dt_y \int d^3y \langle[\mathcal{O}_2(x), \mathcal{O}_2^\dagger(y)]\rangle + h.c. \end{aligned} \tag{4.46}$$

where we have applied the in-in master formula (1.52), which also applies for thermal expectation values, see appendix C for details. The only subtlety here is the definition of the thermal expectation value. Due to the presence of a chemical potential, $\mu_{\rm dist} = \mu$, in the distribution functions, the thermal expectation values should be calculated in the grand canonical ensemble

$$\langle \mathcal{O}(x)\rangle \equiv \mathrm{Tr}\left[\mathcal{O}(x)e^{-\beta(H-\mu Q)}\right], \tag{4.47}$$

where $Q = J^0$ is the χ-number operator. An operator with a definite charge under Q satisfies $[\mathcal{O}(x), Q] = q\mathcal{O}(x)$. This definition implies $\mathcal{O}_1$ and $\mathcal{O}_2$ have charge $q = 4$. The fluctuations are instead given by

$$A \equiv \int d^4y \langle \mathcal{L}'_{\rm int}(x)\mathcal{L}'_{\rm int}(y)\rangle = \frac{16}{f^2}\int d^4y \langle \mathcal{O}_1(x)\mathcal{O}_1^\dagger(y)\rangle + \frac{25}{f^2}\int d^4y \langle \mathcal{O}_2(x)\mathcal{O}_2^\dagger(y)\rangle + h.c. \tag{4.48}$$

Both quantities that we are interested in calculating in warm inflation can be related to correlation functions of operators in the grand canonical ensemble.

In fact, if we define $\mathcal{O}(x) = \chi^4(x)$, then both quantities can be neatly expressed as

$$\gamma\dot{\phi} \equiv \frac{4}{f}\frac{\alpha_1^2}{(4!)^2}\left(\rho_{\mathcal{O}}(k_1) - \rho_{\mathcal{O}^\dagger}(-k_1)\right) + \frac{5}{f}\frac{\alpha_2^2}{(4!)^2}\left(\rho_{\mathcal{O}}(k_2) - \rho_{\mathcal{O}^\dagger}(-k_2)\right) \tag{4.49}$$

$$A \equiv \frac{16\alpha_1^2}{(4!)^2 f^2}\left(\Pi^<_{\mathcal{O}}(k_1) + \Pi^<_{\mathcal{O}^\dagger}(-k_1)\right) + \frac{25\alpha_2^2}{(4!)^2 f^2}\left(\Pi^<_{\mathcal{O}}(k_2) + \Pi^<_{\mathcal{O}^\dagger}(-k_2)\right), \tag{4.50}$$

where $k_1 = (4\dot{\varphi}/f, \vec{0})$, and $k_2 = (5\dot{\varphi}/f, \vec{0})$. The various functions above are the conventionally defined correlation functions [98]

$$\rho_{\mathcal{O}}(k) \equiv \int d^4x e^{ik\cdot(x-y)}\frac{1}{2}\langle[\mathcal{O}(x), \mathcal{O}(y)]\rangle, \tag{4.51}$$

$$\Pi^<_{\mathcal{O}}(x) \equiv \int d^4x e^{ik\cdot(x-y)}\langle\mathcal{O}^\dagger(x)\mathcal{O}(y)\rangle. \tag{4.52}$$

The fluctuation-dissipation theorem, derived in detail in appendix C, is really a relationship between these two types of correlation function

$$\rho_{\mathcal{O}}(k) = \frac{1}{2}\left(e^{\beta(k^0-\mu q)} - 1\right)\Pi^<_{\mathcal{O}}(k). \tag{4.53}$$

In the application of the fluctuation-dissipation theorem to warm inflation we should make sure to recall that we are working to leading order in $\dot{\varphi}/fT$. We Taylor expand the exponential, and set $k \to 0$ in the second expectation value

$$\rho_{\mathcal{O}}(k) \approx \frac{1}{2}\beta(k^0 - \mu q)\Pi^<_{\mathcal{O}}(0). \tag{4.54}$$

If we apply the fluctuation-dissipation theorem to equations (4.49), and (4.50), we arrive at the result (4.44) which was obtained by a direct calculation of the two correlation functions.

This was a good sanity check of our results, and it also teaches us come crucial information about how the fluctuation-dissipation theorem should and should not be used in the context of warm inflation. Firstly, in the literature it is usually quoted that $A = 2\gamma T$. This would be the result if $\mu q = 0$ in equation (4.53). In models that do not have pseudo-scalar couplings, where $\mu = 0$, the usual result therefore applies. However, it is also clear that even in this case $A = 2\gamma T$ is only correct at leading order in a Taylor expansion in $\dot{\varphi}$. Secondly, even in the case $\mu \neq 0$, if μ is the chemical potential for a conserved charge, then it is necessarily the case that $q = 0$, because the operators, $\mathcal{O}(x)$, involved in the calculation of either γ, or A must not break charge conservation as they appear in the Hamiltonian.

The only reason we deviate from the usual expectation is because we are in the somewhat bizarre situation in which the inflaton background is acting as an applied chemical potential for a non-conserved charge. Whilst it seems a little bizarre we should stress that this is in no way a peculiarity of our specific model of warm inflation. In any model with pseudo-scalar couplings for the inflaton, the slow-roll background is a spurion that spontaneously breaks CPT. It will therefore always act as an applied chemical potential, with $\mu_{\text{applied}} \propto \dot{\varphi}$. We therefore expect a similarly large model-dependent deviation of the fluctuation dissipation theorem for any model of warm inflation with pseudo-scalar couplings, such as in [98–102]. The implications of this conclusion for other models in the literature are discussed at length in appendix C.

4.6 The Warm Inflation Power Spectrum

In this section we calculate the prediction for the power spectrum in warm inflation by solving a slightly more complicated version of the Brownian motion equation of motion (4.35).

The power spectrum in warm inflation has been calculated many times with differing results [96, 97, 103–105]. Our results follow closely the treatment in [104].

Firstly, as in the case of cold inflation there is a gauge in which we can drop the metric perturbations as being Planck suppressed, and write the gauge invariant scalar curvature perturbation

$$\zeta_k = -H\frac{\xi_k}{\dot{\varphi}}. \tag{4.55}$$

To compare with the data we need to calculate $\langle \zeta_k \zeta_{-k} \rangle'$.

We redefine the fluctuation of the inflaton $\phi = \varphi + \dot{\varphi}\,\Phi$. Besides the fluctuations of the inflaton we also include fluctuations of the radiation energy density, $T^0_{\chi,0} = -\rho_r(1+\mathcal{E})$, and a perturbation to the scalar component of the momentum density in the radiation sector $\partial_i T^i_{\chi,0} = -\frac{4\rho_r}{3a^2}\nabla^2\Psi$. The equations to solve are:

- Einstein's equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{4.56}$$

  from which one can derive the background Friedmann equation. The perturbed equations only say we can drop the metric perturbation as being suppressed by the Planck scale.

- The inflaton equation of motion

$$\ddot{\phi} + 3H\dot{\phi} - \frac{\nabla^2}{a^2}\phi + V'(\phi) = -\gamma u^\mu \partial_\mu \phi + \Upsilon, \tag{4.57}$$

  where u^μ specifies the rest frame of the thermal bath relative to the inflaton. The background equation here is equation (4.2).

- The conservation of the total stress tensor

$$\nabla_\mu T^\mu_{\ \nu} = 0; \qquad T^\mu_{\ \nu} = T^\mu_{\phi,\nu} + T^\mu_{\chi,\nu}. \tag{4.58}$$

From this and equation (4.57), one can derive the extent to which the radiation stress energy tensor is not conserved due to interactions

$$\nabla_\mu T^\mu_{\chi,\nu} = -\partial_\nu\phi\left(\gamma u^\mu \partial_\mu\phi + \Upsilon\right). \tag{4.59}$$

The background equation here leads directly to the Boltzmann equation (4.3). The $\nu = 0, i$ equations will give two independent equations for the perturbations. Combined with equation (4.57), this gives three equations for three unknowns.

If we expand all three independent equations to leading order in perturbations, assuming Υ is itself first order in perturbations, then we arrive at the following coupled second order differential equations to solve for the perturbations:

$$\Phi'' - \left(\frac{2}{\eta} + \frac{\gamma}{H\eta}\right)\Phi' + k^2\Phi = -\frac{3\gamma\mathcal{E}}{4H^2\eta^2} + \frac{\Upsilon}{H^2\eta^2\dot{\varphi}}, \tag{4.60}$$

$$\mathcal{E}' - \frac{1}{\eta}\mathcal{E} + \frac{4k^2H\eta\Psi}{3} - 8H\Phi' = \frac{4}{\gamma\eta\dot{\varphi}}\Upsilon, \tag{4.61}$$

$$\Psi' - \frac{3}{\eta}\Psi - \frac{1}{4H\eta}\mathcal{E} - \frac{3}{\eta}\Phi = 0. \tag{4.62}$$

To find the primordial power spectrum we numerically solve for the retarded Green's function, assuming the noise term is a source. We then use the two-point function of the noise

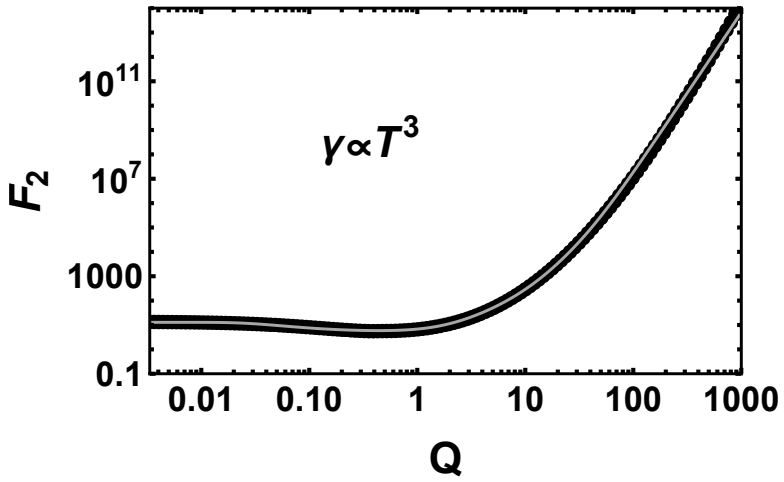


Figure 4.2: taken from [106]/appendix C. Numerical fit to $F_2(Q)$, the function that appears in the primordial power spectrum. In the strong regime the function grows rapidly, but in the weak regime it remains an order 1 number.

term

$$\langle \Upsilon_k(\eta)\Upsilon_{k'}(\eta')\rangle = AH^4\eta^4\delta(\eta-\eta')(2\pi)^3\delta^{(3)}(\vec{k}+\vec{k}'), \tag{4.63}$$

which is equivalent to the two-point function before after changing variables to comoving distance, Fourier transforming, and switching to conformal time. The power spectrum is then given by

$$\langle \zeta_k\zeta_{k'}\rangle = \frac{AH^2}{\dot{\varphi}^2}F_2(Q)\frac{(2\pi)^3\delta^3(k+k')}{k^3}, \tag{4.64}$$

where the function $F_2(Q)$ is found by numerically integrating the Green's function for equation (4.60)

$$\begin{aligned} F_2(Q) &\equiv k^3\int_{-\infty}^{0} d\eta\,\eta^4 G^2_{\Phi_k}(0,\eta) \\ &= 5+7.9\tanh\left(\frac{1}{30Q}\right)+1.5Q^2+7.8\times10^{-3}Q^{4.2}+1.0\times10^{-6}Q^{6.6}, \end{aligned} \tag{4.65}$$

where the final line is the result of a numerical fit, shown in figure 4.2.

Now that we have the expression for the size of the power spectrum in warm inflation, we should compare it to the power spectrum of cold inflation (1.33). Up to order one numbers we can say that parametrically

$$\frac{H^4}{\dot{\varphi}^2} \to \frac{T^4}{f^2}\frac{H^2}{\dot{\varphi}^2}. \tag{4.66}$$

Assuming T and f are not too far apart then we would suspect that thermal fluctuations dominate over their quantum mechanical counterpart as long as $T \gg H$.

4.7 A Working Warm Inflation Data Point

The goal of this project was to come up with a model of warm inflation that is easy to work with, figure out the main subtleties, and show that warm inflation is viable consistent with current constraints. We therefore make no attempt to scan through the entire viable parameter space, and simply seek one data point that works. Firstly, we pick the slow-roll potential

$$V(\phi) = \frac{\lambda}{4}\phi^4. \tag{4.67}$$

To fit the current data we must match the amplitude of the power spectrum

$$A_s = \frac{c_1 T^4}{\pi^2 f^2}(16\alpha_1^2 + 25\alpha_2^2)\frac{H^2}{\dot{\varphi}^2}F_2(Q) = 2.09 \times 10^{-9}. \tag{4.68}$$

In addition to the amplitude of the power spectrum we must fit the spectral tilt, $n_s - 1$. We can calculate the spectral tilt in warm inflation, for a derivation go to appendix C

$$n_s - 1 = 4\frac{2-Q}{1+7Q}\epsilon_V - 6\frac{1+Q}{1+7Q}\eta_V + \frac{d\log F_2}{d\log Q}\frac{1+Q}{1+7Q}(10\epsilon_V - 6\eta_V) = -0.0335. \tag{4.69}$$

We must find a data point that fits these observables, and is consistent with the myriad of assumptions we have made throughout this chapter. There are a few assumptions we have yet to mention. Firstly, we must make sure $\mu_{\rm dist} < m_\chi$, to avoid the formation of a BEC. Secondly, the thermal bath must not excite modes that are out of EFT control $4\pi f > 3T$. Warm inflation can be realised consistent with all the assumptions we have made with the following choice of parameters:

$$\frac{\alpha_1}{16\pi^2} = 0.443, \qquad \frac{\alpha_2}{16\pi^2} = 0.450, \qquad \lambda = 3.98 \times 10^{-18}, \quad f = 9.56 \times 10^{-4} M_{\rm pl}. \tag{4.70}$$

To ensure the complex scalar acts like radiation, and yet the chemical potential is less than the mass we chose $m_\chi/T = 0.55$. The output parameters of warm inflation are

$$\varphi = 17.07 M_{\rm pl}, \quad Q = 0.027, \quad H = 1.68 \times 10^{-7} M_{\rm pl}, \quad T = 8.21 \times 10^{-5} M_{\rm pl}. \tag{4.71}$$

The final constraint from data is an upper bound on the tensor to scalar ratio. The prediction for the tensor perturbations in warm inflation is unmodified from cold inflation, so the tensor to scalar ratio is always highly suppressed

$$r = \left(\frac{\dot\phi}{M_{\rm pl}}\right)^2 \frac{2f^2}{c_1(16\alpha_1^2 + 25\alpha_2^2)T^4 F_2(Q)} = 1.36 \times 10^{-6}, \tag{4.72}$$

which is well below current bounds. In fact, suppressing the tensor-to-scalar ratio is often used to motivate considering warm inflation.

In this chapter we have introduced the idea of warm inflation as an alternative origin

of structure in our universe. We have understood a simple, perturbative, and fully calculable model of warm inflation. In the context of this model we have shown that warm inflation is consistent with our observations of the CMB power spectrum. We are left with the question, is it possible to determine experimentally whether inflation was warm or cold? It turns out that warm inflation may imply a characteristic shape function for NGs, which may be a smoking gun [104]. However, we have left the calculation of NGs for future work.

Chapter 5

Conclusion

In this thesis we have considered several new mechanisms in inflationary cosmology. In the first three quarters of the thesis we have studied cosmological collider physics, the idea that one can detect exotic new particles in precision cosmological data. In chapter 1 we introduced the basic ideas behind inflation, cosmological collider physics, calculated the signal in the simplest scenario, and analysed its main features.

In chapter 2 we were motivated by the current upper bound on the Hubble scale during inflation to hunt for grand unification. We introduced a chemical potential mechanism for spin-1 particles, boosting the prospects of searching for heavy gauge bosons in orbifold GUTs, and opening up the possibility of detecting supersymmetric grand unification.

In chapter 3 we noted that searching for the signal due to the tree level exchange of a massive particle imposes the restriction to search for SM neutral new physics. To consider the prospects of searching for charged particles we identified a model in which a chemical potential is given to a pair of oppositely charged scalars. The leading contribution is due to a calculable

loop diagram. We calculated the parametric dependences of the signal, and found that an experimental detection is challenging, but not impossible, within the region of parameter space in which we have analytic control. We then applied this mechanism to supersymmetric orbifold GUTs, and noticed that relaxing the requirement to have a neutral complex field expands the possible GUT gauge groups one can detect. We also discovered that the multiplicity of fields which are charged under a non-Abelian symmetry group boosts the detection prospects.

Finally in chapter 4 we considered a different scenario called warm inflation. In warm inflation the inflaton interacts with a thermal bath, and the random kicks it acquires due to its interaction provide an alternative origin of structure. In this scenario the observed Gaussian nature of the CMB data is a consequence of the central limit theorem rather than the inflaton effectively acting a free field. We identified a simple model of warm inflation with pseudo-scalar couplings, and checked that warm inflation is consistent with the current CMB data. We also found, and understood from an intuitive physical perspective, disagreements with some of the general claims made in the literature.

We have seen over the course of the thesis that both cold and warm inflation are consistent with current observations of the CMB power spectrum. In the case of cold inflation we have seen that non-Gaussianities probe the small interactions of the inflaton during inflation, and it is in NGs that we may discover exotic new particles produced during inflation. In the case of warm inflation we have stated that NGs are expected to have a characteristic shape that could be a smoking gun, but we have left the verification of this statement to future work. The unifying theme of the thesis is therefore that to understand more about what happened during inflation we will need a detailed understanding of non-Gaussianities, an understanding we can hope to obtain in future large scale structure surveys, and more futuristically in 21-cm

tomography experiments.

Appendix A: Grand Unification at the Cosmological Collider with Chemical Potential

Arushi Bodas[1,2], Edward Broadberry[3], Raman Sundrum[3]

[1] Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA

[2]Particle Theory Department, Fermilab, Batavia, Illinois 60510, USA

[3]Maryland Center for Fundamental Physics, Department of Physics,

University of Maryland, College Park, MD 20742, USA

Published in the Journal of High Energy Physics 01 (2025) 115,

DOI: https://link.springer.com/article/10.1007/JHEP01(2025)115

Abstract:

We introduce a tree-level chemical potential mechanism for spin-1 particles within cosmological collider physics, allowing them to be detected in primordial non-Gaussianities for masses above the inflationary Hubble scale. We apply this mechanism to orbifold grand unification and the massive unification partners of the standard model gauge bosons. Our mechanism requires at least a pair of massive vector fields which are singlets of the standard model, a condition which is satisfied in the classic "trinification" scenario. Assuming that the gauge hierarchy problem is solved by supersymmetry, gauge coupling running points to unification partners at $\sim 10^{15}$ GeV. We show that, within high-scale inflation, chemical potential enhancement can lead to observably strong signals for trinification partners in future cosmological surveys.

A.1 Introduction

The emerging field of cosmological collider physics [1,31–33,107–109] shows how primordial correlators can probe particle physics at inflationary energy scales, which may be many orders of magnitude above that of any terrestrial collider. Current data places an upper bound on the Hubble scale during inflation [17]

$$H \lesssim 5 \times 10^{13} \text{ GeV}, \tag{A.1}$$

informing us of the potential reach of the cosmological collider. The core universal mechanism involves particle production due to spacetime expansion for all species with masses $m \sim O(H)$. If such heavy particles decay into inflatons their propagation will be imprinted on primordial non gaussianities (NGs), with potentially observable signals in CMB data (see [21] for recent data analysis), large scale structure surveys [26–29] (see [20] for recent data analysis), and high-redshift 21-cm tomography [110, 111].

The proto-typical signal for on-shell exchange of a heavy particle is given in the squeezed limit $k_1 \sim k_2 \gg k_3$ by [31, 108]

$$\frac{\langle \mathcal{R}_{\vec{k}_1} \mathcal{R}_{\vec{k}_2} \mathcal{R}_{\vec{k}_3} \rangle'}{\langle \mathcal{R}_{\vec{k}_1} \mathcal{R}_{-\vec{k}_1} \rangle' \langle \mathcal{R}_{\vec{k}_3} \mathcal{R}_{-\vec{k}_3} \rangle'} = e^{-\pi\mu_0} \left(f(\mu_0) \left(\frac{k_1}{k_3} \right)^{-3/2+i\mu_0} + c.c. \right), \tag{A.2}$$

where $\mu_0 = \sqrt{m^2/H^2 - 9/4}$, and $\mathcal{R}$ is the comoving curvature perturbation. The key observable feature is the non-analytic dependence on comoving momentum, k. However, inspection of eq. (A.2) shows the mass must be in a narrow window around $m \sim H$ due to an exponential suppression for large masses, and the non-analytic dependence being unobservable for small masses. Intuitively, the exponential suppression for large masses can be thought of as a Boltz-

mann suppression, where the exponential factor is the square root of a Boltzmann probability for particle excitation with respect to a Hawking temperature of dS-space $T_{\rm H} = H/2\pi$ [112].

Since observability requires a new particle's mass to lie in a narrow range, it is important to consider whether new particles are expected at this scale. A highly motivated class of models predicting new particles at around this energy are grand unified theories (GUTs) [40,41] (see [42] for a review). The main piece of quantitative evidence for GUTs is the approximate unification of the SM gauge couplings if their RG flow is extrapolated to energies of $M_{\rm GUT} \sim 10^{14}$ GeV. The case for detecting the "Higgs" that breaks the GUT gauge group down to the SM in NGs was investigated in [54]. However, the minimal GUT scenarios predict proton decay at a rate ruled out by experiments [44, 45]. Proton decay can be naturally avoided in higher-dimensional "orbifold GUTs" where the GUT symmetry is broken to the standard model by boundary conditions on a $(3+1)$-dimensional boundary [46–48]. Ref. [49] investigated the cosmological collider signals of such models. It was found that the signals for the heavy gauge bosons were highly suppressed, so the largest contribution to NGs was a Planck suppressed signal from the Kaluza Kelin (KK) graviton. Remarkably the KK graviton had small but observable signals. However, this required an even narrower mass range for these spin-2 particles than in the spin-0 case of eq. (A.2).

Minimal GUTs have a few theoretical drawbacks. Firstly, the quantitative evidence in the form of the coupling unification is not so precise. Secondly, the value of the electroweak symmetry breaking (EWSB) scale $v \ll M_{\rm GUT}$ requires fine tuning, that is the gauge hierarchy problem. Both of these problems are solved in supersymmetric (SUSY) GUTs where the coupling unification is more precisely achieved, and a small EWSB scale is radiatively stable. However, bounds on proton decay require the colored GUT partners of the Higgs doublet to be at the

GUT scale. This is called the doublet-triplet splitting problem and corresponds to a significant tree-level fine tuning. On the other hand, orbifold SUSY GUTs suppress proton decay as before, and also solve the doublet triplet splitting by using the boundary conditions to remove the light modes of the colored Higgs triplet whilst retaining the light modes of the SM Higgs. The "cost" of standard supersymmetric grand unification from the perspective of cosmological collider physics is a higher unification scale $M_{\rm GUT} \sim 10^{16}$ GeV, placing it out of reach for the cosmological collider. In orbifold SUSY GUTs however, the KK excited GUT partners can appear at the lower compactification scale $M_C \sim 10^{15}$ GeV [48], as illustrated in figure A.1. However, naively this is still out of reach given eq. (A.1) and the Boltzmann suppression in eq. (A.2).

Nevertheless, there is a "chemical potential" extension of cosmological collider physics involving direct couplings between the heavy particles and the inflaton, in which Boltzmann suppression above H can be evaded [34, 35, 37, 38, 113, 114]. In this paper we will develop a variant of the chemical potential mechanism that allows spin-1 GUT excitations to be seen.

As in thermodynamics, the idea of a chemical potential is that it can overcome Boltzmann suppression. In the inflationary context the mechanism takes advantage of a larger scale than H during inflation, namely the kinetic energy of the inflaton [17]

$$\dot{\phi}_0 \simeq (60H)^2, \tag{A.3}$$

where $\phi_0(t)$ is the inflationary slow-roll background. There exist a class of dimension-5 "chemical potential" couplings schematically of the form

$$\mathcal{L} \sim \frac{i}{\Lambda}\nabla_\mu \phi J^\mu, \tag{A.4}$$

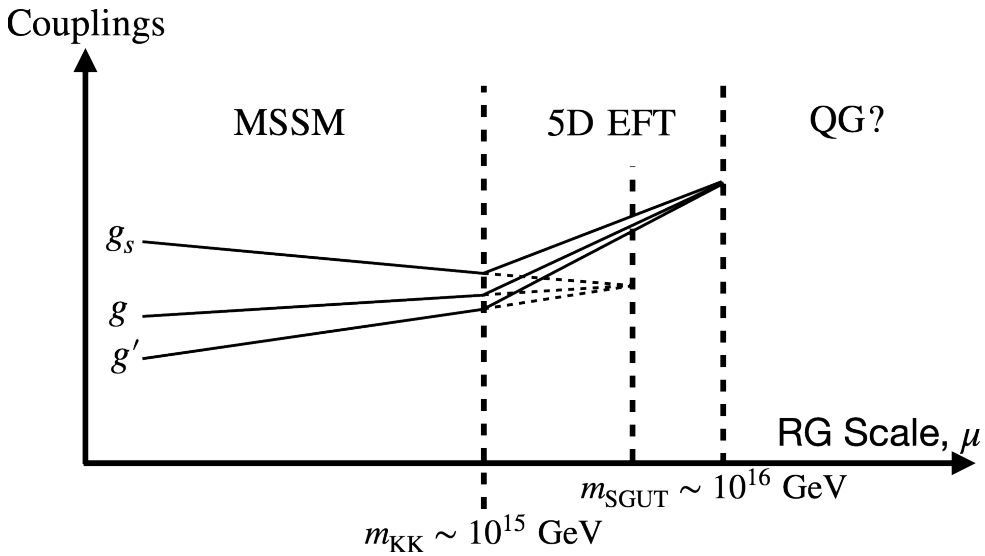


Figure A.1: Adapted from ref. [48]. In orbifold GUTs the gauge couplings need not exactly unify before the extra dimension becomes important. Above this scale the running is modified, and it has been speculated that they unify with the gravitational coupling into a strongly coupled string theory, or some other UV completion of quantum gravity.

where $\Lambda \geq 60H$ in order for the inflationary EFT to be in theoretical control, and J^μ is a current which is quadratic in the heavy field. In the inflationary background, eq. (A.4) reduces to a chemical potential coupling $\mathcal{L} \sim i\lambda J^0$. Such couplings allow for unsuppressed particle production provided $m \leq \lambda$, where

$$\lambda \equiv \dot{\phi}_0/\Lambda \tag{A.5}$$

is the chemical potential. Note, the chemical potential can be as large as the kinetic energy scale

$$\lambda \leq 60H \lesssim 3 \times 10^{15} \text{ GeV}, \tag{A.6}$$

which is above the compactification scale of orbifold SUSY GUTs. The chemical potential mechanism both increases the reach, and widens the mass window probed by cosmological collider physics, thereby allowing us to plausibly directly see GUT particles!

Our mechanism of chemical potential applies to a massive complex vector field, which for a simple tree-level process must be a SM singlet in order to decay to singlet inflatons. This precludes its application to $SU(5)$ and $SO(10)$ GUTs for which there is at most a single SM singlet real vector field among the GUT partners. However, trinification, another classic unification scenario with the gauge group $G = SU(3)_C \times SU(3)_L \times SU(3)_R/\mathbb{Z}_3$ [50], provides a suitable target as it contains several SM singlet GUT partners. The realization of the chemical potential in an orbifold trinification scenario is the focus of this paper, while the extension to fully supersymmetric dynamics is left for future work. We expect the central effect of supersymmetry will be to replace the SM gauge-coupling running by the MSSM running, thereby raising the estimate of the compactification scale to $M_C \sim 10^{15}$ GeV. Here we will simply assume $M_C \sim 10^{15}$ GeV as the target for our chemical potential mechanism.

There has been an earlier implementation of chemical potential for a single real spin-1 field, in which the enhanced/suppressed "charges" correspond to the two helicities (that is, the rough analog of our complex field is made from the two real polarizations) [34, 38]. In this case the gauge-invariant form of the chemical potential interaction displays a tachyonic instability that is cut off by the spin-1 (gauge-noninvariant) mass, requiring some tuning in parameter space to obtain sizeable NG within theoretical control. By contrast our chemical potential interaction is not gauge-invariant in form (gauge-invariance is not mandatory because we are dealing with massive spin-1 particles) and avoids such tuning. Also, the gauge-invariant chemical potential gives rise to NG only at loop-level, whereas our implementation gives rise to NG at tree level with a relatively straightforward analysis. However, since the gauge-invariant realization requires only one real spin-1 field, it might be more readily applicable to $SU(5)$ and $SO(10)$ GUTs. We leave this to future work.

The paper is organized as follows. In section A.2 we introduce our conventions and background material. In particular we summarize the spin-0 chemical potential which we generalise to spin-1. In section A.3 we introduce the model for spin-1 chemical potential. A Stueckelberg analysis is then performed to ensure that the massive spin-1 EFT is within theoretical control, and the relevant interactions for our NG calculation are isolated. Section A.4 introduces the stationary phase approximation for calculating the NG. The central structure of the calculation is presented, with the full details left for appendix A.8.3. We then discuss the apparent late-time divergences due to non-derivative inflaton couplings, explaining why they should cancel based on the shift symmetry under which the inflaton and the heavy field transform. A fuller explanation of this cancellation is presented in appendix A.8.4. We also calculate the corrections to the power spectrum. Requiring this to be subdominant places a bound on the strength of our NG signal. In section A.5 we discuss observational prospects, a method to determine both the mass of the spin-1 particle and the chemical potential, and the possibility of inferring the spin. In section A.6 we study trinification, and show that it contains a candidate complex gauge boson that is uncharged under the SM. We show how to implement the chemical potential coupling, but with a fully supersymmetric implementation left for future work. In section A.7 we end with a summary and a discussion of possible avenues for future research.

A.2 Preliminaries

In this section we establish notational conventions, introduce the in-in formalism for computing cosmological correlators, and give a summary of relevant results from the previous paper on chemical potential [35], which we are generalizing to the spin-1 case.

A.2.1 Conventions

The mostly plus metric signature $(-,+,+,+)$ is used throughout, and the Lagrangian densities are defined with the volume form factored out to facilitate easy integration by parts. During slow-roll inflation, the approximate constancy of the Hubble parameter (H) justifies approximating inflationary spacetime as de-Sitter (dS) space:

$$\begin{aligned} ds^2 &= -dt^2 + a^2(t)d\vec{x}^2 \quad \text{with } a(t) = e^{Ht}, \\ &= \frac{1}{(H\eta)^2}\left(-d\eta^2 + d\vec{x}^2\right). \end{aligned} \tag{A.7}$$

Here, $\eta = \int dt/a(t)$ is the conformal time with $a(t)$ being the scale factor. In dS spacetime, $\eta = -1/(aH)$. For the sake of convenience, we will set units such that the approximately constant Hubble parameter $H \equiv 1$, restoring H explicitly as needed by dimensional analysis.

The inflationary spacetime is most simply realized by the dynamics of a scalar "inflaton" field, ϕ, when it dominates the energy density of the universe. This field can be conveniently separated into a slowly-varying homogeneous background, $\phi_0(t)$, which drives inflation and fluctuations, $\delta\phi(\vec{x},t)$, which seed the inhomogeneities of the later universe after inflation. The fluctuations of the inflaton and the scalar perturbations of the metric, however, are coordinate dependent, with the gauge-invariant combination being the comoving curvature perturbation, $\mathcal{R}$ [115, 116]. Since spatial translation is still a good symmetry of the background inflationary dynamics, the correlations of curvature perturbations have the following form,

$$\langle \mathcal{R}_{\vec{k}_1} \cdots \mathcal{R}_{\vec{k}_n} \rangle = (2\pi)^3 \delta(\vec{k}_1 + \cdots + \vec{k}_n) \langle \mathcal{R}_{\vec{k}_1} \cdots \mathcal{R}_{\vec{k}_n} \rangle', \tag{A.8}$$

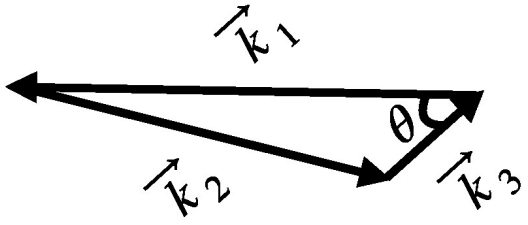


Figure A.2: The shape of the squeezed limit triangle ($k_1 \sim k_2 \gg k_3$) is specified by the degree of squeezing, $p \equiv k_1/k_3$, and the angle between the hard and soft momenta, θ.

where the prime indicates that the comoving momentum-conserving delta function has been factored out.

The CMB data favors a nearly scale-invariant primordial power spectrum [17]

$$\mathcal{P}_{\mathcal{R}}(k) = \langle \mathcal{R}_{\vec{k}} \mathcal{R}_{-\vec{k}} \rangle' = \frac{1}{(\dot{\phi}_0)^2} \frac{1}{2k^3} (k/k_*)^{n_s - 1}, \tag{A.9}$$

where the scalar spectral index $n_s \approx 0.96$, and the reference comoving momentum scale $k_* \approx 0.05\text{Mpc}^{-1}$. Beyond the power spectrum there will be non-Gaussian (NG) higher-point correlations in $\mathcal{R}$.

The focus of this paper is the 3-point correlation function, also called the primordial bispectrum, $\langle \mathcal{R}_{\vec{k}_1} \mathcal{R}_{\vec{k}_2} \mathcal{R}_{\vec{k}_3} \rangle'$. It is conventional to characterise NGs relative to the size of the power spectrum [117]

$$F(k_1, k_2, k_3) = \frac{5}{6} \frac{\langle \mathcal{R}_{\vec{k}_1} \mathcal{R}_{\vec{k}_2} \mathcal{R}_{\vec{k}_3} \rangle'}{\mathcal{P}_{\mathcal{R}}(k_1)\mathcal{P}_{\mathcal{R}}(k_2) + \mathcal{P}_{\mathcal{R}}(k_1)\mathcal{P}_{\mathcal{R}}(k_3) + \mathcal{P}_{\mathcal{R}}(k_2)\mathcal{P}_{\mathcal{R}}(k_3)}. \tag{A.10}$$

Momentum conservation implies that the three comoving momentum vectors can be arranged to form a triangle. There are two limits of the above expression that are of interest to us. Firstly, the size of NG is typically quoted in terms of a quantity, f_{NL}, defined by the equilateral triangle

configuration

$$f_{\rm NL} \equiv F(k,k,k). \tag{A.11}$$

Secondly, the characteristic cosmological collider signal appears in the squeezed limit $k_1 \sim k_2 \gg k_3$, illustrated in figure A.2.

As of now, primordial NG has not been detected in the CMB, which roughly puts the bounds on $f_{\rm NL} \lesssim \mathcal{O}(5-10)$ [117] (the exact constraint is shape-dependent.) However, future datasets are expected to reach higher precision for NGs compared to the CMB. For example, large-scale structure surveys (LSS) are projected to be sensitive to $f_{\rm NL} \sim \mathcal{O}(1)$ [26], while a more futuristic 21-cm tomography could probe NG as low as $f_{\rm NL} \sim \mathcal{O}(10^{-2})$ [110, 111]. These datasets would be sensitive to our benchmark signals in section A.4.

Calculation of the bispectrum is performed in the in-in formalism. The expectation value of a Heisenberg picture operator at a time t is given by [24, 118]

$$\langle \mathcal{O}_H(t)\rangle = \langle 0|\bar{T}\left\{e^{i\int_{-\infty(1+i\epsilon)}^{t} H_I(t')dt'}\right\}\mathcal{O}_I(t)T\left\{e^{-i\int_{-\infty(1-i\epsilon)}^{t} H_I(t')dt'}\right\}|0\rangle, \tag{A.12}$$

where the subscript I indicates an interaction picture operator and $H_I(t)$ is the interaction Hamiltonian. The state $|0\rangle$ is the Bunch-Davies "vacuum" corresponding to the free Minkowski vacuum at asymptotically early times (when each k mode has physical wavenumber $|k\eta| \gg H$), while the $i\epsilon$ prescription projects onto the interacting de-Sitter state as time evolves. The primordial contribution to an expectation value is calculated at the end of inflation (on the reheating surface, for example). Provided the primordial NGs are insensitive to the details of reheating (which is a valid assumption since all the comoving modes that we are interested in

exit the horizon before reheating), one can calculate expectation values in exact dS space, with the end of inflation approximated as $t \to \infty$, or $\eta \to 0$. In this limit the curvature perturbation is most easily calculated in the spatially flat gauge

$$\mathcal{R} = \frac{\delta\phi}{\dot{\phi}_0}, \tag{A.13}$$

where the time derivative of the inflaton background $\dot{\phi}_0 \approx$ constant during inflation. In the squeezed limit we can simplify eq. (A.10), and define $F_{\rm sq}$ as

$$F_{\rm sq} = \frac{5}{12}\dot{\phi}_0 \frac{\langle \delta\phi_{\vec{k}_1}\delta\phi_{\vec{k}_2}\delta\phi_{\vec{k}_3}\rangle'}{\langle \delta\phi_{\vec{k}_1}\delta\phi_{-\vec{k}_1}\rangle' \langle \delta\phi_{\vec{k}_3}\delta\phi_{-\vec{k}_3}\rangle'}\,, \tag{A.14}$$

where eq. (A.13) was used to write the NG in terms of inflaton fluctuations. Henceforth, we will ignore the subdominant slow-roll corrections and work to leading order in the slow-roll background.

A.2.2 Recap: chemical potential for spin-0 particles

The goal of this paper is to extend to spin-1 particles the mechanism of chemical potential that was discussed for spin-0 fields in ref. [35]. We therefore review some of the important lessons from ref. [35]. In the minimal cosmological collider, the exponential suppression $\sim e^{-\pi m/H}$ for fields with $m > H$ is a consequence of the fact that the inflating background typically provides energies of order H, or intuitively a Hawking temperature $T_{\rm H} = H/(2\pi)$. One way to overcome this "Boltzmann" suppression is to incorporate a "chemical potential" that is larger than $T_{\rm H}$, and can allow excitations of particles that are heavier than H. Such a chemical potential can be

provided by the kinetic energy of the rolling inflaton. Coupling the inflaton derivatively to the current of the heavy field can produce a chemical potential-like operator. For a complex scalar χ of mass m, the Lagrangian looks like

$$\begin{aligned}\mathcal{L}_\chi \supset & \left[\left(\nabla_\mu + i\frac{\nabla_\mu\phi}{\Lambda}\right)\chi\right]\left[\left(\nabla^\mu - i\frac{\nabla^\mu\phi}{\Lambda}\right)\chi^\dagger\right] - M^2|\chi|^2 + \alpha\chi + h.c \\ = & -|\nabla_\mu\chi|^2 - M^2|\chi|^2 + \underbrace{i\frac{\nabla_\mu\phi}{\Lambda}J^\mu}_{\text{chem pot}} + \frac{\nabla_\mu\phi\nabla^\mu\phi}{\Lambda^2}|\chi|^2 + \underbrace{\alpha\chi + h.c}_{\text{explicit } \cancel{U(1)}} \,. \end{aligned} \tag{A.15}$$

where ∇_μ is the diffeomorphism-covariant derivative and $J^\mu = i(\chi\nabla^\mu\chi^\dagger - \chi^\dagger\nabla^\mu\chi)$ is the Noether current associated with the global $U(1)$ rephasing symmetry of χ. The role of a small explicit $U(1)$ symmetry-breaking term in eq. (A.15) will becomes clear shortly. After plugging in the background for ϕ, we see that the Lagrangian contains a charge density operator, $\frac{\dot\phi_0}{\Lambda}J^0 = \lambda J^0$, with a chemical potential $\lambda = \frac{\dot\phi_0}{\Lambda}$.

The effect of a chemical potential is more apparent in a "rotated" basis $\chi \mapsto e^{-i\phi/\Lambda}\chi$, where the Lagrangian in eq. (A.15) becomes

$$\begin{aligned}\mathcal{L}_\chi \xrightarrow{\text{rotated}} & -|\nabla_\mu\chi|^2 - M^2|\chi|^2 + \alpha\chi e^{-i\phi/\Lambda} + h.c. \\ = & -|\nabla_\mu\chi|^2 - M^2|\chi|^2 + \alpha\chi e^{-i\lambda t}e^{-i\delta\phi/\Lambda} + h.c.\,, \end{aligned} \tag{A.16}$$

where we have used $e^{-i\phi/\Lambda} = e^{-i(\phi_0+\delta\phi)/\Lambda} = e^{-i\lambda t}e^{-i\delta\phi/\Lambda}$. The importance of symmetry-breaking is now manifest; it acts as a "bath" that allows us to violate χ number conservation. However,

there remains a non-linearly realised symmetry

$$\chi \mapsto e^{ia/\Lambda}\chi,$$
$$\delta\phi \mapsto \delta\phi + a, \tag{A.17}$$

protecting the flatness of the inflaton slow-roll potential, that ensures correlators are free of late time divergences.

The coupling in eq. (A.16) is effectively a linear source term for χ, where the source has frequency λ. This source can now produce particles as heavy as λ without any suppression. This is further illustrated in section A.4.1. The non-Gaussianity due to an exchange of a heavy scalar with $m < \lambda$ was found, up to an unimportant phase, to have the following form

$$F_{\rm sq} = |f_{\rm oscil}(\mu_0, \lambda)| \left(\frac{k_1}{k_3}\right)^{-3/2+i(\mu_0-\lambda)} + c.c. \tag{A.18}$$

The signal is unsuppressed for $\mu_0 < \lambda$ as expected, with a potentially observable $|f_{\rm oscil}| \sim \mathcal{O}(0.1-1)$ [35].

The key takeaways from ref. [35] for implementing the chemical potential are:

1. An explicit breaking of the heavy particle number conservation is required,

2. The chemical potential can be implemented as a phase $e^{\pm i\phi/\Lambda}$ so as to nonlinearly realise the symmetry in eq. (A.17).

We will use these simplified principles in our study to generalize the mechanism to complex vector fields.

A.3 A Chemical Potential for Tree-Level Spin-1 Exchange

A.3.1 Chemical potential in the rotated basis

This section establishes the spin-1 equivalent of eq. (A.16). Consider a complex vector field A_μ of mass m. The $U(1)$ invariant Proca Lagrangian is

$$\mathcal{L}_{\rm pr} = -\frac{1}{2}|F_{\mu\nu}|^2 - m^2|A_\mu|^2. \tag{A.19}$$

As seen in eq. (A.16) in the previous section, a chemical potential can be realized as a ϕ-dependent phase in the $U(1)$ symmetry-breaking term. The simplest $U(1)$ breaking term coupling A_μ linearly to the inflaton is naively $\mathcal{L}_{\rm int} = c\nabla_\mu\phi A^\mu e^{-i\phi/\Lambda} + c.c.$ However, this interaction is redundant because integration by parts yields a term, $\mathcal{L}_{\rm int} = -ic\Lambda(\nabla_\mu A^\mu)e^{-i\phi/\Lambda}$, which is zero in the leading order equations of motion $\nabla_\mu A^\mu = 0$. This coupling can therefore be removed by a field redefinition,

$$A_\mu \mapsto A_\mu + \frac{c^*}{m^2}\nabla_\mu\phi e^{i\phi/\Lambda}. \tag{A.20}$$

The leading non-redundant interaction that is linear in the vector field comes from a dimension-7 operator, [1]

$$\mathcal{L}_{\rm int} = \frac{c}{\Lambda^3}\nabla_\mu\phi A^\mu\nabla_\rho\phi\nabla^\rho\phi e^{-i\phi/\Lambda} + c.c. \tag{A.21}$$

Despite high dimensionality, this interaction is relevant as it contributes to the bispectrum at tree-level, which makes the calculation analytically tractable. The interaction above, along

[1] Such an interaction can arise by integrating out a heavy scalar, σ, which interacts with the inflaton via two couplings of the form $\mathcal{L} \sim \{\sigma\nabla_\mu\phi A^\mu e^{-i\phi/\Lambda}, \sigma(\nabla_\mu\phi)^2\}$.

with the Proca Lagrangian, has the residual nonlinearly realized symmetry,

$$\begin{aligned} \phi &\mapsto \phi + \alpha, \\ A_\mu &\mapsto e^{i\alpha/\Lambda} A_\mu, \end{aligned} \tag{A.22}$$

which protects the flatness of the inflaton potential, and ensures no late-time divergences as in eq. (A.17).

A.3.2 Expansion parameters

Self-interactions of the inflaton follow a derivative expansion $\mathcal{L}(\nabla_\mu \phi/\Lambda^2)$, where Λ is roughly the scale of the EFT breakdown. Requiring this to be a controlled expansion gives the first constraint

$$\Lambda \gtrsim (\dot{\phi}_0)^{1/2} \approx 60H. \tag{A.23}$$

Combining with the definition of the chemical potential in eq. (A.5), eq. (A.23) implies

$$\lambda \lesssim 60H. \tag{A.24}$$

The strength of massive vector interactions at high energies is in general sensitive to both its mass and couplings. Recall the propagator (in Minkowski space)

$$D_{\mu\nu}(p^2) = \frac{i}{p^2 + m^2 - i\epsilon} \left(\eta_{\mu\nu} - \frac{p_\mu p_\nu}{m^2} \right), \tag{A.25}$$

has a term that scales as $1/m^2$, which enhances the interactions of longitudinal modes in the UV. Determining the scale at which massive vector interactions get strong, and thus the EFT breaks down, is most easily done with the Stueckelberg trick. The high-energy behavior of the longitudinal mode can be isolated by restricting to the form [119] (see [39] for a review):

$$A_\mu \to \frac{1}{m}\nabla_\mu \pi. \tag{A.26}$$

With this form, the interaction in eq. (A.21) looks like

$$\mathcal{L}_{\text{int}} \approx \frac{c}{m\Lambda^3}\nabla_\mu\phi\nabla^\mu\pi\nabla_\rho\phi\nabla^\rho\phi e^{-i\phi/\Lambda}. \tag{A.27}$$

As expected, the interactions of the longitudinal mode are enhanced by a factor of cE/m for momenta of typical size E. Requiring that the interactions remain weak up to $E \sim \Lambda$ gives the final constraint

$$c < \frac{m}{\Lambda}. \tag{A.28}$$

A.3.3 Other lower dimension interactions

While we have stated the leading linear interaction between ϕ, and A_μ, there are lower dimension interactions that are quadratic in A_μ. We choose a power counting for these couplings consistent with eq. (A.21) in which every A_μ appearing in interactions with ϕ is multiplied by c. In this way the leading such interactions appear at dimension-6 and are given by

$$\mathcal{L}_6 \sim \frac{|c|^2}{\Lambda^2}A^\dagger_\mu A^\mu \nabla_\nu\phi\nabla^\nu\phi,\ \frac{|c|^2}{\Lambda^2}A^\dagger_\mu A_\nu \nabla^\mu\phi\nabla^\nu\phi. \tag{A.29}$$

These provide generalised mass terms $\delta m^2 \sim |c|^2\lambda^2$. However, due to equations (A.24) and (A.28) these are subleading to the mass term in (A.19), and will be neglected here.

In addition to these higher dimension terms there is the possibility of providing a chemical potential for pairs of vectors within loop level processes

$$\mathcal{L}_{\rm int} \sim |c^2|\Lambda^2 A_\mu A^\mu e^{-2i\phi/\Lambda} + h.c. \tag{A.30}$$

This is an interesting possibility which we will explore in future work, but in this paper we will focus on the tree-level effect due to the linear coupling.

A.3.4 Interactions

The interactions can be obtained from eq. (A.21) by separating the fields into their homogeneous backgrounds and fluctuations, $\phi(\vec{x},\eta) = \phi_0(\eta) + \delta\phi(\vec{x},\eta)$, $A_\mu(\vec{x},\eta) = \mathcal{A}_\mu(\eta) + A_\mu(\vec{x},\eta)$. To avoid confusion with the time component of a 4-vector we have called the vector background "$\mathcal{A}_\mu$", and we have reused "A_μ" (rather than "δA_μ") for the fluctuation to avoid overly cumbersome notation.

When the background for the vector is plugged into eq. (A.21), we are only left with inflaton self-interactions. While these interactions are important for the evaluation of the full shape of non-Gaussianity, they do not contribute to the non-analytic part as it requires on-shell propagation of a heavy particle. Therefore, we will not discuss the effects of inflaton self-interactions in this work. The full set of interactions and the expression for the homogeneous background of the vector field can be found in appendix A.8.2. Below we only collect terms that are relevant for the calculation of the non-analytic part of the NG. For the bispectrum, both

mixing and three-point vertices are important.

The quadratic mixing terms between the two fields are as follows,

$$\mathcal{L}_{\rm mix} = -\mathcal{H}_{\rm mix} = \frac{c}{\Lambda}(-\eta)^{i\lambda}\left\{i\eta\lambda^3 A_0\delta\phi + 3\eta^2\lambda^2 A_0\delta\phi' - \lambda^2\eta^2 A_i\partial_i\delta\phi\right\} + h.c.\,, \tag{A.31}$$

where the prime indicates $' \equiv \partial/\partial\eta$.

At tree-level, the relevant three-point interactions are those that contain two inflaton fluctuations and one vector fluctuation,

$$\mathcal{L}_{\rm A\phi\phi} = -\mathcal{H}_{\rm A\phi\phi} = \frac{c}{\Lambda^2}(-\eta)^{i\lambda}\left\{\frac{\lambda^3\eta}{2}A_0\delta\phi^2 - 3i\lambda^2\eta^2 A_0\delta\phi\delta\phi' - 3\lambda\eta^3 A_0(\delta\phi')^2 \right.$$
$$\left. + \lambda\eta^3 A_0(\partial_i\delta\phi)^2 + i\lambda^2\eta^2 A_i\delta\phi\ \partial_i\delta\phi + 2\lambda\eta^3 A_i\delta\phi'\ \partial_i\delta\phi\right\} + h.c. \tag{A.32}$$

In the next section we derive the central features of the NG calculation, while the detailed calculation is performed in appendix A.8.3.

A.4 Non-Gaussianity due to Massive Spin-1 Exchange

A.4.1 The stationary phase approximation

The bispectrum can be computed by plugging interactions into the in-in formula of eq. (A.12),

$$\langle\delta\phi^3\rangle|_{\eta\to 0} = \langle 0|\left\{\bar{T}e^{i\int_{-\infty(1+i\epsilon)}^{0}H_I(\eta')d\eta'}\right\}(\delta\phi^3)|_{\eta\to 0}\left\{Te^{-i\int_{-\infty(1-i\epsilon)}^{0}H_I(\eta')d\eta'}\right\}|0\rangle, \tag{A.33}$$

where we have used a shorthand $\delta\phi^3 \equiv \delta\phi_{\vec{k}_1}\delta\phi_{\vec{k}_2}\delta\phi_{\vec{k}_3}$. We restrict to the relevant interactions described in eqs. (A.31) and (A.32) in the previous section, $(\mathcal{H}_{\rm mix},\ \mathcal{H}_{\rm A\phi\phi}) \subset \mathcal{H}_I$. By expanding

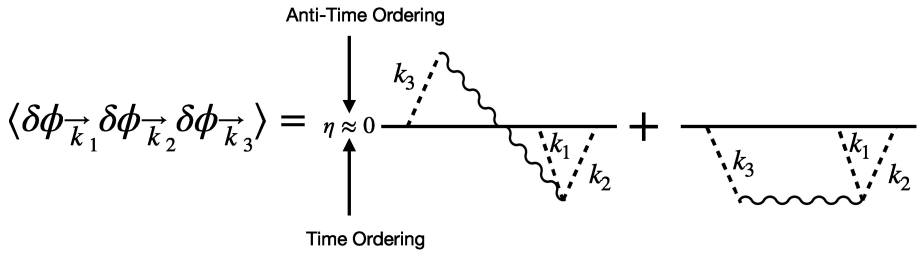


Figure A.3: The tree level contributions to the three-point function due to exchange of A_μ. Vertices above $\eta \approx 0$ come from the expansion of the anti time-ordered exponential in eq. (A.12), whilst the vertices below come from the time ordered exponential. The wavy line is the heavy vector, and the dashed lines are inflaton fluctuations.

the exponential, we can see that the dominant non-analytic contribution occurs at tree-level with two vertices and four types of diagrams,

$$\begin{aligned} \langle \delta\phi_{\vec{k}_1} \delta\phi_{\vec{k}_2} \delta\phi_{\vec{k}_3} \rangle'_{\eta\to 0} &= I_{-+} + I_{+-} + I_{++} + I_{--} \\ &= 2\mathrm{Re}(I_{-+} + I_{++}), \end{aligned} \tag{A.34}$$

where the $+/-$ notation indicates whether the mixing and the 3-point vertices are chosen from the time-ordered/anti time-ordered exponential respectively. It is easy to check that I_{+-} and I_{--} are the complex conjugates of I_{-+} and I_{++} respectively, leading to the simplification in the second line in eq. (A.34). The corresponding diagrams are shown in figure A.3.

Let us first focus on the I_{-+} diagram, which is easier to analyse and is useful in understanding the role of the chemical potential. Since the time ordering is trivial in this diagram, the calculation factorizes into two time integrals, one at each vertex,

$$I_{-+} = \frac{|c|^2}{\Lambda^3} I^-_{k_3} I^+_{k_{12}} + c.c.\,. \tag{A.35}$$

Here, $I^{-}_{k_3}$ corresponds to the integral at the mixing vertex, while $I_{k_{12}}$ corresponds to the integral at the cubic vertex. These integrals typically contain polynomial functions accompanied by highly oscillatory phase factors,

$$\begin{aligned} I^{-}_{k_3} &= \int_0^\infty dx f_-(x) e^{ig(x)}, \\ I^{+}_{k_{12}} &= \int_0^\infty dx f_+(x) e^{ih(x,p)}, \end{aligned} \tag{A.36}$$

where a convenient variable of integration, $x \equiv -k_3\eta$, is chosen and

$$p \equiv (k_1 + k_2)/k_3, \tag{A.37}$$

is the squeezing parameter.

Let us focus on the oscillatory phases as they turn out to be important in determining the presence (or absence) of exponential suppression. Firstly, the chemical potential provides a high frequency source at each vertex as seen in eqs. (A.31) and (A.32),

$$\mathcal{H}_I \propto e^{\pm i\lambda t} = e^{\mp i\lambda \log(-\eta)}. \tag{A.38}$$

Secondly, the mode functions for the fields also have oscillatory time-dependence (equivalent of e^{-iEt} in Minkowski space),

$$\begin{aligned} &\delta\phi(\vec{k},\eta) \propto e^{\mp ik\eta}, \\ &A_\mu(\vec{k},\eta) \propto e^{\mp i\int^\eta \sqrt{(-k\eta')^2+\mu^2}\,\frac{d\eta'}{\eta'}} = e^{\pm i\int^t dt' E_k(t')}, \end{aligned} \tag{A.39}$$

where $\mu = \sqrt{m^2 - 1/4} \approx m$ for $m \gg H$, and $+(-)$ in the phases corresponds to particle creation (annihilation). The full expressions for the mode functions of A_μ are given by Hankel functions, as described in appendix A.8.1. However, for $m \gg H$, they can be approximated by the WKB solutions to a harmonic oscillator with time-dependent frequency $E_k(\eta) = \sqrt{(-k\eta)^2 + m^2} = \sqrt{k_{\rm phy}^2 + m^2}$, which is used in the second line of eq. (A.39). This is essentially the locally Minkowskian energy of a massive particle in the expanding dS background. Putting these together, the phases in eq. (A.36) can be written as

$$\begin{aligned} g(x) &= -\lambda \ln x + x + \int^x \frac{dx'}{x'} \sqrt{x'^2 + \mu^2}, \\ h(x) &= \lambda \ln x - px - \int^x \frac{dx'}{x'} \sqrt{x'^2 + \mu^2}. \end{aligned} \tag{A.40}$$

We see that due to the highly oscillatory nature of the exponential in these time-integrals, the contribution is suppressed except where the phase is stationary. In $I^-_{k_3}$ integral, $g'(x) = 0$ occurs when

$$\begin{aligned} &(-k_3\eta) + \sqrt{(-k_3\eta)^2 + \mu^2} = \pm\lambda \\ &\rightarrow \; x_- \equiv (-k_3\eta_-) = \frac{\lambda^2 - \mu^2}{2\lambda}. \end{aligned} \tag{A.41}$$

Similarly, for $I^+_{k_{12}}$ integral, $h'(x) = 0$ happens when

$$\begin{aligned} &(-k_1\eta) + (-k_2\eta) + \sqrt{(-k_3\eta)^2 + \mu^2} = \pm\lambda \\ &\rightarrow \; x_+ \equiv (-k_3\eta_+) = \frac{\lambda - \mu}{p}. \end{aligned} \tag{A.42}$$

The interpretation is clear: the integrals are dominated by the time when the energy flowing through the vertex is conserved in the local Minkowski frame. The intermediate heavy particle is therefore produced on-shell, and can propagate over a long distance leading to the characteristic non-analytic signature. If the time integrals do not include a region where energy can be conserved modulo H (supplied by the expanding spacetime), putting the heavy particle on-shell will lead to an exponential suppression analogous to Boltzmann suppression with the chemical potential taken into account [35]. This is also the reason behind the Boltzmann-like suppression $e^{-\pi\mu}$ in the minimal cosmological collider signal, as a stationary point does not exist on the real axis for eqs. (A.41) and (A.42) if $\lambda = 0$. The chemical potential, through the factors of $e^{\pm i\lambda t}$, is effectively injecting or removing energy at each vertex, and is therefore crucial for overcoming the standard Boltzmann suppression of cosmological collider signals.

The non-analytic dependence can be pinned down from the phase at the stationary points,

$$I_{-+} \propto e^{ig(x_-)} e^{ih(x_+)}. \tag{A.43}$$

Since x_- only depends on μ and λ, $e^{ig(x_-)}$ only provides an overall phase. However, x_+ depends on the degree of squeezing p, and $e^{ih(x_+)}$ will contain the non-analytic dependence on p. The $h(x)$ in eq. (A.40) can be straightforwardly evaluated to be

$$h(x_+) = \lambda \ln x_+ - p x_+ - \sqrt{x_+^2 + \mu^2} - \mu \ln \left(\frac{x_+}{\mu^2 + \mu\sqrt{x_+^2 + \mu^2}} \right). \tag{A.44}$$

In the squeezed limit $p \gg 1$, every term is approximately independent of p besides the two logs,

$(\lambda - \mu)\ln(x_+)$. The result therefore has a non-analytic dependence coming from

$$e^{ih(x_+)} = p^{i(\mu-\lambda)} e^{i\delta(\mu,\lambda)}. \tag{A.45}$$

Determining the prefactors in the integrals requires considering the polynomial functions in eq. (A.36) and performing the Gaussian integral around the stationary point. The normalised bispectrum given in eq. (A.14) has the form,

$$F_{\rm sq} = \frac{1}{2} f_{\rm oscil}(\mu, \lambda, \cos\theta) p^{-5/2+i(\mu-\lambda)} + c.c. \tag{A.46}$$

Following the outline described above, the full calculation of the non-analytic component of the bispectrum is performed in appendix A.8.3. Here we simply report the final answer for $f_{\rm oscil}$:

$$f_{\rm oscil}(\mu, \lambda, \cos\theta) = \frac{40\pi|c^2|}{3\sqrt{2}m^2}\dot{\phi}_0 \frac{\lambda^3}{\Lambda^3}\sqrt{\frac{\lambda}{\mu}} \frac{C(\mu,\lambda)\left[A(\mu,\lambda) + B(\mu,\lambda)\cos^2\theta\right]}{(\lambda-\mu)(\lambda^2-\mu^2)}, \tag{A.47}$$

where functions $A(\lambda,\mu)$, $B(\lambda,\mu)$, and $C(\lambda,\mu)$ are given by

$$\begin{aligned}
A &= \frac{1}{8}\left(-3\lambda^4+8\lambda^3\mu-\lambda^2(5\mu^2+2i\mu+6)-2\lambda\mu(\mu^2-2)+2\mu^2(\mu^2+i\mu-1)\right)\\
&= -\frac{3\lambda^4}{8}\left(1+\mathcal{O}\left(\frac{\mu}{\lambda}\right)\right),\\
B &= \frac{1}{2}\left(\frac{1}{2}i(\lambda-\mu)^2-\frac{1}{4}\lambda(2i-\lambda+\mu)^2\right)\left(\sqrt{\frac{(\lambda-\mu)^2}{p^2}+\mu^2}+\frac{i\left(\frac{2(\lambda-\mu)^2}{p^2}+\mu^2\right)}{2\left(\frac{(\lambda-\mu)^2}{p^2}+\mu^2\right)}\right)\\
&= -\frac{\lambda^3\mu}{8}\left(1+\mathcal{O}\left(\frac{\lambda^2}{\mu^2p^2}\right)\right),\\
C &= -\frac{3(\lambda^2-\mu^2)^2}{4\lambda^2}+i\left(1-\frac{i(\lambda^2-\mu^2)}{2\lambda}\right)\left(\lambda+\frac{1}{2}\sqrt{\frac{(\lambda^2+\mu^2)^2}{\lambda^2}}+\frac{i(\lambda^4+\mu^4)}{(\lambda^2+\mu^2)^2}\right)\\
&= 2i\lambda\left(1+\mathcal{O}\left(\frac{\mu^2}{\lambda^2}\right)\right),
\end{aligned} \tag{A.48}$$

where we used $\lambda \gg \mu$, and $p > \lambda/\mu$ for the approximations. Note that B is smaller than A by a factor of μ/λ, which is relevant for the determination of spin as discussed in section A.5.3. Also in the smaller squeezing regime, $p < \lambda/\mu$, B will have some modest p-dependence. Observing the non-analytic signal in this regime will be useful in determining the particles mass, and will be discussed in the next section.

A.4.2 Other subdominant diagrams

The vector propagator in all four diagrams requires an A_μ to contract with an $A_\mu^\dagger$, which means one vertex gets $e^{i\lambda t}$, whilst the other gets $e^{-i\lambda t}$. In the I_{-+} diagram λ is either injected or removed at both vertices, which corresponds to the propagation of an anti-particle or a particle, respectively. We studied the case of anti-particle propagation above, where energy is injected at both vertices, and a stationary point therefore exists in both integrals. It is then

straightforward to conclude that stationary points do not exist for particle propagation, where energy is instead removed from both vertices. This diagram (and its complex conjugate in I_{+-}) is therefore exponentially suppressed as $\sim e^{-2\pi(\lambda+\mu)}$. For the I_{++} and I_{--} diagrams, the energy from the chemical potential can only be injected at one vertex, and must be taken away at the other. Therefore, there is always one time-integral in which a stationary point does not exist. This gives a slightly less severe but still exponential suppression of the form $\sim e^{-\pi(\lambda-\mu)}$. For more detailed discussion, we refer the reader to ref. [35]. We will henceforth only focus on the dominant diagram I_{-+}.

A.4.3 Apparent late-time divergences

Lastly, we would like to address the apparent divergences in the integrals. It seems that the inflaton fluctuations in the vertices that come from the expansion of the exponential phase $e^{-i\delta\phi/\Lambda} = 1 - i\frac{\delta\phi}{\Lambda} + \cdots$ are not derivatively coupled. Non-derivative couplings of the inflaton typically lead to late-time ($\eta \to 0$) divergences in the integrals,

$$I = \int_{-\infty}^{0} \frac{d\eta}{\eta^4} f(\eta) e^{ig(\eta)}. \tag{A.49}$$

This is because the inflaton fluctuations become constant at late time (i.e. superhorizon freezing) and $f(\eta) \propto \eta^{3/2}$ coming from the heavy field dilution is not sufficient to nullify the divergence $1/\eta^4$ from the metric factor. However, the divergences in our calculations are purely artifacts of the rotated basis that we are using. The residual symmetry in eq. (A.22) ensures that the late-time diverges cancel when we sum over all Feynman diagrams, while the stationary phase contributions that are coming from early time remain. The reader is referred to appendix A.8.4

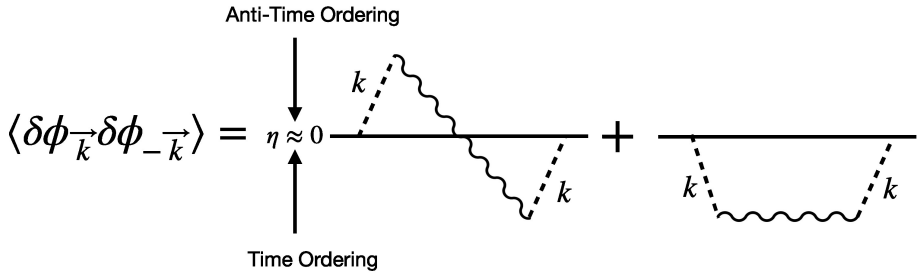


Figure A.4: The tree level contributions to the power spectrum involving the exchange of the vector. The wavy line is the heavy vector, and the dashed lines are inflaton fluctuations. Kinematic considerations show that the I_{-+} diagram on the left is dominant, and the result can be calculcated in the stationary phase approximation, as was done for the NG.

for more details.

A.4.4 Correction to the power spectrum

Ensuring corrections to the two-point function are small compared to the usual primordial power spectrum provides a further constraint on $|c|$. To estimate the leading correction we can first consider the case $m > \lambda$. While this case is not the focus of this paper, the calculation can be used to infer the constraint in the region of interest $m < \lambda$. For $m > \lambda$, the vector can be straightforwardly integrated out by solving the equations of motion for A_μ in eq. (A.21) and ignoring the kinetic term

$$A^\mu = \frac{\bar{c}}{m^2} \frac{1}{\Lambda^3} \nabla^\mu \phi (\nabla_\rho \phi)^2 e^{-i\phi/\Lambda}. \tag{A.50}$$

Plugging back into the Lagrangian renormalizes the purely inflaton pieces. Expanding about the slow-roll background leads to

$$\mathcal{L}_{\text{eff}} = -\frac{1}{2} \mathcal{Z} (\nabla_\mu \delta\phi)^2 + \mathcal{L}_{\text{int}}(\delta\phi), \tag{A.51}$$

where $\mathcal{Z}-1=\mathcal{O}(1)|c|^2\lambda^4/m^2\Lambda^2$. The fluctuations should be canonically normalised by $\delta\phi\to\delta\phi/\sqrt{\mathcal{Z}}$. This has an effect on the power spectrum

$$\langle\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}\rangle'=\frac{1}{\mathcal{Z}}\frac{1}{2k^3}=\frac{1}{2k^3}\left[1+\mathcal{O}(1)\frac{|c|^2\Lambda^2}{m^2}\frac{\lambda^4}{\Lambda^4}\right]. \tag{A.52}$$

Since eq. (A.52) is scale invariant, sizable corrections to the power spectrum are tolerable. For example, imposing the conservative constraint that the correction is at most 10% of the free field contribution yields

$$|c|^2\frac{\Lambda^2}{m^2}<0.1\frac{\Lambda^4}{\lambda^4}. \tag{A.53}$$

In the scenario of interest for cosmological collider physics, $m\lesssim\lambda$, there are two regimes: in one regime the heavy particle is significantly off shell, just as in the $m\gtrsim\lambda$ case, and we can expect its contribution to be of similar size. There is also the possibility of the heavy particle being produced on shell, dominated by the I_{-+} diagram for similar kinematic reasons to the NG. The tree level contributions involved the exchange of the heavy vector are shown in figure A.4. The I_{-+} diagram factorises and can be readily calculated, yielding

$$I_{-+}=\mathcal{O}(1)\frac{1}{2k^3}\frac{|c|^2}{m^2\Lambda^2}\frac{\lambda^5}{(\lambda^2-\mu^2)^2}|C(\mu,\lambda)|^2\xrightarrow{\mu\ll\lambda}\mathcal{O}(1)\frac{1}{2k^3}\frac{|c|^2}{m^2\Lambda^2}4\lambda^3, \tag{A.54}$$

where $C(\mu,\lambda)$ is given in eq. (A.48). Comparing with eq. (A.52), the on-shell contribution is subdominant for masses $\mu<\lambda/2$. Therefore, we obtain the same conservative constraint that we derived from $m\gtrsim\lambda$, namely equation eq. (A.53).

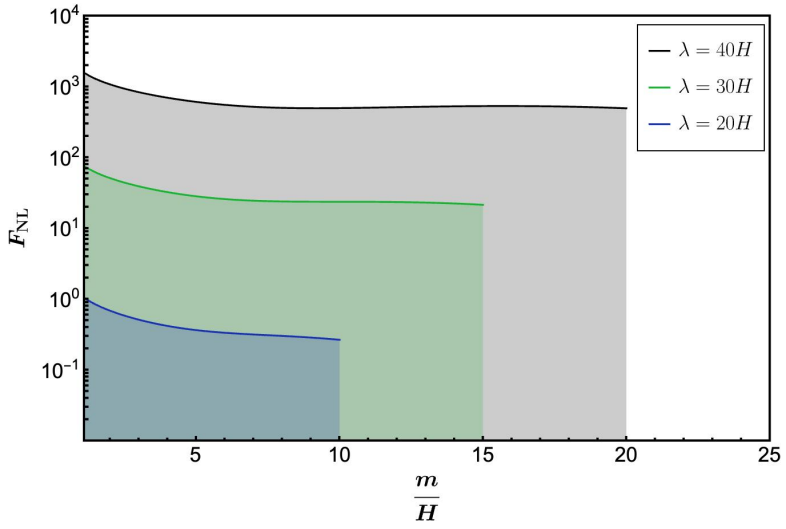


Figure A.5: A plot of the maximum possible signal size $|f_{\rm oscil}|$ within theoretical control as a function of the mass m of the heavy vector field in inflationary Hubble units. We have fixed $\cos\theta = 0.5$. For $\mu \lesssim \lambda/2$, the signal strength is within the sensitivity of future LSS and 21-cm experiments.

A.5 Signal Strength and Key Features

A.5.1 Signal size

The size of the signal is characterized by the amplitude $|f_{\rm oscil}|$ in eq. (A.47). In figure A.5 we plot the size of the signal $|f_{\rm oscil}|$ for the largest value of $|c|/m$ consistent with eqs. (A.28) and (A.53). We also restrict to values of $\lambda \leq 40H$, in accord with the constraint eq. (A.24) from the requirement of perturbative control.

As expected, the exponential suppression is nullified and $|f_{\rm oscil}|$ is enhanced for $\mu \lesssim \lambda$. For heavier masses beyond the chemical potential, the nonanalyticity in the NG drops exponentially as the on-shell production becomes inefficient. The increase in the maximum allowed value of the coupling c with mass is compensated by the $1/m^2$ factor in the NG, as seen in eq. (A.47), which leads to roughly flat profiles in figure A.5. We see that depending on the value of the

chemical potential, a signal size of

$$|f_{\rm oscil}| \sim \mathcal{O}(1 - 10^3) \tag{A.55}$$

is theoretically possible for a wide range of masses. A large portion of this signal space is not excluded by current data, whilst still being able to be probed by future large-scale structure and 21-cm experiments [26, 110, 111]. Note that the stationary phase approximation should not be trusted for values of μ close to λ, since in this limit the stationary points are pushed to late times, $x_-, x_+ \to 0$, and the non-analyticity becomes approximately analytic. Therefore, the non-analytic signal would not be separable from the background analytic part leftover after the cancellation of IR divergences and that coming from the inflaton self-interactions. For this reason, we restrict to regions $\mu < \lambda/2$ in figure A.5, with the largest mass $m \sim 20H \sim 10^{15}\,\text{GeV}$. Even with this conservative choice, the chemical potential allows us to reach the typical mass of the KK gauge bosons in a SUSY orbifold GUT scenario, to be discussed in section A.6.

A.5.2 Inferring the mass and the chemical potential

As seen in eq. (A.47), the non-analytic dependence in the squeezed limit has the form

$$F_{\rm sq} \propto p^{i(\mu-\lambda)}, \tag{A.56}$$

which allows for an observation of the combination $(\mu - \lambda)$. Figure A.6 shows the non analytic dependence on squeezing for $\mu = 20H$, with $\lambda = 40H$.

However, this form is correct only in the very squeezed limit. To see this, let us see more

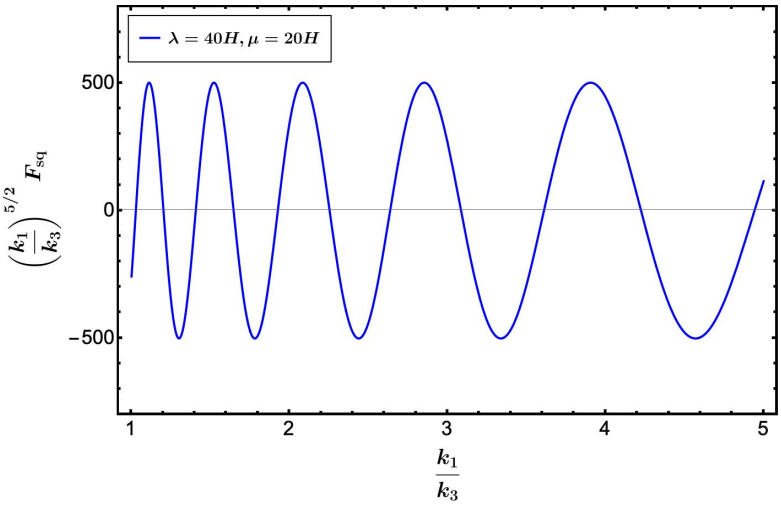


Figure A.6: A plot of the non analytic dependence of $F_{\rm sq}$ as a function of the squeezing (k_1/k_3) with $\cos\theta = 1/2$. A measurement of the frequency of oscillation will allow a determination of the combination $\mu - \lambda$.

carefully how the stationary point at k_{12}-vertex (x_+) changes with squeezing. Using $m \approx \mu$, eq. (A.42) can be solved to get

$$x_+ = \frac{p\lambda}{p^2-1} - \frac{\sqrt{\lambda^2 + p^2\mu^2 - \mu^2}}{p^2-1}. \tag{A.57}$$

At large squeezing such that $p > \lambda/\mu$, all but the second term in the square root can be dropped. On the other hand, in the regime of modest squeezing with $p < \lambda/\mu$, all but the first term can be neglected. Therefore,

$$x_+ = \begin{cases} \frac{(\lambda-\mu)}{p}\,, & p \gg \lambda/\mu \\ \\ \frac{\lambda}{1+p}\,, & p \ll \lambda/\mu. \end{cases} \tag{A.58}$$

Since $p > 2$, such a regime can only exist if $\lambda > 2\mu$, which we are restricting to anyway.

This change in the expression of the stationary point at k_{12}-vertex in different regions of

the squeezing parameter p has a direct effect on the imaginary exponent in $F_{\rm sq}$. To see this, let us zoom in on the non-analytic dependence that comes from the phase in eq. (A.44),

$$h(x_+) = \lambda \ln x_+ - p x_+ - \sqrt{x_+^2 + \mu^2} - \mu \ln \left(\frac{x_+}{\mu^2 + \mu \sqrt{x_+^2 + \mu^2}} \right). \tag{A.59}$$

For $p \gg \lambda/\mu$ and $\lambda > \mu$, $x_+ \ll \mu$, and the phase simplifies to $(\lambda - \mu)\ln(x_+)$. However, in the modest squeezing regime, $x_+ \gg \mu$, and the second log term in the above expression becomes approximately independent of p. The non-analytic dependence in this case comes entirely from the first log term, which depends only on λ. Therefore, the non-analyticity in these two regimes goes as

$$F_{\rm sq} \propto \begin{cases} p^{i(\mu-\lambda)}\,, & p \gg \lambda/\mu \\ (1+p)^{i\lambda}\,, & p \ll \lambda/\mu. \end{cases} \tag{A.60}$$

Observation of this change in the non-analytic exponent as we change the amount of squeezing can be used to infer both the mass and the chemical potential, paralleling the results of spin-0 case from ref. [35].

A.5.3 Constraining the spin

In ref. [1] it was claimed that NG due to spin-j exchange would have a characteristic dependence on the shape of the momentum-conservation triangle. They used the minimal coupling

$$\mathcal{L} \sim (\nabla^{\mu_1}\phi)\cdots(\nabla^{\mu_j}\phi) A_{\mu_1\cdots\mu_j}, \tag{A.61}$$

leading to a signal of the form

$$F_{\rm sq} \propto P_j(\cos\theta)\left(\frac{k_1}{k_3}\right)^{-3/2+i\mu} + c.c.\,, \tag{A.62}$$

where P_j is the Legendre polynomial of a degree that is determined by the spin j. Observing the signal as a function of $\cos\theta$ would then allow determination of the spin.

However, it was pointed out in ref. [33] that for odd spin particles, a cancellation between diagrams leads to a more severe dilution factor of $(k_1/k_3)^{-5/2}$. In [49, 120] the authors found a $\sin^2\theta$ dependence for NG mediated by spin-1, by considering a minimal higher dimension interaction. Here, because of the chemical potential, we find a different shape dependence for spin-1 of the form

$$F_{\rm sq} = \left[\mathcal{F}_1(\mu,\lambda) + \mathcal{F}_2(\mu,\lambda)\cos^2\theta\right]\left(\frac{k_1}{k_3}\right)^{-5/2+i(\mu-\lambda)}, \tag{A.63}$$

where $\mathcal{F}_i(\mu,\lambda)$ can be read off from eq. (A.47).

It appears that an observation of the dilution will indicate the exchange of an odd spin particle, but it may be hard to pin down the exact spin since there is no general form for the shape dependence. Further investigation is therefore needed to know if the shape dependence can unambiguously determine the spin in cosmological collider physics. It turns out that in our model, $\mathcal{F}_2$ is smaller than $\mathcal{F}_1$ by a factor of μ/λ, so it may be difficult to observe the θ dependence. However, an observation of any dependence on $\cos\theta$ for the non-analytic signal will be a smoking gun of non-zero spin.

A.6 Embedding the Chemical Potential within Orbifold Unification

A.6.1 Trinification

Implementing a coupling of the form in equation eq. (A.21) requires the vector be a standard model (SM) gauge singlet because the inflaton is uncharged under the SM. The simplest $SU(5)$, $SO(10)$ GUTs do not include a complex vector SM singlet. However, the classic trinification[2] scenario with a gauge group [50]

$$G_{\rm GUT} = SU(3)_C \times SU(3)_L \times SU(3)_R/\mathbb{Z}_3, \tag{A.64}$$

breaking to $G_{\rm SM} = SU(3)_C \times SU(2)_L \times U(1)_Y$ contains several neutral heavy gauge bosons. The discrete $\mathbb{Z}_3$ exchanges the three $SU(3)$ groups ensuring a single gauge coupling.

The symmetry breaking pattern for the gauge bosons is as follows

$$\begin{aligned}
(8,1,1) \quad &\to (8,1)_0, \\
(1,8,1) \quad &\to (1,3)_0 \oplus (1,2)_{1/2} \oplus (1,2)_{-1/2} \oplus (1,1)_0, \\
(1,1,8) \quad &\to 4(1,1)_0 \oplus 2(1,1)_1 \oplus 2(1,1)_{-1},
\end{aligned} \tag{A.65}$$

where the numbers indicate the dimension of the representations under the GUT (left) or SM gauge group (right) with the subscripts indicating the hypercharge. Of the 5 SM singlets one linear combination is the hypercharge gauge boson, leaving 4 real, or 2 complex massive vectors as candidates for a cosmological collider signal.

[2]Note that all three unified groups mentioned are subgroups of E_6

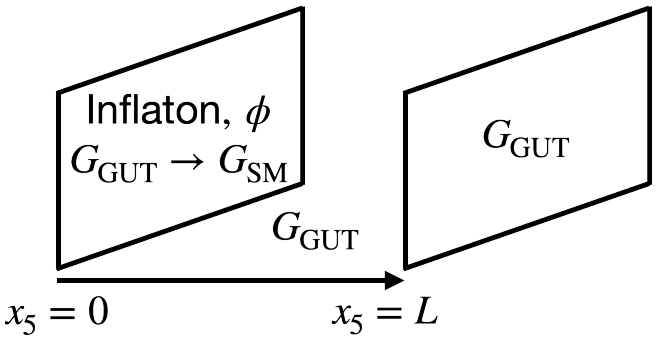


Figure A.7: In orbifold GUTs the symmetry breaking occurs via boundary condition in an extra dimension. We consider the gauge fields to live on an interval of length L. We must impose either Dirichlet or Neumann BCs at either end of the interval. Imposing Dirichlet BCs will remove the zero mode, explicitly breaking the gauge invariance in the IR for that particular generator. Neumann BCs at both ends will leave a zero mode, such that after integrating out the extra dimension, the IR theory respects the gauge invariance for that generator.

A.6.2 Realising the chemical potential coupling

The coupling in eq. (A.21) can be realised in the context of orbifold GUTs within 5-dimensional effective field theory. In orbifold GUTs, the symmetry breaking, illustrated in figure A.7, is achieved by choice of boundary conditions (BC) on one or more (3+1) dimensional boundaries of a higher dimensional "bulk". The full GUT gauge symmetry is respected by the bulk action, but only the SM gauge invariance is respected on the SM boundary at $x_5 = 0$. The inflaton is a purely 4D field localised on the boundary at $x_5 = 0$ [49].

The orbifold GUT symmetry breaking from trinification down to the standard model gauge group was studied in [51, 52]. On the SM boundary, we impose the SM-gauge-invariant

conditions

$$\nabla_5 A_\mu^{\rm SM}|_{x_5=0} = 0, \tag{A.66}$$

$$A_\mu^{\not{\rm SM}}|_{x_5=0} = 0, \tag{A.67}$$

where $\not{\rm SM}$ indicates broken GUT generators. The boundary at $x_5 = L$ respects the full GUT gauge invariance, and Neumann BCs are imposed on all generators

$$\nabla_5 A_\mu|_{x_5=L} = 0. \tag{A.68}$$

After performing a KK decomposition, the BCs in eqs. (A.66) and (A.68) ensure the existence of a zero mode for the SM gauge bosons, whereas eq. (A.67) removes the zero mode for the gauge bosons of the broken generators, leaving only the massive KK modes.

Let us denote by $Y_\mu = (A_1^{\not{\rm SM}} + iA_2^{\not{\rm SM}})_\mu$ one of the candidate complex vectors corresponding to two broken generators that are also SM singlets. The chemical potential coupling can then be implemented via the interaction:

$$S_{\rm boundary} \supset \frac{c_1}{\Lambda_{5D}^{4.5}} \int d^4x \nabla_\mu \phi (\nabla_\rho \phi \nabla^\rho \phi) e^{-i\phi/\Lambda} \nabla_5 Y_\mu|_{x_5=0}, \tag{A.69}$$

where the 5D effective gauge theory strong coupling scale $\Lambda_{5D} \sim 16\pi^2/N g_5^2$ with g_5 the 5D gauge coupling, and $N = 3$ are the number of colors of the various trinification groups. Notice that since Y_μ is a SM gauge singlet the coupling respects the SM gauge symmetry, but not the gauge symmetry associated with the broken generators, which is allowed for terms localised to

the boundary at $x_5 = 0$.

The 5D gauge field KK-decomposes as

$$Y_\mu(x, x_5) = \sqrt{\frac{2}{L}} \sum_{n=1}^{\infty} \sin\left(\frac{(n - \frac{1}{2})\pi x_5}{L}\right) Y_\mu^{(n)}(x). \tag{A.70}$$

We will focus on the first KK mode with $n = 1$. The coupling in eq. (A.21) can therefore be identified with

$$c = \frac{c_1 \pi}{\sqrt{2}} \left(\frac{\Lambda}{\Lambda_{5D}}\right)^{3/2} (\Lambda L)^{-3/2}, \tag{A.71}$$

where Λ is the cutoff of the inflationary EFT which must be before the 5D gauge theory cutoff $\Lambda \leq \Lambda_{5D}$.

In the extra-dimensional framework the $5D$ gauge coupling, g_5, is non-renormalisable. Below the compactification scale $M_C = \pi/L$ we match on to the SM or MSSM effective theory. Tree level matching predicts gauge coupling unification at, M_C

$$g_i(M_C) = \frac{g_5}{\sqrt{L}}, \tag{A.72}$$

where $g_{i=1,2,3}$ are the SM gauge couplings.

If we run SM gauge couplings up to high energies (where hypercharge is given its usual GUT normalization), they unify crudely around the scale $M_{\rm GUT} \sim 10^{14}$ GeV. The discrepancy may then be plausibly accounted for by loop level and SM brane-localised corrections to eq. (A.72) that depend on the index i. A unification can be achieved with a compactification scale of $M_C \sim 5 \times 10^{13}$ GeV. As the reader may have noticed, this scale is very close to the inflationary Hubble parameter in high-scale inflation from eq. (A.1). The cosmological collider prospects for

such orbifold GUTs were explored in ref. [49]. One downside of this setup however is that the corrections must be quite large since the unification in the SM is imperfect.

This issue is dramatically improved in the minimal supersymmetric standard model (MSSM) as the running gauge couplings come much closer to unifying than in the non-supersymmetric case, and SUSY helps to solve the gauge hierarchy problem between the unification and electroweak scales. In the SUSY scenario, only small corrections to equation eq. (A.72) are sufficient to fit the data to the unification hypothesis. A detailed accounting of corrections to eq. (A.72) was given in [48, 51, 52], which points to a compactification scale $M_C \sim 2 \times 10^{15}$ GeV in the SUSY case. Interestingly, as pointed out in ref. [51], this particular value of the compactification scale leads to the gauge couplings unifying at the 5D Planck scale. The UV completion of such a model may unify all fundamental forces including gravity. While the compactification scale here is higher than the Hubble scale of inflation H, it can still be brought within the reach of cosmological collider discovery via our chemical potential mechanism.

The full model-building of an orbifold GUT coupled to the inflaton with our chemical potential within a supersymmetric (supergravity) framework is left to future work. But we expect this to be readily achieved, with the central relevant features already described above. For the sake of simplification, we have also assumed that only the first KK mode is below the chemical potential scale, i.e. $m_n < \lambda$ only holds for $n = 1$. Here m_n is the mass of the nth KK mode ($Y_\mu^{(n)}$),

$$m_n = \left(n - \frac{1}{2}\right) M_C. \tag{A.73}$$

Notice that the analysis of ref. [51] points to a mass for the first KK mode of $m_1 \sim 10^{15}$ GeV. Combined with the restriction $m_1 \lesssim \lambda/2$ to trust the stationary phase calculation this amounts

to requiring $m_1 \in (\lambda/3, \lambda/2)$. The analysis of the previous sections now hold for $Y_\mu^{(1)}$. It would be interesting for future work to consider the case where more than one of the lowest lying KK modes are accessible, $m_n < \lambda$ for more than one n.

A.7 Discussion

In this paper, we have demonstrated that a chemical potential mechanism can be given to a complex spin-1 particle coupled to the inflaton while maintaining EFT control, allowing its propagation to be detected at tree-level. In line with the results of ref. [35] for spin-0, we find that the chemical potential can lead to non-Boltzmann-suppressed signals in non-Gaussianities (NGs) for masses up to the chemical potential, which can be as large as $60H$. The size of the signal can be within the reach of future experiments in a large part of the parameter space.

With the extended reach for single-production of massive spin-1 particles offered by our chemical potential mechanism, the heavy gauge bosons of orbifold-trinification can be targeted. While this also includes the more motivated supersymmetric unification, which has a higher GUT scale compared to the non-supersymmetric case, further model-building is required to realize inflation and chemical potential within full supersymmetric dynamics. While in this paper we have focused on a chemical potential mechanism that allows single-production (therefore at tree level) of heavy complex fields, we saw in section A.3 that there is also a $U(1)$-breaking term quadratic in the complex vector field which would give rise to pair-production (at loop level). Like the inflaton, the pair would have to be a SM-singlet, but now each member of the pair may be SM-charged. We leave this exploration to future work.

We see that an observation of the non-analyticity at intermediate and large squeezing can

be used to determine both the mass and chemical potential. Our results are also consistent with the general expectation that NGs of odd-spin heavy particle exchange come with a dilution factor of $(k_1/k_3)^{-5/2}$ [33]. It would therefore be possible to say that a cosmological collider signal is due to exchange of a particle with odd spin, but we found a dependence on the shape of the triangle that seems to suggest it may not be possible to say for certain the particle is spin-1. In future work it would be interesting to explore the possibility of determining the spin of a particle observed in cosmological collider physics. In addition to this, the chemical potential mechanism seems generalisable to higher spins, and we hope to apply this methodology to search for higher spin states in the future, including those that might arise in superstring theory.

Acknowledgments

We would like to thank Soubhik Kumar for helpful discussions. The authors were supported by NSF grant PHY-2210361 and the Maryland Center for Fundamental Physics. In addition, AB was supported by DOE grant DE-SC-0013642 and by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

A.8 Appendix

A.8.1 Mode functions

The mode functions for the inflaton and a massive spin-1 particle in dS space are given below (see ref. [33] for a full derivation). After quantisation, the Fourier transformed inflaton

fluctuation is

$$\delta\phi_{\vec{k}}(\eta) = \bar{u}_k(\eta)c_{-\vec{k}} + u_k(\eta)c^\dagger_{\vec{k}}\,, \tag{A.74}$$

where $c^\dagger$, c are the creation and annihilation operators, respectively. The mode function is given by

$$u_k(\eta) = \frac{1}{\sqrt{2k^3}}(1 - ik\eta)e^{ik\eta}, \tag{A.75}$$

where $k = |\vec{k}|$.

For the vector, we focus on the longitudinal mode, which is the only polarization that is affected by the chemical potential in our implementation. The longitudinal mode of a complex vector is quantized as

$$A^L_\mu(\vec{k},\eta) = \bar{\sigma}_\mu(k,\eta)\hat{a}_{-\vec{k}} + \sigma_\mu(k,\eta)\hat{b}^\dagger_{\vec{k}}, \tag{A.76}$$

where $\hat{a}_{\vec{k}}, \hat{b}_{\vec{k}}$ are the annihilation operators for the longitudinal mode of the particle, antiparticle respectively. The bar indicates complex conjugation. For the longitudinal mode

$$\bar{\sigma}_\mu = \left(\bar{\sigma}_0(k,\eta), \bar{\sigma}_1(k,\eta)\vec{\epsilon}_{\vec{k}}\right); \quad \vec{\epsilon}_{\vec{k}} = \frac{\vec{k}}{k}, \tag{A.77}$$

with explicit forms

$$\begin{aligned}
\bar{\sigma}_0 &= \mathcal{A}N_0(-k\eta)^{3/2}H^{(1)}_{i\mu}(-k\eta), \\
\bar{\sigma}_1 &= \frac{i}{2}\mathcal{A}N_0(-k\eta)^{1/2}\left[-H^{(1)}_{i\mu}(-k\eta) - (-k\eta)\left(H^{(1)}_{i\mu+1}(-k\eta) - H^{(1)}_{i\mu-1}(-k\eta)\right)\right],
\end{aligned} \tag{A.78}$$

where $\mu = \sqrt{m^2 - 1/4}$. The prefactors are determined by normalisation,

$$N_0 = \sqrt{\frac{\pi}{2}} \frac{1}{\sqrt{2k}} \frac{1}{m},$$
$$\mathcal{A} = e^{i\pi/4} e^{-\pi\mu/2}. \tag{A.79}$$

If the particles mass is large compared to Hubble, $\mu \gg H$, we can approximate the mode functions by their adiabatic (WKB) solution

$$\bar{\sigma}_0(x = -k\eta) \approx \frac{1}{m\sqrt{2k}} \frac{x^{3/2}}{(x^2+\mu^2)^{1/4}} e^{i\int^x \sqrt{1+\mu^2/x'^2}dx'}, \tag{A.80}$$
$$\bar{\sigma}_1(x = -k\eta) \approx -\frac{1}{m\sqrt{2k}} \left(\frac{i}{2} \frac{2x^2+\mu^2}{x^2+\mu^2} + \sqrt{x^2+\mu^2} \right) \frac{x^{1/2}}{(x^2+\mu^2)^{1/4}} e^{i\int^x \sqrt{1+\mu^2/x'^2}dx'}. \tag{A.81}$$

We will use these forms instead of eq. (A.78) henceforth. This will greatly simplify the calculation as we will see in appendix A.8.3.

A.8.2 The full set of interactions

The full Lagrangian along with the interaction given in eq. (A.21) is

$$\mathcal{L} = -\frac{1}{2}(\nabla_\mu \phi)^2 - V_{\rm s.r}(\phi) - \frac{1}{2}|F_{\mu\nu}|^2 - m^2|A_\mu^2| + \frac{c}{\Lambda^3} \nabla_\mu \phi A^\mu (\nabla_\rho \phi)^2 e^{-i\phi/\Lambda} + h.c.\,, \tag{A.82}$$

where $V_{\rm s.r.}(\phi)$ is the slow roll potential that drives inflation. It is easiest to calculate the background fields by performing the field redefinition

$$A_\mu \mapsto A_\mu + \frac{c^*}{m^2\Lambda^3} (\nabla_\rho \phi)^2 \nabla_\mu \phi e^{i\phi/\Lambda}, \tag{A.83}$$

which removes the interaction term and replaces it with

$$\mathcal{L}_{\text{int}} = \frac{c}{m^2\Lambda^3} F^{\mu\nu}\nabla_\mu\phi\nabla_\nu((\nabla_\rho\phi)^2)e^{-i\phi/\Lambda} + h.c. \tag{A.84}$$

In this basis, the interaction becomes trivial if all three ϕ's are replaced by the background, $\phi_0(\eta)$. To see this, notice that $(\nabla_\rho\phi)^2$ is a scalar, so the second covariant derivative can be replaced by a partial derivative, ∂_ν. If both fields are set to their background configuration, $(\nabla_\rho\phi_0)^2 = (\lambda\Lambda)^2$ is constant. The partial derivative then gives $\mathcal{L}_{\text{int}}(\phi_0) = 0$. Therefore, after the field definition the vector cannot obtain a VEV. Undoing the field redefinition in eq. (A.83) then implies that in the original basis the vector field has a background $\mathcal{A}_\mu(\eta)$, which can be found by plugging ϕ_0 into the shifting term in eq. (A.83),

$$\mathcal{A}_\mu(\eta) = \delta^0_\mu \frac{c^*}{m^2}\frac{\lambda^3}{\eta}(-\eta)^{-i\lambda}. \tag{A.85}$$

Expanding both fields about their background, $\phi \mapsto \phi_0 + \delta\phi$ and $A_\mu \mapsto \mathcal{A}_\mu + A_\mu$, gives the full Lagrangian

$$\begin{aligned}
\mathcal{L} \quad &= \mathcal{L}_{\text{free}} + c(-\eta)\lambda^3(-\eta)^{i\lambda}A_0\left(e^{-i\delta\phi/\Lambda} - 1\right) - 2\frac{|c|^2}{m^2}\lambda^6\cos(\delta\phi/\Lambda) \\
&\quad -\frac{c}{\Lambda}\lambda^2(-\eta)^{i\lambda}\nabla_\mu\delta\phi A^\mu e^{-i\delta\phi/\Lambda} + 2\frac{c}{\Lambda}\lambda^2\eta^2(-\eta)^{i\lambda}A_0\delta\phi' e^{-i\delta\phi/\Lambda} \\
&\quad -6\frac{|c|^2}{m^2}\frac{\lambda^5}{\Lambda}(-\eta)\delta\phi'\cos(\delta\phi/\Lambda) - \frac{c}{\Lambda^2}\lambda(-\eta)(-\eta)^{i\lambda}A_0(\nabla_\rho\delta\phi)^2 e^{-i\delta\phi/\Lambda} \\
&\quad -2\frac{c}{\Lambda^2}\lambda(-\eta)(-\eta)^{i\lambda}\nabla_\mu\delta\phi A^\mu\delta\phi' e^{-i\delta\phi/\Lambda} + 2\frac{|c|^2}{m^2}\frac{\lambda^4}{\Lambda^2}(\nabla_\rho\delta\phi)^2\cos(\delta\phi/\Lambda) \\
&\quad -4\frac{|c|^2}{m^2}\frac{\lambda^4}{\Lambda^2}(\delta\phi')^2\cos(\delta\phi/\Lambda) + \frac{c}{\Lambda^3}(-\eta)^{i\lambda}\nabla_\mu\delta\phi A^\mu(\nabla_\rho\delta\phi)^2 e^{-i\delta\phi/\Lambda} \\
&\quad -2\frac{|c|^2}{m^2}\frac{\lambda^3}{\Lambda^3}\eta\delta\phi'(\nabla_\rho\delta\phi)^2\cos(\delta\phi/\Lambda) + h.c.
\end{aligned} \tag{A.86}$$

A.8.3 Detailed stationary phase calculation

The leading contribution to the three point function is calculated by expanding eq. (A.12) to first order in perturbations. As argued in the main text, the leading contribution comes from the I_{-+} diagram

$$I_{-+} = \langle 0| \int d^4x_- \sqrt{-g} i\mathcal{H}_{\rm mix} \delta\phi_{\vec{k}_3} \delta\phi_{\vec{k}_1} \delta\phi_{\vec{k}_2} \int d^4x_+ \sqrt{-g}(-i)\mathcal{H}_{A\phi\phi}|0\rangle + c.c. \tag{A.87}$$

After contracting all fields and factoring out momentum-conserving delta functions, the result factorises into two integrals

$$I_{-+} = \frac{|c|^2}{\Lambda^3} I^-_{k_3} I^+_{k_{12}} + c.c. \tag{A.88}$$

The first integral over the two point mixing vertex is

$$I^-_{k_3} = \frac{1}{\sqrt{2k_3^3}} \int_{-\infty}^{0} \frac{d\eta}{\eta^4} (-\eta)^{-i\lambda} \left\{ (-i\eta\lambda^3 \bar{u}_{k_3} + 3\eta^2\lambda^2 \bar{u}'_{k_3}) \bar{\sigma}_0(k_3,\eta) - i\lambda^2\eta^2 k_3 \bar{u}_{k_3} \bar{\sigma}_1(k_3,\eta) \right\} \tag{A.89}$$

First insert the mode functions from eqs. (A.75), (A.80), and (A.81), then make the substitution $x = -k_3\eta$ to massage this into a form appropriate for applying stationary phase

$$I^-_{k_3} = \frac{\lambda^2}{2\sqrt{2}mk_3^{3/2-i\lambda}} \int_0^\infty dx \frac{C(x)}{x^{3/2}(x^2+\mu^2)^{1/4}} e^{ig(x)}. \tag{A.90}$$

The two functions are given by

$$C(x) = i(1-ix)\left(\lambda+\sqrt{x^2+\mu^2}+\frac{i}{2}\frac{2x^2+\mu^2}{x^2+\mu^2}\right)-3x^2, \tag{A.91}$$

$$g(x) = -\lambda\ln x + x + \int^x \sqrt{1+\mu^2/y^2}dy. \tag{A.92}$$

This integral has a late-time divergence in the limit $x \to 0$. However, the residual shift symmetry ensures that the late-time divergences cancel in the sum over all Feynman diagrams. This is discussed further in appendix A.8.4. Since the stationary phase occurs at early time, well away from the $\eta \to 0$ boundary, the contribution from around the stationary point is still a valid method to evaluate these integrals.

The stationary phase approximation $g'(x_-) = 0$, yields $x_- = (\lambda^2-\mu^2)/2\lambda$, and $g''(x_-) = 4\lambda^3/(\lambda^4-\mu^4)$ for the time and (inverse of the) variance respectively. The result is

$$I^-_{k_3} \approx \frac{\lambda^2}{2\sqrt{2}mk_3^{3/2-i\lambda}}\sqrt{\frac{2\pi}{|g''(x_-)|}}\frac{C(x_-)e^{ig(x_-)}}{x_-^{3/2}(x_-^2+\mu^2)^{1/4}}, \tag{A.93}$$

$$= \frac{\sqrt{\pi}\lambda^2}{mk_3^{3/2-i\lambda}}\frac{\sqrt{\lambda}}{\lambda^2-\mu^2}C(x_-)e^{i\delta(\mu,\lambda)} \tag{A.94}$$

where $e^{i\delta(\mu,\lambda)}$ is an overall phase that can be neglected. Notice that the time calculated from stationary phase is identical to that estimated by energy conservation in the locally Minkowski frame, with the mass m replaced by $\mu = \sqrt{m^2-1/4} \approx m$.

A.8.3.1 Squeezed limit

The second integral comes from the three-point vertex, and carries the dependence on the shape of the triangle. There is a sum over both contractions of the external legs with the inflaton

fields in the vertex

$$I^{+}_{k_{12}} = \frac{1}{\sqrt{4k_1^3k_2^3}}\int_{-\infty}^{0}\frac{d\eta}{\eta^4}(-\eta)^{i\lambda}\left\{\left(\left(\frac{\lambda^3\eta}{2}+\lambda\eta^3k_1k_2\right)u_{k_1}u_{k_2}-3i\lambda^2\eta^2u_{k_1}u'_{k_2}-3\lambda\eta^3u'_{k_1}u'_{k_2}\right)\sigma_0\right.$$
$$\left.+\vec{k}_2\cdot\hat{k}_3\left(-\lambda^2\eta^2u_{k_1}u_{k_2}+2i\lambda\eta^3u'_{k_1}u_{k_2}\right)\sigma_1(k_3,\eta)+(k_1\leftrightarrow k_2)\right\}. \tag{A.95}$$

Equation eq. (A.95) is solved by again substituting $x=-k_3\eta$, and the mode functions from eqs. (A.80) and (A.81).

The term multiplying σ_1 is almost antisymmetric in the exchange of k_1 and k_2, $\vec{k}_1\cdot\vec{k}_3=-k_1k_3\cos\theta$ whereas $\vec{k}_2\cdot\vec{k}_3=k_2k_3\cos\theta$. In the squeezed limit there is a slight difference between k_1 and k_2. Using the cosine rule and Taylor expanding

$$k_2=k_1\left(1-\frac{k_3}{k_1}\cos\theta+O\left(\frac{k_3}{k_1}\right)^2\right)\implies k_1-k_2=k_3\cos\theta. \tag{A.96}$$

The second integral can be massaged into a similar form to the first

$$I^{+}_{k_{12}}=\frac{\lambda}{2\sqrt{2}mk_1^3k_2^3k_3^{-3/2+i\lambda}}\int_0^\infty dx\frac{A(x)+B(x)\cos^2\theta}{x^{3/2}(x^2+\mu^2)^{1/4}}e^{ih(x)}, \tag{A.97}$$

where

$$A(x) = -\frac{\lambda^2}{2}-i\lambda^2\frac{px}{2}+\left(\frac{\lambda^2}{2}+3i\lambda-1\right)\left(\frac{px}{2}\right)^2-(6\lambda+2i)\left(\frac{px}{2}\right)^3+4\left(\frac{px}{2}\right)^4 \tag{A.98}$$

$$B(x) = \frac{1}{2}\left(\sqrt{x^2+\mu^2}+\frac{i}{2}\frac{2x^2+\mu^2}{x^2+\mu^2}\right)\left(\lambda\left(1+\frac{ipx}{2}\right)^2+2i\left(\frac{px}{2}\right)^2\right), \tag{A.99}$$

$$h(x) = \lambda\ln x-px-\int^x\sqrt{1+\mu^2/x'^2}dx'. \tag{A.100}$$

where $k_1 = k_2$ was taken in all terms besides those with a factor of $k_1 - k_2$.

The stationary phase approximation in the squeezed limit gives $x_+ \approx (\lambda - \mu)/p$, and $|h''(x_+)| \approx p^2/(\lambda - \mu)$. In this integral the phase depends on p since

$$h(x_+) = \lambda \ln x_+ - p x_+ - \sqrt{x_+^2 + \mu^2} - \mu \ln \left(\frac{x_+}{\mu^2 + \mu \sqrt{x_+^2 + \mu^2}} \right). \tag{A.101}$$

In the squeezed limit every term is approximately independent of p besides the two logs, $(\lambda - \mu) \ln(x_+)$. The result there has a non-analytic dependence coming from

$$e^{ih(x_+)} = p^{i(\mu-\lambda)} e^{i\delta'(\mu,\lambda)}. \tag{A.102}$$

The final result is given by

$$I^+_{k_{12}} = \frac{\sqrt{\pi}\lambda e^{i\delta'(\mu,\lambda)}}{2\sqrt{2} m k_1^3 k_2^3 k_3^{-3/2+i\lambda}} \frac{A(x_+) + B(x_+)\cos^2\theta}{\sqrt{\mu}(\lambda - \mu)} p^{1/2+i(\mu-\lambda)}. \tag{A.103}$$

Combining equations (A.88), (A.94), and (A.103) leads to a result for the three point function

$$I_{-+} \approx \frac{\pi |c^2|}{2\sqrt{2} m^2 k_1^3 k_2^3} p^{1/2+i(\mu-\lambda)} \frac{\lambda^3}{\Lambda^3} \sqrt{\frac{\lambda}{\mu}} \frac{C(x_-)\left[A(x_+) + B(x_+)\cos^2\theta\right]}{(\lambda - \mu)(\lambda^2 - \mu^2)}. \tag{A.104}$$

Equation (A.14) is then used to relate the three point function to the observable quantity, F_{sq}

$$F_{\text{sq}} = f_{\text{oscil}}(c, \mu, \lambda) p^{-5/2+i(\mu-\lambda)} + c.c. \tag{A.105}$$

where the amplitude of oscillations is given by

$$f_{\rm oscil}(c,\mu,\lambda)=\frac{20\pi|c^2|}{3\sqrt{2}m^2}\dot{\phi}_0\frac{\lambda^3}{\Lambda^3}\sqrt{\frac{\lambda}{\mu}}\frac{C(x_-)\left[A(x_+)+B(x_+)\cos^2\theta\right]}{(\lambda-\mu)(\lambda^2-\mu^2)}. \tag{A.106}$$

A.8.4 Cancellation of late-time divergences

In this section it is shown that the apparent late-time divergences in calculations due to the non-derivative inflaton couplings are ultimately absent. It is demonstrated that the leading late-time divergences cancel in the sum over all Feynman diagrams, and that the stationary phase approximation extracts the non-analytic signal relevant for cosmological collider physics.

A.8.4.1 Justifying the stationary phase approximation

The divergent integrals appearing in the NG calculation are of the form

$$I(p)=p^m\int_0^\infty dx x^{-n\pm i(\mu-\lambda)}e^{ipx}e^{-if(x,\mu,\lambda)}, \tag{A.107}$$

where $x=-k\eta$. For $n\geq 1$, the integrand in eq. (A.107) diverges as $x\to 0$. In the vector case, $n=3/2$, while in the scalar case of eq. (A.16), $n=5/2$ or $3/2$. Therefore, it appears that there are late-time divergences in our integrals. Before we demonstrate finiteness, notice that if we differentiate with respect to the squeezing parameter p enough (say 'k' times), the effective n can be made small enough such that the integrand is rendered finite,

$$\frac{\partial^k}{\partial p^k}I(p)<\infty. \tag{A.108}$$

Therefore, the form of $I(p)$ must be

$$I(p) = I_1 + I_2 p + \cdots + I_k p^{k-1} + \text{finite}, \tag{A.109}$$

where the I_i are potentially divergent constants of integration. From this, we learn that the apparent late-time divergences (which eventually cancel with other diagrams) are analytic in the squeezing parameter, p. In cosmological collider physics we are interested in the non-analytic dependence, which only comes from the region at early times near the saddle point of eq. (A.107). The issue of late-time divergences therefore does not affect the calculation of non-analytic dependence in the diagrams. It is therefore justified to only consider the diagrams for which a saddle point exists, or equivalently for which energy conservation can be satisfied at each vertex, and perform the stationary phase approximation, ignoring the region of integration near $x \to 0$.

The reader may still worry about the validity of the model since the analytic part of the NG seems to diverge. Below, we justify why the apparent divergences cancel, at least for the leading non-Boltzmann suppressed contributions.

A.8.4.2 "Integrating out" the vector field

In the following sections we treat the tadpole of the heavy particle perturbatively. Here we show that despite the individual diagrams appearing to be divergent, the leading late-time

divergences must cancel when all diagrams are summed. This is a reflection of the shift symmetry

$$\delta\phi \mapsto \delta\phi + \alpha, \tag{A.110}$$

$$A_\mu \mapsto e^{i\alpha/\Lambda} A_\mu, \tag{A.111}$$

which protects the slow-roll potential of the inflaton.

As seen in the main text, the tree level contribution to the inflaton three-point function is of the form

$$\langle \delta\phi^3 \rangle' = \int_{-\infty}^{0} \frac{d\eta}{\eta^{5/2}} \int_{-\infty}^{0} \frac{d\eta'}{\eta'^{5/2}} f(\eta, \eta'), \tag{A.112}$$

where f is some function satisfying $f(0,0) = \text{const}$, and the integrals may or may not be (anti)time-ordered. These integrals are divergent at late-times, and one may worry about two possibilities:

1. There are regions of integration in which $\eta \gg \eta'$ (or vice versa).

2. An overall divergence in the integration region $\eta \approx \eta' \to 0$.

Since we are treating the tadpole perturbatively all diagrams involve a single internal line of the heavy particle. In the first case the particle is propagating for a very long time and therefore must be on-shell. The vertex at late times cannot then satisfy energy conservation so these will be exponentially suppressed diagrams.

In the region of integration $\eta \approx \eta' \to 0$, the particle is propagating for a short time and is therefore off-shell. In this scenario we have too much energy ($m < \lambda$) to produce the vector on-shell, and its exchange can be integrated out to give a local vertex.

To understand how this works we go to the local Minkowski frame at $\eta \approx \eta' \to 0$. We

consider the Minkowski space Feynman propagator with 4-momentum, k. Due to the λ injection we can say that the energy going through the intermediate line is $k_0 = \lambda + \delta k_0$. In this case there is a well-defined expansion of the heavy off-shell propagator

$$\frac{i}{k^2 - m^2} = \frac{i}{\lambda^2 - m^2 + 2\lambda\delta k_0 + \delta k_0^2 - |\vec{k}|^2} \tag{A.113}$$

$$= \frac{i}{\lambda^2 - m^2}\left[1 - \frac{2\lambda\delta k_0}{\lambda^2 - m^2} + \cdots\right], \tag{A.114}$$

where successive terms are suppressed by $1/(\lambda^2 - m^2)$, rather than the usual $1/m^2$. The contributions to a correlator in the region $\eta \approx \eta' \to 0$ can therefore be reproduced with local effective vertices (polynomial in δk_μ) involving only $\delta\phi$. The symmetry eq. (A.111) is then simply a shift symmetry for $\delta\phi$, leading to manifestly divergence free derivative couplings.

Appendix B: Charged Loops at the Cosmological Collider with Chemical Potential

Arushi Bodas[1,2], Edward Broadberry[3], Raman Sundrum[3], Zhaohui Xu[3]

[1] Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA

[2]Particle Theory Department, Fermilab, Batavia, Illinois 60510, USA

[3]Maryland Center for Fundamental Physics, Department of Physics,

University of Maryland, College Park, MD 20742, USA

Published in the Journal of High Energy Physics 01 (2026) 083,

DOI: https://link.springer.com/article/10.1007/JHEP01(2026)083

Abstract:

Cosmological collider physics allows the detection of heavy particles at inflationary scales through their imprints on primordial non-Gaussianities. We study the chemical potential mechanism applied to a pair of charged scalars. We analytically evaluate the resulting one-loop contribution to the bispectrum, using the spectral decomposition. In this way we are able to determine the parametric dependencies for both the signal and the background. We show that a signal strength $f_{\rm NL} \sim \mathcal{O}(0.01)$ can be obtained within theoretical control, potentially reachable by 21 cm tomography. As an application we consider the colored Higgs bosons in SU(5) supersymmetric orbifold grand unification with masses $M \lesssim 10^{15}\,\text{GeV}$.

B.1 Introduction

Heavy particles produced on-shell during inflation, if they decay into inflatons, could have imprinted themselves on cosmological observables such as the Cosmic Microwave Background (CMB) and Large Scale Structure (LSS) [1, 31, 33, 107–109]. The mechanism, dubbed cosmological collider physics, offers an alternative approach to search for new particles. Compared to terrestrial colliders at the TeV-scale, the "cosmological collider" probes inflationary energy scales, possibly as high as $\sim 10^{14}\,\mathrm{GeV}$. Therefore, with increasing sensitivity from upcoming LSS experiments [26–29, 121, 122] and future 21 cm tomography [111, 123, 124], there is a possibility to detect heavy particles far beyond the reach of upcoming colliders.

In the minimal version of cosmological collider physics particles are created by the expansion of spacetime, characterized by the inflationary Hubble energy scale, H. During inflation, this could be as large as [17]

$$H < 2.5 \times 10^{-5} M_{\mathrm{pl}} \approx 6 \times 10^{13}\,\mathrm{GeV}. \tag{B.1}$$

If the heavy particles can decay into inflatons, they will give the following non-zero contribution to the non-Gaussianity (NG) in the 3-point correlator of the comoving curvature perturbation, $\mathcal{R}$, [1, 31, 108]

$$\frac{\langle \mathcal{R}_{\mathbf{k}_1}\mathcal{R}_{\mathbf{k}_2}\mathcal{R}_{\mathbf{k}_3}\rangle'}{\langle \mathcal{R}_{\mathbf{k}_1}\mathcal{R}_{-\mathbf{k}_1}\rangle'\langle \mathcal{R}_{\mathbf{k}_3}\mathcal{R}_{-\mathbf{k}_3}\rangle'} \sim e^{-\pi\mu}\left(\frac{k_3}{k_1}\right)^{-3/2-i\mu} + \text{c.c.}, \qquad \mu \coloneqq \sqrt{\frac{M^2}{H^2} - \frac{9}{4}}, \tag{B.2}$$

in the squeezed limit, $k_1 \sim k_2 \gg k_3$. Depending on the value of the heavy mass M, this signal

can be either oscillatory ($M > 3H/2$), or a fractional power-law ($M < 3H/2$). Both behaviors, especially the oscillatory one, cannot be mimicked by local contact interactions of inflatons, and can serve as strong evidence for particle production.

Despite the exciting opportunity in this direction, eq. (B.2) contains an exponential suppression $e^{-\pi\mu}$ for $M > 3H/2$, which leads to a very narrow observable mass window for oscillatory signals. Fortunately, this suppression can be overcome through an extension of the minimal mechanism, where the kinetic energy of the inflaton background can give rise to a chemical potential for heavy fields through a dimension-5 coupling [18, 34, 35, 37, 38, 114, 125, 126]:

$$\mathcal{L} \supset -\frac{1}{\Lambda}\nabla_\mu \phi J^\mu \supset \lambda J^0 \tag{B.3}$$

with the chemical potential defined by $\lambda \coloneqq \dot{\phi}_0/\Lambda$, and J^μ is a non-conserved current made from the heavy fields. In order to maintain in effective field theory control, Λ^2 is required to be larger than $\dot{\phi}_0$, i.e. the chemical potential is bounded by $\lambda < \sqrt{\dot{\phi}_0}$ where [17]

$$\dot{\phi}_0 \approx (60H)^2. \tag{B.4}$$

The chemical potential leads to unsuppressed particle production for all fields in J whose masses, $M < \lambda$. This can greatly widen the mass window for oscillatory signatures. The simplest signals of this type, both theoretically and phenomenologically, appear at tree-level, and have been explored in ref. [18, 35].

Loop level processes for cosmological correlators are poorly understood, because we usually lack the theoretical control required to calculate them. However, since the inflatons on the

external legs are neutral, one-loop processes can be important because particles with nontrivial quantum numbers must be produced and annihilated in pairs. Consequently, including one-loop effects can broaden the range of new physics targets to include charged particles and fermions.

There have been several efforts towards computing loops during inflation:[1]

- Rough estimates based on "loop factors × propagators × vertices" have been accomplished in [34, 35, 37, 38, 53, 54]. There are also approximate one-loop calculations based on a late-time expansion of propagators in [34, 55]. Nonetheless, as mentioned in [34], these results cannot predict the extra parametric dependence on the chemical potential, λ, and the particles masses coming from the loop integrals.

- In [56], a numerical calculation of a one-loop process was performed. It is found that the brute-force method is time-consuming because of the high dimensionality and the oscillatory nature of the integrand. Furthermore, the underlying physics is less clear than in an analytic calculation.

- There are also analytic results in [57–63]. In particular, ref. [57, 60] applied the spectral decomposition technique to one-loop corrections to the power spectrum, ref. [63] applied this to multi-loops in the special case of conformally coupled scalar, while ref. [59] further applied this to a particular one-loop contribution to the bispectrum without chemical potential, reducing it to a one-dimensional integral over the invariant mass or a sum over residues. However, most models will also produce triangle diagrams which cannot be captured by the spectral decomposition, a problem which has not so far been addressed. Compared to the numerical approach, this method relies on de Sitter isometry, which is

[1] This is far from an exhaustive list of papers on loops during inflation, here are a few more papers that focus on similar diagrams to the one we have considered [62, 64–67].

broken down to scale invariance by the inflaton background.

In this paper we evaluate the 3-point correlator in a model where the chemical potential effect is perturbative, such that we can still take advantage of the de Sitter isometry. Furthermore, in this particular model all leading order effects can be computed using the spectral representation.

It turns out that in the limit of heavy particles with a large chemical potential, the signal size from the loop process has a surprisingly simple form: for two scalars with masses M_1, M_2, the amplitude for the 3-point function is related to the amplitude that one would obtain from the tree level exchange of a single scalar by

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k_2}} \mathcal{R}_{\mathbf{k_3}} \rangle'_{\text{1-loop,signal}} \simeq \frac{e^{3\pi i/4}}{2\pi^{3/2}} \sqrt{\frac{M_1 M_2}{M_{12}}} \left(\frac{2M_{12}}{\lambda + M_{12}} \frac{2k_3}{k_1} \right)^{-3/2} \times \langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k_2}} \mathcal{R}_{\mathbf{k_3}} \rangle'_{\text{tree}} \big|_{M=M_{12}} \tag{B.5}$$

in the squeezed limit, where $M_{12} \coloneqq M_1 + M_2$ and λ is the chemical potential. A quick derivation of this relation is given in section B.3.3. In addition to an oscillatory signal, the model predicts the existence of a smooth background. We also evaluate the analytic background, showing that it comes from the resonance limit in the corresponding effective tree-level model. By obtaining the correct parametric dependence on both the chemical potential and the masses, we remove one obstacle for future phenomenological studies in this direction.

As a phenomenological application we consider the prospect of detecting the colored Higgs of supersymmetric SU(5) orbifold grand unification. The framework of grand unification is a well motivated scenario for physics beyond the standard model. In GUTs there exist new

particles with masses around the scales that cosmological collider physics can probe [40–42]. Some of these possibilities have been investigated in [53, 54]. However, the simple models of grand unification are in tension with proton decay bounds, which can be avoided in the higher dimensional "orbifold GUTs" [46–48]. The cosmological collider signatures of orbifold GUTs have been investigated in [18, 49], while the idea of detecting supersymmetry in general has been considered in [32, 127]. In ref. [18], the authors also considered the signal from an orbifold SUSY GUT with a chemical potential. However, they looked at tree level exchanges, which require the heavy particle to be neutral. This is not possible at tree level with the simple SU(5) gauge group, but required more exotic grand unification, e.g. trinification. The fact that we can calculate a loop level process allows us to consider the simpler SU(5) gauge group. We find that the one-loop signal could be within the sensitivity of future 21 cm experiments.

The rest of the paper is organized as follows. In section B.2, we introduce our notation and review the necessary background. Section B.3 elaborates on the implementation of a chemical potential and the method of spectral representation, which is crucial for evaluating loop diagrams involving a pair of charged scalars. We also present an intuitive calculation of the non-analytic part of the bispectrum, which constitutes the cosmological collider signal. Sections B.4 and B.5 then provide a more systematic evaluation of both the non-analytic contribution and the smooth analytic background. In section B.6, we analyze theoretical and observational constraints and estimate the maximum signal strength achievable. As a concrete application of our chemical potential model, we assess the detectability of the colored Higgs bosons predicted by the SU(5) orbifold SUSY GUT model in section B.7. We conclude with a discussion of future directions in section B.8. The appendices contain supplementary material relevant to our calculations.

B.2 Preliminaries

In this section, we define our notation and briefly review the in-in formalism for cosmological correlators. The mostly plus signature $(-,+,+,+)$ for the metric is used throughout this paper.

B.2.1 Primordial non-Gaussianity

During slow-roll inflation the spacetime is approximately described by the de Sitter metric:

$$\begin{aligned} \mathrm{d}s^2 &= -\,\mathrm{d}t^2 + e^{2Ht}\,\mathrm{d}\mathbf{x}^2, \\ &= \frac{-\,\mathrm{d}\eta^2 + \mathrm{d}\mathbf{x}^2}{H^2\eta^2}, \end{aligned} \tag{B.6}$$

where t is the proper time, H is the Hubble parameter during inflation and $\eta = -e^{-Ht}/H$ is the conformal time.

Inflation is usually realized by a scalar field ϕ called the "inflaton". It can be decomposed into a homogeneous classical background, ϕ_0, that controls the inflation dynamics, plus a quantum fluctuation, $\delta\phi$, that is responsible for the primordial fluctuations after inflation. In the spatially flat gauge, $\delta\phi$ is related to the gauge-invariant comoving curvature perturbation, $\mathcal{R}$, by [128]

$$\mathcal{R} = -H\frac{\delta\phi}{\dot{\phi}_0} \tag{B.7}$$

with the overdot representing the derivatives with respect to the proper time t. Since the inflationary dynamics is invariant under spatial translation, all momentum-space correlators

contain a total-momentum δ-function:

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \cdots \mathcal{R}_{\mathbf{k}_n} \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \cdots + \mathbf{k}_n) \cdot \langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \cdots \mathcal{R}_{\mathbf{k}_n} \rangle'. \tag{B.8}$$

Starting from the 2-point correlation function, current CMB data favors a nearly scale-invariant primordial power spectrum given by

$$\mathcal{P}_{\mathcal{R}}(k) \coloneqq \langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{-\mathbf{k}} \rangle' = \frac{H^4}{2\dot{\phi}_0^2 k^3} \left(\frac{k}{k_*} \right)^{n_s - 1}. \tag{B.9}$$

The scalar spectral index, n_s, has been measured to be around 0.96 at the reference comoving momentum scale, $k_* \approx 0.05\,\mathrm{Mpc}^{-1}$. In this paper, we assume exact scale invariance during inflation as an approximation, i.e. we will treat H and $\dot{\phi}_0$ as constants. From now on we work in units with $H = 1$ in all equations unless stated otherwise. Note that with these two assumptions, the spacetime scale transformation:

$$\eta, \mathbf{x} \mapsto \lambda\eta, \lambda\mathbf{x} \tag{B.10}$$

remains to be a good symmetry.

While scale invariance forces $\mathcal{P}_{\mathcal{R}}(k)$ to be proportional to $1/2k^3$, higher-point correlation functions, such as the primordial bispectrum $\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle'$, may contain more intricate momentum dependence. The bispectrum is conventionally normalized relative to the size of the

power spectrum in defining the NG as [120]

$$F(k_1,k_2,k_3) \coloneqq \frac{5}{6}\frac{\langle \mathcal{R}_{\mathbf{k}_1}\mathcal{R}_{\mathbf{k}_2}\mathcal{R}_{\mathbf{k}_3}\rangle'}{\mathcal{P}_\mathcal{R}(k_1)\mathcal{P}_\mathcal{R}(k_2)+\mathcal{P}_\mathcal{R}(k_2)\mathcal{P}_\mathcal{R}(k_3)+\mathcal{P}_\mathcal{R}(k_3)\mathcal{P}_\mathcal{R}(k_1)} = -\frac{10}{3}\frac{\dot{\phi}_0}{}\cdot\frac{k_1^3k_2^3k_3^3}{k_1^3+k_2^3+k_3^3}B(k_1,k_2,k_3), \tag{B.11}$$

where $B(k_1,k_2,k_3) = \langle \delta\phi_{\mathbf{k}_1}\delta\phi_{\mathbf{k}_2}\delta\phi_{\mathbf{k}_3}\rangle'$ is the 3-point correlator of the inflaton fluctuations. The current bound from Planck on the 3-point correlator in the equilateral configuration [117] is

$$f_{\rm NL} \coloneqq F(k,k,k) \lesssim \mathcal{O}(10), \tag{B.12}$$

where the detailed value depends on the shape of the NG. Meanwhile, future LSS experiments are expected to be able to probe these signals at $f_{\rm NL}\sim\mathcal{O}(1)$ [26], and more futuristic 21 cm tomography can further bring this precision down to $f_{\rm NL}\sim\mathcal{O}(0.01)$ [111, 123, 124].

The characteristic cosmological collider signal appears in the squeezed limit $k_1\sim k_2\gg k_3$. We define the squeezing parameter p and the angle parameter χ to characterize the shape of the squeezed triangle up to similarity as follows:

$$p \coloneqq \frac{k_1+k_2}{k_3}, \qquad\qquad \chi \coloneqq \frac{k_1-k_2}{k_3}. \tag{B.13}$$

The corresponding momentum triangle is illustrated in figure B.1. In the limit $p\gg 1$, $\chi = \cos\theta_1+\mathcal{O}(p^{-1}) = -\cos\theta_2+\mathcal{O}(p^{-1})$, whereas triangular inequality implies $p\geq 1$ and $|\chi|\leq 1$ regardless of whether the squeezed limit is taken. Under the (p,χ) parametrization, eq. (B.11)

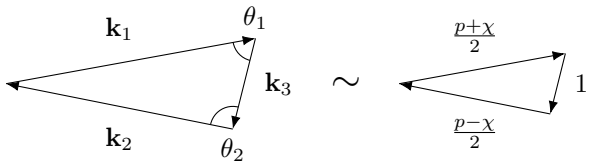


Figure B.1: The shape of the momentum triangle and the two parameters. In the squeezed limit $p \gg 1$, the angle parameter $\chi \simeq \cos\theta_1 \simeq -\cos\theta_2$.

becomes

$$F(k_1,k_2,k_3) = -\frac{10}{3}\dot{\phi}_0 \cdot \frac{4}{p^3+3\chi^2 p+4} \cdot k_1^3 k_2^3 B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$$
$$\xrightarrow{p\to\infty} -\frac{40}{3}\dot{\phi}_0 \cdot p^{-3} \cdot k_1^3 k_2^3 B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3). \tag{B.14}$$

Since the angular dependence does not play a significant role in scalar models, $\chi = 0$ is illustrated in all the plots unless otherwise stated.

B.2.2 In-in formalism and mode functions

In inflationary cosmology, we are mostly interested in the in-in correlators, i.e., the expectation values of operators at a specific time (t_0) in the vacuum state [118]:

$$\langle\Omega_-|\mathcal{O}_H(t_0)|\Omega_-\rangle = \langle 0|\Big(\overline{T}e^{i\int_{-\infty_+}^{t_0} H_{\rm int}(t)\,{\rm d}t}\Big)\mathcal{O}_I(t_0)\Big(Te^{-i\int_{-\infty_-}^{t_0} H_{\rm int}(t)\,{\rm d}t}\Big)|0\rangle, \tag{B.15}$$

where $\infty_\pm := \infty(1\pm i\epsilon)$ contains the $i\epsilon$-prescription, which we leave implicit in all expressions henceforth, and T ($\overline{T}$) represents time (anti-time) ordering. This differs from the usual in-out correlators by having the anti-time ordered exponential. After expanding the two exponentials in power, one can evaluate in-in correlators using ordinary Feynman rules with two sets of vertices from the time ordered exponential (denoted as +) and the anti-time ordered exponen-

tial (denoted as $-$) respectively. Meanwhile, depending on the types of vertices connected in eq. (B.15), there are four sets of propagators:

$$\begin{aligned} G_{++}(t_1,t_2) &\coloneqq \langle 0|T[\phi(t_1)\phi(t_2)]|0\rangle, \\ G_{--}(t_1,t_2) &\coloneqq \big\langle 0|\overline{T}[\phi(t_1)\phi(t_2)]|0\big\rangle, \\ G_{-+}(t_1,t_2) &\coloneqq \langle 0|\phi(t_1)\phi(t_2)|0\rangle, \\ G_{+-}(t_1,t_2) &\coloneqq \langle 0|\phi(t_2)\phi(t_1)|0\rangle. \end{aligned} \tag{B.16}$$

The G_{++} propagator is identical to the Feynman propagator in the usual in-out formalism, while other propagators are either its hermitian conjugate (G_{--}) or can be obtained from its analytic continuation (G_{-+} and G_{+-}).

In de Sitter spacetime it is simpler to work with conformal time, η, and comoving spatial momentum. To calculate the matrix elements in eq. (B.16) for scalar fields, we represent the scalar field operator in terms of creation and annihilation operators as

$$\phi(\mathbf{k},\eta) = f_k(\eta)a_{\mathbf{k}} + \bar{f}_k(\eta)a^\dagger_{-\mathbf{k}}, \qquad\qquad [a_{\mathbf{k}'}, a^\dagger_{\mathbf{k}}] = (2\pi)^3\delta^3(\mathbf{k}'-\mathbf{k}), \tag{B.17}$$

where the mode function $f_k(\eta)$ is a normalized solution to the free field equation

$$\begin{cases} \left(\partial_\eta^2 - \dfrac{2}{\eta}\partial_\eta + k^2 + \dfrac{M^2}{\eta^2}\right) f_k = 0, \\ \qquad -i\big(f_k\partial_\eta\bar{f}_k - \bar{f}_k\partial_\eta f_k\big) = \eta^2. \end{cases} \tag{B.18}$$

With this representation, the four propagators in eq. (B.16) are given by

$$\begin{aligned}
G_{++}(k,\eta_1,\eta_2) &\coloneqq f_k(\eta_1)\bar{f}_k(\eta_2)\theta(\eta_1-\eta_2)+\bar{f}_k(\eta_1)f_k(\eta_2)\theta(\eta_2-\eta_1),\\
G_{--}(k,\eta_1,\eta_2) &\coloneqq \bar{f}_k(\eta_1)f_k(\eta_2)\theta(\eta_1-\eta_2)+f_k(\eta_1)\bar{f}_k(\eta_2)\theta(\eta_2-\eta_1),\\
G_{-+}(k,\eta_1,\eta_2) &\coloneqq f_k(\eta_1)\bar{f}_k(\eta_2),\\
G_{+-}(k,\eta_1,\eta_2) &\coloneqq \bar{f}_k(\eta_1)f_k(\eta_2),
\end{aligned} \tag{B.19}$$

where θ is the Heaviside step function. To fix the solution, we must further impose the Bunch-Davies initial condition at $\eta \to -\infty$:

$$f_k(\eta) \sim e^{-ik\eta}, \qquad \eta \to -\infty. \tag{B.20}$$

This choice ensures that the spacetime contains no particles in the infinite past. The solution to eq. (B.18) is then given by

$$f_k(\eta) = \sqrt{\frac{\pi}{4}} e^{\pi i/4 - \pi\mu/2} (-\eta)^{3/2} H^{(1)}_{i\mu}(-k\eta), \tag{B.21}$$

where $H^{(1)}$ is the Hankel function of first kind and μ is the de Sitter mass parameter defined by

$$i\mu \coloneqq \begin{cases} i\sqrt{M^2 - \dfrac{9}{4}}, & M > 3/2, \\ \sqrt{\dfrac{9}{4} - M^2}, & M < 3/2. \end{cases} \tag{B.22}$$

In the special case of the inflaton, we have $M = 0$ and

$$f_k(\eta) = \frac{i}{\sqrt{2k^3}}(1 + ik\eta)e^{-ik\eta}. \tag{B.23}$$

The primordial fluctuations are measured at the end of inflation, i.e., $\eta \to 0$, so we also need the bulk-to-boundary propagators for the inflaton:

$$D_\pm(k, \eta) \coloneqq G_{+\pm}(k, 0, \eta) = \frac{1}{2k^3}(1 \mp ik\eta)e^{\pm ik\eta}. \tag{B.24}$$

In particular, the leading-order inflaton power spectrum is given by

$$\mathcal{P}_\phi(k) \coloneqq \langle \delta\phi_{\mathbf{k}} \delta\phi_{-\mathbf{k}} \rangle' = D_+(k, 0) = \frac{1}{2k^3}. \tag{B.25}$$

B.3 Charged-pair Chemical Potential Model and Spectral Decomposition

In this section we construct a model for a pair of complex scalars (χ_1 and χ_2) with a chemical potential, and opposite charges under all symmetry groups. We then derive the Källén-Lehmann representation for the 3-point correlation function, and discuss some features of the spectral density.

B.3.1 Chemical potential for a pair of charged scalars

For two complex scalar fields, χ_1, and, χ_2, a chemical potential-like coupling is achieved by coupling their U(1) currents to the derivative of the inflaton field, ϕ, as follows [35]:

$$\mathcal{L} \supset -\sum_{I=1,2} \frac{1}{2\Lambda} \nabla_\mu \phi J_I^\mu, \qquad\qquad J_I^\mu := i(\bar{\chi}_I \nabla^\mu \chi_I - \chi_I \nabla^\mu \bar{\chi}_I). \tag{B.26}$$

After plugging in the classical background, ϕ_0, the Lagrangian density contains a term $\frac{\lambda}{2} J_0$ for each field, giving a chemical potential difference of

$$\lambda := \frac{\dot{\phi}_0}{\Lambda}, \tag{B.27}$$

between χ_1, χ_2 and their conjugates $\bar{\chi}_1$, $\bar{\chi}_2$. An alternative way to write the chemical potential of eq. (B.26) is through the following Lagrangian:

$$\begin{aligned} \mathcal{L}_\chi \supset & -\sum_{I=1,2} \left[\left(\nabla^\mu + i \frac{\nabla^\mu \phi}{2\Lambda} \right) \bar{\chi}_I \left(\nabla_\mu - i \frac{\nabla_\mu \phi}{2\Lambda} \right) \chi_I + M_I^2 \left| \chi_I \right|^2 \right] \\ = & -\sum_{I=1,2} \left[\left| \nabla \chi_I \right|^2 + M_I^2 \left| \chi_I \right|^2 + \frac{1}{2\Lambda} \nabla_\mu \phi J_I^\mu + \frac{1}{4\Lambda^2} (\nabla \phi)^2 \left| \chi_I \right|^2 \right], \end{aligned} \tag{B.28}$$

which gives the same dimension-5 interaction up to a dimension-6 coupling.

In the form of eq. (B.28), it is clear that the chemical potential has no nontrivial effect if the full theory is invariant under the following $\mathrm{U}(1)_A$ symmetry:

$$\mathrm{U}(1)_A : \quad \chi_1 \to \chi_1 e^{i\theta}, \qquad \chi_2 \to \chi_2 e^{i\theta}. \tag{B.29}$$

In this case a field redefinition $\chi_I \to \chi_I e^{i\phi/2\Lambda}$ could eliminate the chemical potential term, while leaving any non-derivative couplings of χ_I invariant. Absent any further symmetry considerations, the lowest dimensional symmetry-breaking term would be a linear term in χ_I, leading to the model discussed in ref. [35], where a heavy particle is exchanged at tree-level. In this paper we will focus on charged scalars, where we assume that χ_1 and χ_2 have opposite charges under all other symmetries. For example if the scalars are also charged under some other $\mathrm{U}(1)_V$, they should transform under its action as[2]

$$\mathrm{U}(1)_V: \quad \chi_1 \to \chi_1 e^{i\theta}, \qquad \chi_2 \to \chi_2 e^{-i\theta}. \tag{B.30}$$

In this case, the lowest dimensional term that breaks $\mathrm{U}(1)_A$ is given by $\chi_1\chi_2$ and we arrive at the following Lagrangian:

$$\mathcal{L}_\chi = -\sum_{I=1,2}\left[|\nabla\chi_I|^2 + M_I^2|\chi_I|^2 + \underbrace{\frac{1}{2\Lambda}\nabla_\mu\phi J_I^\mu}_{\text{chemical potential}} + \frac{1}{4\Lambda^2}(\nabla\phi)^2|\chi_I|^2\right] - \underbrace{\alpha\chi_1\chi_2 - \bar{\alpha}\bar{\chi}_1\bar{\chi}_2}_{\text{explicit } \cancel{\mathrm{U}(1)}_A}. \tag{B.31}$$

The flatness of the inflaton potential is preserved due to the derivative ϕ couplings.

The role of the explicit $\mathrm{U}(1)_A$ breaking term is more apparent after the field redefinition

[2]One can also build a model with $\chi_1 = \chi_2 = \chi$ as a single complex scalar. In this case the same argument applies if χ is odd under a $\mathbb{Z}_2$ symmetry, or it transforms under a nontrivial real representation of some symmetry group. Equivalently, one can start with two real scalars with the same set of symmetries, in which the $\mathrm{U}(1)_A$ breaking term $\alpha\chi^2$ naturally comes from the mass splitting between the two real scalars.

$\chi_I \mapsto \chi_I e^{i\phi/2\Lambda}$, where the Lagrangian in eq. (B.31) becomes [3]

$$\begin{aligned}\mathcal{L}_\chi &= -\sum_{I=1,2}\left(|\nabla\chi_I|^2 + M_I^2\,|\chi_I|^2\right) - \alpha\chi_1\chi_2 e^{i\phi/\Lambda} - \bar{\alpha}\bar{\chi}_1\bar{\chi}_2 e^{-i\phi/\Lambda},\\ &= -\sum_{I=1,2}\left(|\nabla\chi_I|^2 + M_I^2\,|\chi_I|^2\right) - \alpha\chi_1\chi_2 e^{i\lambda t}e^{i\delta\phi/\Lambda} - \bar{\alpha}\bar{\chi}_1\bar{\chi}_2 e^{-i\lambda t}e^{-i\delta\phi/\Lambda}.\end{aligned} \tag{B.32}$$

The $\mathrm{U}(1)_A$ breaking term effectively acts as an external source that injects (extracts) energy of order λ into (from) the system and creates (annihilates) $\chi_1\chi_2$ pairs. The Lagrangian in this basis still contains a shift symmetry for the inflaton:

$$\phi \to \phi + c, \qquad \chi_I \to \chi_I e^{ic/2\Lambda}. \tag{B.33}$$

This symmetry plays the same role as the original shift symmetry in preserving the flatness of the inflaton potential.

B.3.2 One-loop diagram under spectral representation

The time-dependent coupling in eq. (B.32) explicitly breaks de Sitter covariance down to scale invariance. However, the breaking effects will remain in control if the mixing, α, can be treated perturbatively. Since $\chi_1\chi_2$ are created and annihilated in pairs, the leading-order contribution to the bispectrum is $\mathcal{O}(\alpha^2)$ at one-loop, as illustrated in figure B.2. In the in-in

[3] One could have started with χ_I having different charges. This does not make any difference if only $\chi_1\chi_2$ term is considered. When one consider the $\chi_1\bar{\chi}_2$ term together with uneven charges, one can write down a model with two sources injecting different energies.

formalism, it is given by

$$\langle \delta\phi_{\mathbf{k}_1}\delta\phi_{\mathbf{k}_2}\delta\phi_{\mathbf{k}_3}\rangle = |\alpha|^2 \sum_{s_1,s_2=\pm}(-is_1)(-is_2)\int\sqrt{-g_1}\,\mathrm{d}^4x_1\int\sqrt{-g_2}\,\mathrm{d}^4x_2(-\eta_1)^{-i\lambda}(-\eta_2)^{i\lambda}$$
$$\times\left\langle \delta\phi_{\mathbf{k}_1}\delta\phi_{\mathbf{k}_2}\delta\phi_{\mathbf{k}_3}\cdot e^{i\delta\phi(x_1)/\Lambda}e^{-i\delta\phi(x_2)/\Lambda}\right\rangle_{s_1s_2}\left\langle\mathcal{O}(x_1)\overline{\mathcal{O}}(x_2)\right\rangle_{s_1s_2}, \tag{B.34}$$

where $\mathcal{O}=\chi_1\chi_2$ and the two signs $s_1,s_2=+/-$ indicate whether the vertex comes from the time or anti-time ordered exponential. At leading order, the two correlators of $\delta\phi$ and χ_I can be evaluated as in the free field theory, where the second correlator gives rise to a loop in momentum space:

$$\left\langle\mathcal{O}(\mathbf{k},\eta_1)\overline{\mathcal{O}}(\mathbf{k},\eta_2)\right\rangle'_{s_1s_2} = \int\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}G_{s_1s_2}(|\mathbf{q}|,\eta_1,\eta_2;\mu_1)G_{s_1s_2}(|\mathbf{k}-\mathbf{q}|,\eta_1,\eta_2;\mu_2), \tag{B.35}$$

with $G_{s_1s_2}$ defined in eqs. (B.19) and (B.21). Note that at $\mathcal{O}(\alpha^2)$, there is only two insertions of the exponentials $e^{\pm i\delta\phi(x_1)/\Lambda}$ with which $\delta\phi_{\mathbf{k}}$ can contract, as shown in figure B.3. In particular, triangle diagrams do not appear at $\mathcal{O}(\alpha^2)$ and are thus higher order in the α-expansion.

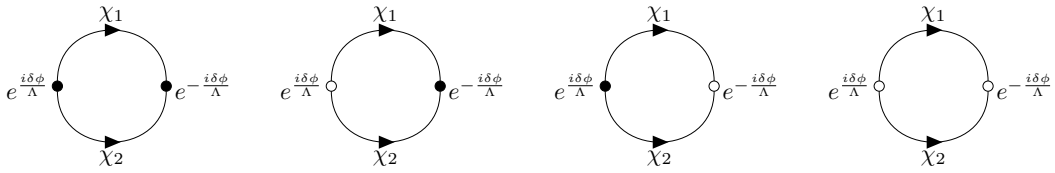
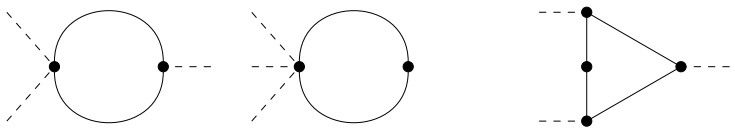


Figure B.2: The 1PI one-loop subdiagrams contributing at $\mathcal{O}(\alpha^2)$. The solid dots, •, represents (+)-type vertices, while the empty dots, ∘, represent (−)-type vertices.

Loop integrations such as the one in eq. (B.35) are, in general, hard to evaluate either analytically or numerically. Nevertheless, since the propagators appearing in eq. (B.35) are free propagators respecting the de Sitter isometries, it can be simplified through the spectral repre-



$\mathcal{O}(\alpha^2)$ $\mathcal{O}(\alpha^4)$

Figure B.3: (Left) At $\mathcal{O}(\alpha^2)$, there are only two types of topologies contributing to the 3-point function and both of them admit the spectral density representation. (Right) Triangle diagram contributions first appear at $\mathcal{O}(\alpha^4)$. These contributions are subdominant though cannot be represented by a single spectral integral.

sentation. The idea is similar to the Källén–Lehmann representation in Minkowski spacetime: to calculate the matrix element $\langle \mathcal{O}\overline{\mathcal{O}} \rangle_{-+} = \langle 0|\mathcal{O}\overline{\mathcal{O}}|0\rangle$, we first insert a complete set of states to get

$$\langle \mathcal{O}(\mathbf{k},\eta_1)\overline{\mathcal{O}}(\mathbf{k},\eta_2)\rangle_{-+} = \sum\!\!\!\!\!\!\int \mathrm{d}n \int \frac{\mathrm{d}^3\mathbf{k}'}{(2\pi)^3} \langle 0|\mathcal{O}(\mathbf{k},\eta_1)|\mathbf{k}',\mu_n\rangle \overline{\langle 0|\mathcal{O}(\mathbf{k},\eta_2)|\mathbf{k}',\mu_n\rangle}, \tag{B.36}$$

where $|\mathbf{k},\mu_n\rangle$ is a state with momentum, $\mathbf{k}$, and a mass parameter, μ_n, that can either be a single-particle state or a multi-particle state with a definite invariant mass. The de Sitter isometry then implies that the matrix element has a unique form up to a multiplicative constant:[4]

$$\langle 0|\mathcal{O}(\mathbf{k}',\eta)|\mathbf{k},\mu_n\rangle = \langle 0|\mathcal{O}|\mu_n\rangle \cdot f_k(\eta;\mu_n) \cdot (2\pi)^3\delta^3(\mathbf{k}'-\mathbf{k}), \tag{B.37}$$

where $\langle 0|\mathcal{O}|\mu_n\rangle$ is the reduced matrix element and $f_k(\eta;\mu)$ is defined by eq. (B.21). Plugging

[4]The easiest way to derive this is to consider the de Sitter Casimir operator C_1, which has a constant value $\mu^2+9/4$ over irreducible representations, while acting on scalar operators as the d'Alembert operator:

$$[C_1,\mathcal{O}(x)] = \Box\mathcal{O}(x) = \eta^2\left(-\partial_\eta^2 + \frac{2}{\eta}\partial_\eta + \nabla^2\right)\mathcal{O}(x).$$

This implies that the matrix element $\langle 0|\mathcal{O}(\mathbf{k}',\eta)|\mathbf{k},\mu_n\rangle$ satisfies the differential equation eq. (B.18) with $M^2 = \mu_n^2+9/4$.

this into eq. (B.36) gives

$$\begin{aligned}\left\langle \mathcal{O}(\mathbf{k},\eta_1)\overline{\mathcal{O}}(\mathbf{k},\eta_2)\right\rangle'_{-+} =& \sum\!\!\!\!\!\!\int |\langle 0|\mathcal{O}|\mu_n\rangle|^2\, G_{-+}(k,\eta_1,\eta_2;\mu_n)\,\mathrm{d}n \\ =:& \int_0^\infty \rho^{\mathrm{dS}}_{\mathcal{O}}(\mu) G_{-+}(k,\eta_1,\eta_2;\mu)\,\mathrm{d}\mu, \end{aligned} \tag{B.38}$$

where

$$\rho^{\mathrm{dS}}_{\mathcal{O}}(\mu) := \sum\!\!\!\!\!\!\int |\langle 0|\mathcal{O}|\mu_n\rangle|^2\, \delta(\mu_n-\mu)\,\mathrm{d}n \tag{B.39}$$

is the de Sitter spectral density of $\mathcal{O}$. Equivalently:[5]

$$\left\langle \mathcal{O}(x_1)\overline{\mathcal{O}}(x_2)\right\rangle_{s_1 s_2} = \int_0^\infty \rho^{\mathrm{dS}}_{\mathcal{O}}(\mu)\langle \chi_\mu(x_1)\bar{\chi}_\mu(x_2)\rangle_{s_1 s_2}\,\mathrm{d}\mu, \tag{B.40}$$

where χ_μ is a fictitious scalar with mass parameter μ. We then arrive at the following relationship between the one-loop correlator and a fictitious tree-level correlator:

$$B_{\text{1-loop}}(k_1,k_2,k_3) = \int_0^\infty \rho^{\mathrm{dS}}_{\mu_1\mu_2}(\mu) B_{\mathrm{tree}}(k_1,k_2,k_3;\mu)\,\mathrm{d}\mu, \tag{B.41}$$

This is shown diagrammatically as in figure B.4.

The corresponding de Sitter spectral density for $\mathcal{O}=\chi_1\chi_2$ has been evaluated and has a

[5] This form of spectral decomposition is only correct when either at least one scalar has $M_I \geq 3/2$, or if both scalars have $M_I < 3/2$ then they must satisfy

$$\sqrt{\frac{9}{4}-M_1^2}+\sqrt{\frac{9}{4}-M_2^2} \leq \frac{3}{2}.$$

For these values of M_I, the multi-particle states all have $M \geq 3/2$.

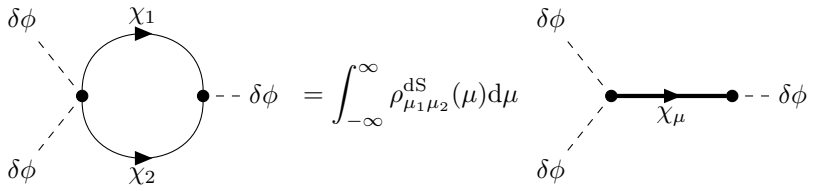


Figure B.4: The spectral representation of the 3-point correlator at one-loop. This diagram only shows one of the ways $\delta\phi$ can contract with $e^{\pm i\delta\phi/\Lambda}$.

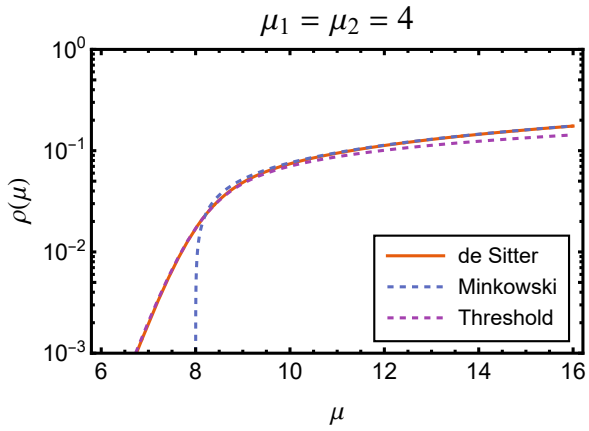


Figure B.5: The de Sitter spectral density eq. (B.42) compared with the Minkowski spectral density eq. (B.44), and the near-threshold approximation eq. (B.46), for $\mu_1 = \mu_2 = 4$.

surprisingly simple form [68–70]:[6]

$$\rho^{\rm dS}_{\mu_1\mu_2}(\mu) = \frac{1}{8\pi^4}\frac{\mu\sinh\pi\mu}{\pi}\frac{\prod_{\epsilon,\epsilon_1,\epsilon_2=\pm}\Gamma\left[\frac{3}{4}+\frac{i}{2}\epsilon(\mu+\epsilon_1\mu_1+\epsilon_2\mu_2)\right]}{\Gamma\left(\frac{3}{2}-i\mu\right)\Gamma\left(\frac{3}{2}+i\mu\right)}. \tag{B.42}$$

In this paper we will not go through the derivation of this expression, but only discuss some of its features:

- At small distances the spacetime curvature is unimportant and we should expect that eq. (B.42) recovers its Minkowski form. This is achieved by taking the limit $\mu, \mu_1, \mu_2 \gg 1$, where all gamma functions in eq. (B.42) can be replaced by their Stirling approximation. One obtains

$$\rho^{\rm dS}_{\mu_1\mu_2}(\mu) \simeq \frac{1}{8\pi^2\mu}\prod_{\epsilon_1,\epsilon_2=\pm}\sqrt{|\mu+\epsilon_1\mu_1+\epsilon_2\mu_2|} \times \begin{cases} 1, & \mu > \mu_{12}, \\ e^{-\pi(\mu_{12}-\mu)}, & |\mu_1-\mu_2| < \mu < \mu_{12}, \\ e^{-2\pi(\max\{\mu_1,\mu_2\}-\mu)}, & \mu < |\mu_1-\mu_2|\,, \end{cases} \tag{B.43}$$

where $\mu_{12} \coloneqq \mu_1 + \mu_2$. For comparison, we quote the Minkowski spectral density for

[6] We are also aware of another expression for the spectral density in terms of the hypergeometric function ${}_7F_6$ in [57, 59]. It can be numerically verified that our spectral density is related to the spectral density in [59], denoted as $\mathrm{P}^{\rm dS}_{\mu_0}(\mu)$, through the following equation:

$$\rho^{\rm dS}_{\mu_0\mu_0}(\mu) = \frac{2\mu}{\pi}\,\mathrm{Im}\,\mathrm{P}^{\rm dS}_{\mu_0}(\mu).$$

$\mathcal{O} = \chi_1\chi_2$:

$$\rho^{\text{Mink}}_{M_1M_2}(M) = \frac{1}{8\pi^2 M}\prod_{\epsilon_1,\epsilon_2=\pm}\sqrt{M+\epsilon_1 M_1+\epsilon_2 M_2}\cdot\theta(M-M_1-M_2). \tag{B.44}$$

The comparison between the two spectral densities is presented in figure B.5. For $\mu_1, \mu_2 \gg 1$, the two spectral densities are close for $\mu - \mu_1 - \mu_2 \gg 1$, while they have major difference for $\mu < \mu_1 + \mu_2$. In particular, the Minkowski spectral density eq. (B.44) has a sharp cutoff at $\mu = \mu_1 + \mu_2$, whereas the de Sitter spectral density decays exponentially for $\mu < \mu_1 + \mu_2$.

- As will be shown in the next section, the oscillatory signal is determined by the threshold limit of the spectral density, i.e. $\mu = \mu_{12} + \varepsilon$ with $\varepsilon = \mathcal{O}(1) \ll \mu_1, \mu_2$. In this limit all gamma functions in eq. (B.42) can be approximated except

$$\Gamma\left[\frac{3}{4} \pm \frac{i}{2}(\mu - \mu_{12})\right] = \Gamma\left(\frac{3}{4} \pm \frac{i\varepsilon}{2}\right), \tag{B.45}$$

giving

$$\rho^{\text{dS}}_{\mu_1\mu_2}(\mu_{12}+\varepsilon) \simeq \frac{1}{4\pi^3}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}}\Gamma\left(\frac{3}{4}+\frac{i\varepsilon}{2}\right)\Gamma\left(\frac{3}{4}-\frac{i\varepsilon}{2}\right)e^{\pi\varepsilon/2}. \tag{B.46}$$

This form of the spectral density is plotted in figure B.5, showing good agreement with the general form for small values of μ. It also differs from the flat space case eq. (B.44), where one has

$$\rho^{\text{Mink}}_{\mu_1\mu_2}(M_{12}+E) \simeq \frac{1}{4\pi^2}\sqrt{\frac{M_1M_2}{M_{12}}}\sqrt{2E}\cdot\theta(E). \tag{B.47}$$

This difference is important in our chemical potential model because we will show later that the proper time elapsed from the creation to the annihilation of the heavy particle(s) is

larger than H^{-1} in the squeezed limit. Consequently, the particle(s) can see the curvature in the time direction and distinguish de Sitter from flat spacetime. Eq. (B.46) can be better understood by deriving it from the non-relativistic limit of de Sitter spacetime. Since this deviates from the main purpose of this paper, we leave the discussion in appendix B.9.3.

- As a function of complex μ, $\rho^{\rm dS}_{\mu_1\mu_2}(\mu)$ is meromorphic, and only has contains simple poles at

$$\mu = \pm\mu_1 \pm \mu_2 \pm \left(\frac{3}{2} + 2\mathbb{N}\right) i. \tag{B.48}$$

Meanwhile, the late-time behavior of the mode function eq. (B.21) is given by

$$f_k(\eta;\mu) = \sum_{k=0}^{\infty}\Big[A_k(-\eta)^{\Delta^+ +2k} + B_k(-\eta)^{\Delta^- +2k}\Big], \qquad \begin{cases} \Delta^+ := \dfrac{3}{2} + i\mu, \\ \Delta^- := \dfrac{3}{2} - i\mu. \end{cases} \tag{B.49}$$

Consequently, poles in the upper half plane correspond to the μ values when $\Delta^+ = \Delta_1^\pm + \Delta_2^\pm$. From the dS/CFT perspective, these poles can further be attributed to different late-time operators appearing in the operator product expansion of $\chi_1\chi_2$ [70].

B.3.3 Central result for the signal

A systematic calculation of all contributions to the one-loop bispectrum will be carried out in the next section. However, in this section we give a quick derivation of the non-analytic piece in the squeezed limit. This piece is the oscillatory signal in cosmological collider physics.

Our starting point will be the tree-level amplitude in (B.41), which is given by

$$B_{\rm tree}(k_1,k_2,k_3;\mu) = |\alpha|^2 \sum_{s_1,s_2}(is_1)(is_2)\int\sqrt{-g_1}\,\mathrm{d}^4x_1\int\sqrt{-g_2}\,\mathrm{d}^4x_2(-\eta_1)^{-i\lambda}(-\eta_2)^{i\lambda}$$
$$\times\left\langle\delta\phi_{\mathbf{k}_1}\delta\phi_{\mathbf{k}_2}\delta\phi_{\mathbf{k}_3}\cdot e^{i\delta\phi(x_1)/\Lambda}e^{-i\delta\phi(x_2)/\Lambda}\right\rangle_{s_1s_2}\langle\chi_\mu(x_1)\bar{\chi}_\mu(x_2)\rangle_{s_1s_2}. \tag{B.50}$$

The amplitude for the fictitious scalar has been computed in ref. [35] as part of a tree-level model. The dominant contribution comes from the $(-+)$ diagram.

In the left panel of figure B.6 we plot the integrand in eq. (B.41) for the $(-+)$ diagram. The integrand has the form of an oscillatory function multiplied by a smooth function. In

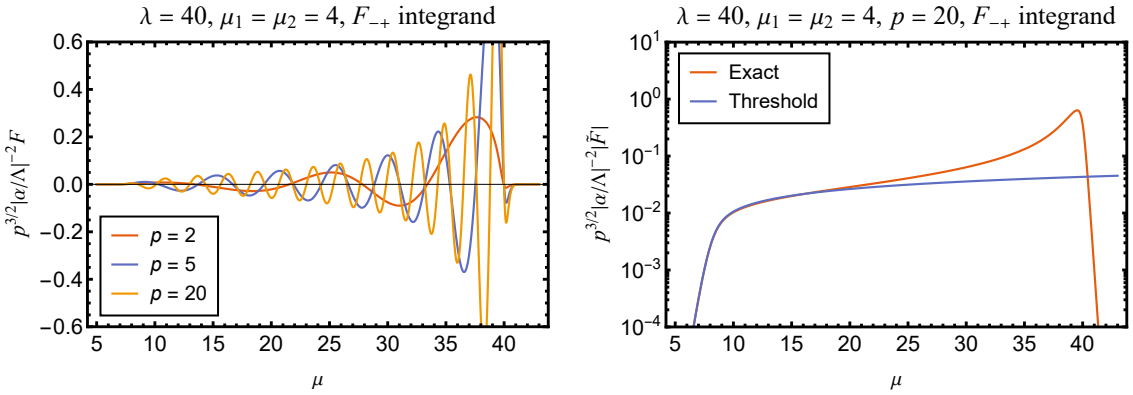


Figure B.6: (Left) The integrand in eq. (B.41) for $\lambda = 40$, $\mu_1 = \mu_2 = 4$ and various p. The normalized bispectrum eq. (B.11) is used instead as the y-axis when making the plot. (Right) The magnitude of the integrand for $p = 20$. The orange line represents the exact integrand while the blue line represents the modified integrand to extract the near-threshold approximation in eq. (B.55).

particular, as we will see shortly, the phase does not have a stationary point. However, the smooth envelope changes rapidly around $\mu = \lambda$, and $\mu = \mu_{12}$. Near $\mu = \lambda$, the tree-level amplitude $B_{\rm tree}$ peaks and then decays exponentially once $\mu > \lambda$. Near $\mu = \mu_{12}$, the spectral density also decays exponentially for $\mu < \mu_{12}$. The integral is therefore dominated by the two regions where $\mu \simeq \mu_{12}, \lambda$.

For values of the mass parameter, $\mu \lesssim \lambda$, the squeezed limit tree level amplitude can be approximately calculated in the Stationary Phase Approximation (SPA), and it is given by [35]

$$B_{\text{tree,SPA}}(k_1, k_2, k_3; \mu) = \frac{\pi}{16 k_1^3 k_2^3} \frac{|\alpha|^2}{\Lambda^3} \frac{\lambda^{1/2} \mu^{-1/2}}{\lambda^2 - \mu^2} p^{3/2 - i(\lambda - \mu)} e^{i \vartheta_*(\lambda, \mu)} + \text{c.c.}, \tag{B.51}$$

where the phase ϑ_* is given by[7]

$$\vartheta_*(\lambda, \mu) \coloneqq \lambda \log 2\lambda + \mu \log 2\mu - (\lambda + \mu) \log(\lambda + \mu). \tag{B.52}$$

Plugging this into eq. (B.41), we obtain

$$\begin{aligned} B_{\text{1-loop}}(k_1, k_2, k_3) &\sim \int_0^\infty \rho^{\text{dS}}_{\mu_1 \mu_2}(\mu) B^{(\text{SPA})}_{\text{tree}}(k_1, k_2, k_3; \mu) \\ &\sim \frac{\pi}{16 k_1^3 k_2^3} \frac{|\alpha|^2}{\Lambda^3} \int_0^\lambda \frac{\lambda^{1/2} \mu^{-1/2}}{\lambda^2 - \mu^2} p^{3/2 - i(\lambda - \mu)} e^{i \vartheta_*(\lambda, \mu)} \rho^{\text{dS}}_{\mu_1 \mu_2}(\mu) \, \mathrm{d}\mu \,. \end{aligned} \tag{B.53}$$

Note that the superficial divergence when $\mu \to \lambda$ is an artifact of SPA and does not exist in the exact expression.

The full phase[8]

$$\vartheta(\lambda, \mu, p) \coloneqq -(\lambda - \mu) \log p + \vartheta_*(\lambda, \mu), \qquad \frac{\partial \vartheta}{\partial \mu} = \log \frac{2\mu p}{\lambda + \mu} > 0 \tag{B.54}$$

[7] Ref. [35] does not present an expression for this phase as it is unimportant for their computation. We will derive it in section B.4.2.

[8] Even for moderate p, ϑ is given by eq. (B.84) and its derivative is given by

$$\frac{\partial \vartheta}{\partial \mu} = \frac{1}{2} \log \frac{\lambda + \mu - p x_*}{\lambda - \mu - p x_*} - \log \frac{\lambda + \mu}{\lambda - \mu} > 0,$$

where $x_* = x_*^-$ is given by eq. (B.81). The latter is clearly positive-definite.

does not, as we mentioned before, have a stationary point on its own. Inspection of eq. (B.53) tells us that the non-analyticity is proportional to $p^{-i(\lambda-\mu)}$. Thus, only the $\mu \to \mu_{12}$ limit contribution has oscillating behavior as a function of p.

To isolate this limit, we consider a modified integral:

$$B^{(\mu_{12})}_{\text{1-loop}}(k_1,k_2,k_3) \simeq B^{(\text{SPA})}_{\text{tree}}(k_1,k_2,k_3;\mu_{12}) \int_{-\infty}^{\infty+} \rho^{\text{dS}}_{\mu_1\mu_2}(\mu_{12}+\varepsilon)\beta^{i\varepsilon}\,\mathrm{d}\varepsilon, \tag{B.55}$$

$$\beta \coloneqq \exp\left(\frac{\partial\vartheta}{\partial\mu}\right)\bigg|_{\mu=\mu_{12}} = \frac{2\mu_{12}p}{\lambda+\mu_{12}}. \tag{B.56}$$

As shown in the right panel of figure B.6, this new integral leaves the envelope of the integrand near $\mu \simeq \mu_{12}$ unmodified, but also lets it continues to oscillate beyond $\mu = \lambda$, eliminating the contribution near that limit. To simplify the calculation, we have also expanded the full phase eq. (B.54) to linear order around $\mu = \mu_{12}$.

We can now further approximate $\rho^{\text{dS}}_{\mu_1\mu_2}(\mu)$ by eq. (B.46) and obtain an approximate form for the signal

$$\begin{aligned} B^{(\mu_{12})}_{\text{1-loop,sq}}(k_1,k_2,k_3) &\simeq \frac{1}{2\pi^{3/2}}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}}\left(\frac{2\mu_{12}p}{\lambda+\mu_{12}}\right)^{-3/2} e^{3\pi i/4} B^{(\text{SPA})}_{\text{tree}}(k_1,k_2,k_3;\mu_{12}) \\ &= \frac{1}{32\sqrt{\pi}k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda^{1/2}\mu_{12}^{-1/2}}{\lambda^2-\mu_{12}^2}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}} \\ &\quad\times\left(\frac{2\mu_{12}}{\lambda+\mu_{12}}\right)^{-3/2} p^{-i(\lambda-\mu_{12})} e^{i\vartheta_*(\lambda,\mu_{12})+3\pi i/4}. \end{aligned} \tag{B.57}$$

Here we have used the Gamma function integral identity in eq. (B.184) to perform the integral over ϵ. Matching to the bispectrum parametrized by eq. (B.11), the non-analytic piece of the

signal is given by

$$F^{(\mu_{12})}_{\text{1-loop,sq}}(k_1,k_2,k_3) = \widetilde{F}^{(\mu_{12})}_{\text{1-loop,sq}}(k_1,k_2,k_3) + \text{c.c.}, \tag{B.58}$$

$$\widetilde{F}^{(\mu_{12})}_{\text{1-loop,sq}}(k_1,k_2,k_3) \simeq f_{\text{oscil}}(\lambda,\mu_1,\mu_2)p^{-3-i(\lambda-\mu_{12})}, \tag{B.59}$$

where the function, f_{oscil}, is given by

$$f_{\text{oscil}} = -\frac{5}{12\sqrt{\pi}}\frac{|\alpha|^2}{\Lambda^2}\frac{\lambda^{3/2}\mu_{12}^{-1/2}}{\lambda^2-\mu_{12}^2}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}}\left(\frac{2\mu_{12}}{\lambda+\mu_{12}}\right)^{-3/2} e^{i\vartheta_*(\lambda,\mu_{12})+3\pi i/4}. \tag{B.60}$$

Before going through a more systematic derivation, we list some important features of the signal here:

- The β-factor defined in eq. (B.55) has a clear physical meaning: it measures the proper time lapse between the creation and annihilation of the heavy particle(s) in its own reference frame. This can be seen by noting that the μ-dependence in the phase $\vartheta(\lambda,\mu,p)$ comes from the time evolution of the heavy particle(s), thus

$$\log\beta = \frac{\partial}{\partial\mu_{12}}\int_{\eta_{1*}}^{\eta_{2*}} E(\eta)\frac{\mathrm{d}\eta}{(-\eta)} = \int_{\eta_{1*}}^{\eta_{2*}}\frac{\mu_{12}}{E(\eta)}\frac{\mathrm{d}\eta}{(-\eta)} = \Delta T, \tag{B.61}$$

  where

$$E(\eta) \coloneqq \sqrt{k_3^2\eta^2+\mu_{12}^2} \tag{B.62}$$

  is the local energy of the heavy particle(s). This shows that $e^{\Delta T}$ is proportional to the squeeze parameter p in the squeezed limit.

- Compared to the $p^{-3/2}$ dilution in tree-level exchanges, the one-loop signal decays faster as p^{-3}. This is because $p^{-3/2}$ is a result of the matter-like dilution of the heavy particle between its production and decay. Due to the presence of two heavy particles in the loop, the cost of dilution is doubled.

- At large p the oscillation frequency of the one-loop amplitude is the same as would be obtained from a tree-level exchange, except for a phase shift $3\pi i/4$. As has been discussed in ref. [58], this phase can be used to separately determine λ and μ from the data, when combined with the oscillating frequency. In ref. [18, 35], it is also suggested that such separation can be achieved by measuring the frequency of the non-analytic component as a function of p for small to moderate values of p. For this purpose, we will also derive the SPA approximation for the signal applicable to moderate p in section B.5.1.

In this section we have approximately calculated the non-analytic piece of the bispectrum, but we did note that the integral in eq. (B.41) has a large contribution from the region $\mu \to \lambda$. In this region the contribution is neither oscillatory, nor suppressed. Therefore, this model is a good case study where the exchange of the heavy scalars provides both the signal and background.

The following two sections will be dedicated to a systematic calculation of all one-loop contributions to the bispectrum, including the backgrounds.

B.4 Systematic Computation: Tree-level Model Revisited

In this section we will evaluate the tree-level amplitude eq. (B.50) systematically. The amplitude eq. (B.50), when evaluated by series expanding $e^{\pm i\delta\phi/\Lambda}$, contains late-time divergence in individual contributions. We will perform a change of basis in the effective tree-level La-

grangian to eliminate this divergence explicitly. Besides the results in ref. [35], we will also derive some extra quantities necessary for the one-loop calculation, such as the full form of the phase in eq. (B.84), the improved approximated form in eq. (B.82), and the contributions from the (++)-diagrams in section B.4.3.

B.4.1 Perturbative expansion under late-time-finite basis

In ref. [35], the amplitude eq. (B.50) is evaluated by series expanding $e^{\pm i\delta\phi/\Lambda}$. This gives rise to two diagrams, shown in figure B.7. It turns out that both diagrams contain divergence in the late-time region. To understand its origin note that, at leading order in α, the tree-level amplitude can be derived from the following Lagrangian:

$$\begin{aligned}\mathcal{L}_\chi &= -\left|\nabla\chi_\mu\right|^2 - M^2\left|\chi_\mu\right|^2 - \alpha_\mu\chi_\mu e^{i\phi/\Lambda} - \bar{\alpha}_\mu\bar{\chi}_\mu e^{-i\phi/\Lambda}, \\ &= -\left|\nabla\chi_\mu\right|^2 - M^2\left|\chi_\mu\right|^2 - \alpha_\mu(-\eta)^{-i\lambda}\chi_\mu e^{i\delta\phi/\Lambda} - \bar{\alpha}_\mu(-\eta)^{i\lambda}\bar{\chi}_\mu e^{-i\delta\phi/\Lambda},\end{aligned} \tag{B.63}$$

where $t = -\log(-\eta)$ and $M^2 = \mu^2 + 9/4$. Note that the μ-dependent coupling α_μ should account for the variation in spectral density: by matching to the tree-level result, we should have

$$\left|\alpha_\mu\right|^2 = \rho^{\mathrm{dS}}_{\mu_1\mu_2}(\mu)\left|\alpha\right|^2. \tag{B.64}$$

In the rest of this section, we will drop the subscript μ on χ_μ and α_μ for simplicity.

The Lagrangian in eq. (B.63) is invariant under

$$\phi \to \phi + c, \qquad \chi \to \chi e^{ic/\Lambda}, \tag{B.65}$$

following from eq. (B.33), which ensures no late-time divergence. However, the shift symmetry is not manifest if one expands $e^{\pm i\delta\phi/\Lambda}$ in a power series for perturbative calculations. As a result, late-time divergences are present in individual diagrams, and only cancel after summing over all diagrams at the same order.

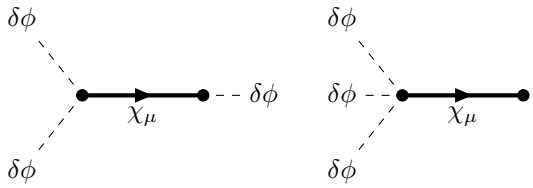


Figure B.7: The effective tree-level diagrams from the spectral decomposition. There are another two diagrams with χ_μ flowing in the opposite direction. In the new basis described by eq. (B.70), only the first diagram is present.

We are free to perform field redefinitions of χ in the Lagrangian eq. (B.63) because the contribution to the amplitude in eq. (B.50) will remain unchanged. To avoid complications with late-time divergences, consider the field redefinition $\chi \mapsto \chi e^{-i\phi/\Lambda}$ that brings the Lagrangian back to its chemical potential form:

$$\begin{aligned}\mathcal{L}_\chi &= -\left(\nabla^\mu + i\frac{\nabla^\mu\phi}{\Lambda}\right)\bar{\chi}\left(\nabla_\mu - i\frac{\nabla_\mu\phi}{\Lambda}\right)\chi - M^2\left|\chi\right|^2 - \alpha\chi - \bar{\alpha}\bar{\chi} \\ &= -\left|\nabla\chi\right|^2 - M^2\left|\chi\right|^2 - \alpha\chi - \bar{\alpha}\bar{\chi} + \frac{1}{\Lambda}\nabla_\mu\phi J^\mu - \frac{1}{\Lambda^2}(\nabla\phi)^2\left|\chi\right|^2 . \end{aligned} \tag{B.66}$$

Treating the last two terms as interactions, the equation of motion of χ at free field order is given by

$$\left[-\left(\nabla_\mu - i\frac{\nabla_\mu\phi_0}{\Lambda}\right)^2 + M^2\right]\chi = -\bar{\alpha}, \qquad \phi_0 = \dot{\phi}_0 t = -\dot{\phi}_0\log(-\eta). \tag{B.67}$$

The inhomogeneous term on the r.h.s. gives rise to a homogeneous classical background for χ:

$$\chi_0 = \frac{\bar{\alpha}}{\lambda^2 + 3i\lambda - M^2} = -\frac{\bar{\alpha}}{\left(\frac{3}{2} - i\lambda\right)^2 + \mu^2}. \tag{B.68}$$

Now, consider the following field redefinition:

$$\chi \mapsto \chi_0 + \delta\chi e^{i\phi/\Lambda}. \tag{B.69}$$

Substituting this into eq. (B.66) gives

$$\begin{aligned}
\mathcal{L}_{\delta\chi} &= -\left|\nabla\delta\chi\right| - M^2\left|\delta\chi\right|^2 - \frac{\left|\chi_0\right|^2}{\Lambda^2}(\nabla\phi)^2 + \mathcal{L}^{(\text{int})}_{\delta\chi} \\
&\supset -\left|\nabla\delta\chi\right| - M^2\left|\delta\chi\right|^2 - \frac{\left|\chi_0\right|^2}{\Lambda^2}(\nabla\delta\phi)^2 + \mathcal{L}^{(\text{int})}_{\delta\chi},
\end{aligned} \tag{B.70}$$

$$\begin{aligned}
\mathcal{L}^{(\text{int})}_{\delta\chi} &= \bar{\chi}_0(-\eta)^{-i\lambda}\left[\frac{i}{\Lambda}\Box\delta\phi + 2(-\eta)\frac{\lambda}{\Lambda}\delta\phi' - \frac{1}{\Lambda^2}(\nabla\delta\phi)^2\right]e^{i\delta\phi/\Lambda}\delta\chi + \text{c.c.} \\
&= \bar{\chi}_0(-\eta)^{-i\lambda}\left[\frac{i}{\Lambda}\Box\delta\phi + 2(-\eta)\frac{\lambda}{\Lambda}\delta\phi' - \frac{1}{\Lambda^2}(\nabla\delta\phi)^2\right]\left(1 + \frac{i}{\Lambda}\delta\phi + \cdots\right)\delta\chi + \text{c.c.},
\end{aligned} \tag{B.71}$$

where we have dropped all constant terms and tadpoles in $\delta\phi$ at the second line of $\mathcal{L}_{\delta\chi}$. This new basis is advantageous over the exponential basis eq. (B.63) as at least one of the $\delta\phi$ in each interaction has a derivative acting on it. This improves the late-time behavior and leads to finite results for each individual diagram.

In the new basis eq. (B.70) onlythe first diagram in figure B.7 contributes to the bispectrum. In the in-in formalism this breaks up into 8 Feynmann diagrams as illustrated in

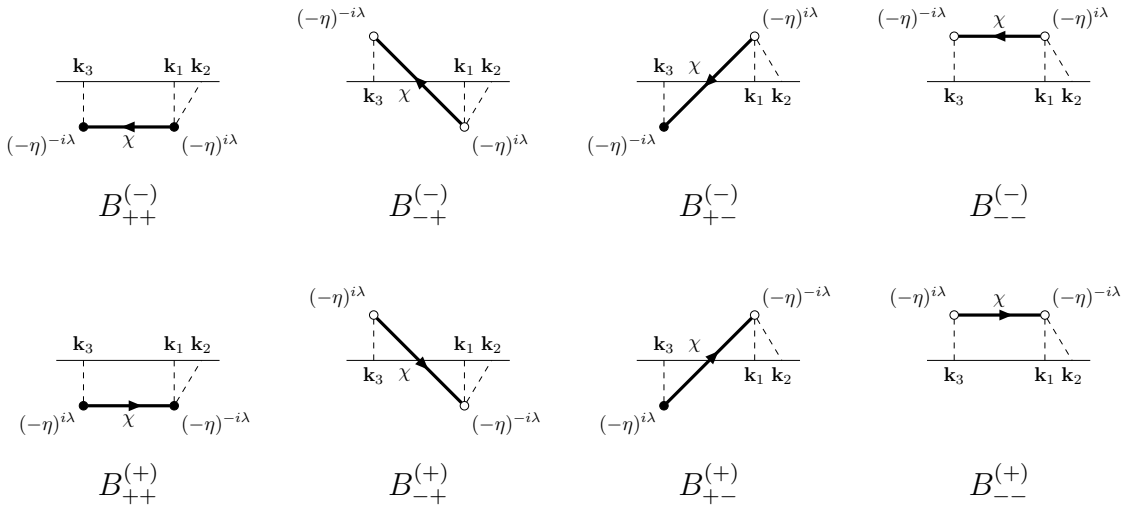


Figure B.8: The various effective tree diagrams contributing to the bispectrum. The solid dots • represent (+)-type vertices, while the empty dots ∘ represent (−)-type vertices. For each diagram shown, there are also another two diagrams where $\mathbf{k}_3$ is permuted with $\mathbf{k}_1$ or $\mathbf{k_2}$. Here we also use the upper/lower half plane notation for the two types of vertices to indicate the direction of time: (+)-type vertices are in the lower half plane with time direction pointing upward, while (−)-type vertices are in the upper half plane with time direction pointing downward. The signs in parentheses correspond to the direction of χ number flow, indicated in the diagram by the arrow.

figure B.8. The total contribution to the bispectrum at this order is thus given by

$$B(k_1,k_2,k_3)=\sum_{s_1,s_2,s_3=\pm}B^{(s_3)}_{s_1s_2}(k_1,k_2,k_3)+(k_3\to k_1,k_2), \tag{B.72}$$

where s_1 gives the index of the 2-point vertex, s_2 gives the index of the 3-point vertex, and s_3 indicates the direction of χ number flow. One can derive the following two relations between the 8 contributions:

$$\begin{aligned} B^{(-s_3)}_{-s_1,-s_2}(k_1,k_2,k_3)&=B^{(s_3)*}_{s_1s_2}(k_1,k_2,k_3),\\ B^{(-s_3)}_{s_1s_2}(k_1,k_2,k_3)&=-\,B^{(s_3)}_{s_1s_2}(k_1,k_2,k_3)\big|_{\lambda\to-\lambda}\,. \end{aligned} \tag{B.73}$$

Therefore, only $B^{(+)}_{++}$ and $B^{(-)}_{-+}$ need to be evaluated explicitly. Nonetheless, since the approximated form relies on the sign of λ, $B^{(\pm)}_{++}$ and $B^{(\pm)}_{-+}$ will be treated separately in approximate calculations. In both cases, we have

$$B^{(\pm)}_{-+}=\frac{\lambda\,|\chi_0|^2}{4k_1^3k_2^3\Lambda^3}\int_0^\infty \mathrm{d}x_2\int_0^\infty \mathrm{d}x_1\,\mathcal{I}^{(\pm)}_{1,-+}(\lambda;x_1)\mathcal{I}^{(\pm)}_{2,-+}(\lambda,p,\chi;x_2), \tag{B.74}$$

$$\begin{aligned} B^{(\pm)}_{++}=-\frac{\lambda\,|\chi_0|^2}{4k_1^3k_2^3\Lambda^3}\Bigg[&\int_0^\infty \mathrm{d}x_2\int_0^{x_2} \mathrm{d}x_1\,\mathcal{I}^{(\pm)}_{1,++}(\lambda;x_1)\mathcal{I}^{(\pm)}_{2,-+}(\lambda,p,\chi;x_2)\\ &+\int_0^\infty \mathrm{d}x_2\int_{x_2}^\infty \mathrm{d}x_1\,\mathcal{I}^{(\mp)*}_{1,-+}(\lambda;x_1)\mathcal{I}^{(\mp)*}_{2,++}(\lambda,p,\chi;x_2)\Bigg], \end{aligned} \tag{B.75}$$

where we have made the substitution $x_{1,2}\coloneqq -k_3\eta_{1,2}$. The functions, $\mathcal{I}_{1,2}$, are given by

$$\begin{aligned} \mathcal{I}^{(s_3)}_{1,s_1+}(\lambda;x_1)&\coloneqq e^{-\pi\mu/2}\sqrt{\frac{\pi}{2}}x_1^{-1/2+is_3\lambda}e^{-is_1x_1}H^{(1)}_{i\mu}(x_1),\\ \mathcal{I}^{(s_3)}_{2,s_1+}(\lambda,p,\chi;x_2)&\coloneqq e^{-\pi\mu/2}\sqrt{\frac{\pi}{2}}\sum_{n=0}^{2}(-s_1i)^n\cdot c_n^{s_3}\cdot x_2^{-1/2+n-is_3\lambda}e^{is_1px_2}H^{(1)*}_{i\mu}(x_2). \end{aligned} \tag{B.76}$$

The coefficients inside $\mathcal{I}_2$ are given by

$$\begin{aligned}
c_0^{\pm} &:= -\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_3^2} \pm i\lambda\frac{k_1^2+k_2^2}{k_3^2} = \frac{1}{2}\big[(1\pm 2i\lambda)\big(p^2-2Q\big)-1\big],\\
c_1^{\pm} &:= -\frac{\mathbf{k}_1\cdot\mathbf{k}_2(k_1+k_2)}{k_3^3} \pm \frac{i\lambda k_1k_2(k_1+k_2)}{k_3^3} = \frac{1}{2}p\big[p^2-2(1\mp i\lambda)Q-1\big],\\
c_2^{\pm} &:= -\frac{(\mathbf{k}_1\cdot\mathbf{k}_2)k_1k_2}{k_3^4} + \frac{k_1^2k_2^2}{k_3^4} = \frac{1}{2}Q(p^2-1),
\end{aligned} \tag{B.77}$$

where $Q := k_1k_2/k_3^2 = \frac{1}{4}(p^2-\chi^2)$. Notice in eq. (B.76) that as $x_i\to 0$, the integrands go as $x_i^{-1/2+n\pm i\lambda}$ with $n\geq 0$. The amplitude is now expressed in a form that is manifestly free of late-time divergences.

B.4.2 Contributions from B_{-+}

For $\lambda,\mu,\lambda-\mu\gg 1$, the SPA can be used to calculate eqs. (B.74) and (B.75). Starting from the Hankel functions $H_{i\mu}^{(1)}(x)$ with $\mu\gg 1$, we have Debye's asymptotic representation of the Hankel function:

$$\begin{aligned}
H_{i\mu}^{(1)}(x) &\simeq \sqrt{\frac{2}{\pi}}e^{\pi\mu/2}\cdot\frac{e^{ih(x,\mu)}}{(x^2+\mu^2)^{1/4}}\left\{1+\mathcal{O}\left[\frac{\mu^2}{(x^2+\mu^2)^{3/2}}\right]\right\},\\
h(x,\mu) &= -\frac{\pi}{4}+\sqrt{x^2+\mu^2}-\mu\,\mathrm{arcsinh}\,\frac{\mu}{x}, \qquad \frac{\partial h}{\partial x} = \sqrt{1+\frac{\mu^2}{x^2}},
\end{aligned} \tag{B.78}$$

which follows directly from the WKB method. Substituting into eq. (B.76), the two integrands factorize into slowly-varing parts and highly oscillatory parts:

$$\begin{aligned}
\mathcal{I}_{1,s_1+}^{(s_3)}(\lambda;x_1) &\simeq \frac{1}{x_1^{1/2}(x_1^2+\mu^2)^{1/4}}e^{ig_{1,s_1}^{(s_3)}(x_1,\lambda,\mu)},\\
\mathcal{I}_{2,s_1+}^{(s_3)}(\lambda,p,\chi;x_2) &\simeq \frac{\sum_{n=0}^{2}(-s_1ix_2)^n c_n^{s_3}}{x_2^{1/2}(x_2^2+\mu^2)^{1/4}}e^{ig_{2,s_1}^{(s_3)}(x_2,\lambda,\mu,p)},
\end{aligned} \tag{B.79}$$

where the phases are given by

$$\begin{aligned} g^{(s_3)}_{1,s_1}(x_1,\lambda,\mu) &\coloneqq -s_1 x_1 + s_3\lambda\log x_1 + h(x_1,\mu), \quad \frac{\partial g^{(s_3)}_{1,s_1}}{\partial x_1} = -s_1 + \frac{s_3\lambda}{x_1} + \sqrt{1+\frac{\mu^2}{x_1^2}}, \\ g^{(s_3)}_{2,s_1}(x_2,\lambda,\mu,p) &\coloneqq s_1 p x_2 - s_3\lambda\log x_2 - h(x_2,\mu), \quad \frac{\partial g^{(s_3)}_{2,s_1}}{\partial x_2} = s_1 p - \frac{s_3\lambda}{x_2} - \sqrt{1+\frac{\mu^2}{x_2^2}}. \end{aligned} \tag{B.80}$$

The dominant contributions to eqs. (B.74) and (B.75) then come from those diagrams for which a stationary phase, $g'(x_*) = 0$, with $x_* > 0$, exists in both integrals. The stationary points, along with the condition for their existence, are tabulated in table B.1. For the x_2 integrals the stationary point is given by

$$x_*^{\pm} \coloneqq \frac{\lambda p \pm \sqrt{\lambda^2 + \mu^2(p^2-1)}}{p^2-1} \xrightarrow{p\gg\lambda/\mu} \frac{\lambda\pm\mu}{p} \tag{B.81}$$

in the squeezed limit. For ease of notation we give only the squeezed limit expression in table B.1.

Integrand	Stationary phase point $x_{1,2}$	Conditions
$\mathcal{I}^{(s_3)}_{1,-+}$	$\dfrac{\lambda^2-\mu^2}{2\lambda}$	$s_3 = -, \quad \lambda > \mu$
$\mathcal{I}^{(s_3)}_{1,++}$	$\dfrac{\mu^2-\lambda^2}{2\lambda}$	$s_3 = -, \quad \lambda < \mu$
$\mathcal{I}^{(s_3)}_{2,-+}$	$\dfrac{\lambda-\mu}{p}$	$s_3 = -, \quad \lambda > \mu$
$\mathcal{I}^{(s_3)}_{2,++}$	$-\dfrac{\lambda-\mu}{p}$	$s_3 = -, \quad \lambda < \mu$
	$\dfrac{\lambda+\mu}{p}$	$s_3 = +$

Table B.1: The stationary phase points for the integrands in eq. (B.79) in the squeezed limit. The third column states the conditions for a stationary phase point to exist on the positive real axis.

By comparing table B.1 with eqs. (B.74) and (B.75), it is manifest that only $B^{(-)}_{-+}$ has stationary points for both the x_1- and the x_2-integral, and is the dominant contribution to the whole $B_{\rm tree}$. Using SPA

$$\begin{aligned} B_{\rm tree}(k_1,k_2,k_3) &\simeq B^{(-)}_{-+}(k_1,k_2,k_3) \\ &\simeq \frac{\pi}{4k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda^{1/2}e^{i\vartheta(\lambda,\mu,p)}}{(\lambda^2-\mu^2)[\lambda^2+\mu^2(p^2-1)]^{1/4}}\sum_{n=0}^{2} d_n (ix_*)^{n-2}, \end{aligned} \tag{B.82}$$

where $x_* \coloneqq x_*^-$ is defined by (B.81),

$$d_0 \coloneqq 1, \qquad d_1 \coloneqq p, \qquad d_2 \coloneqq \frac{1}{4}\big(p^2-\chi^2\big). \tag{B.83}$$

The phase, ϑ, is given by

$$\begin{aligned} \vartheta(\lambda,\mu,p) &= \lambda\log\frac{2\lambda x_*}{\lambda^2-\mu^2} - \mu\log\frac{\lambda+\mu}{\lambda-\mu} + \mu\,\text{arcsinh}\,\frac{\mu}{x_*} \\ &\simeq \begin{cases} -(\lambda-\mu)\log p + \vartheta_*(\lambda,\mu), & p \gg \lambda/\mu, \\ -\lambda\log\dfrac{p+1}{2}, & p \ll \lambda/\mu,\ \lambda \gg \mu. \end{cases} \end{aligned} \tag{B.84}$$

In the squeezed limit, this reduces to eq. (B.51).

The amplitude, $B^{(-)}_{-+}$, can also be calculated exactly by noting that the two time integrals in eq. (B.74) factorize and each of them can be reduced to a standard integral given in eq. (B.177)

$$\begin{aligned} B^{(-)}_{-+}(k_1,k_2,k_3) &= \frac{1}{4k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda e^{\pi\lambda}}{\big(\lambda^2-\mu^2-\frac{9}{4}\big)^2+9\lambda^2}\mathbf{F}^{\frac{1}{2}-i\lambda}_{i\mu}(1)\sum_{n=0}^{2}\frac{c_n^-}{2^n}\mathbf{F}^{\frac{1}{2}+n+i\lambda}_{i\mu}(p) \\ &= -\frac{1}{2k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda e^{\pi\lambda}}{\big(\frac{3}{2}+i\lambda\big)^2+\mu^2}\mathbf{F}^{\frac{1}{2}-i\lambda}_{i\mu}(1)\sum_{n=0}^{2}\frac{d_n}{2^n}\mathbf{F}^{-\frac{3}{2}+n+i\lambda}_{i\mu}(p), \end{aligned} \tag{B.85}$$

where[9]

$$\mathbf{F}^{\rho}_{i\mu}(p) \coloneqq \Gamma\begin{pmatrix}\rho+i\mu,\rho-i\mu\\ \frac{1}{2}+\rho\end{pmatrix} F\begin{pmatrix}\rho+i\mu,\rho-i\mu\\ \frac{1}{2}+\rho\end{pmatrix}; \frac{1-p}{2}\Bigg), \tag{B.86}$$

and $F\left(\begin{smallmatrix}a,b\\ c\end{smallmatrix}; z\right) = {}_2F_1(a,b,c;z)$ is the hypergeometric function. In the second line in eq. (B.85) the recurrence relation given in eq. (B.183) for $\mathbf{F}^{\rho}_{i\mu}(p)$ has been used, after which our result is equivalent to the one in ref. [35] evaluated in the basis in eq. (B.63). Our calculation is manifestly finite at every step compared to the one in the exponential basis.

The exact expression in eq. (B.85) reduces back to the stationary phase expression in eq. (B.82) for $\lambda, \mu, \lambda-\mu \gg 1$, as shown in figure B.9. Meanwhile, the naive singularity at $\mu \to \lambda$ in eqs. (B.51) and (B.82) becomes a finite resonance peak in the exact expression. The other contribution, $B^{(+)}_{-+}$, can be obtained from $B^{(-)}_{-+}$ using eq. (B.73):

$$B^{(+)}_{-+}(k_1,k_2,k_3) = -\left.B^{(-)}_{-+}(k_1,k_2,k_3)\right|_{\lambda\to-\lambda} \sim e^{-2\pi\lambda} B^{(-)}_{-+}(k_1,k_2,k_3), \tag{B.87}$$

i.e. it is always exponentially suppressed for $\lambda \gg 1$.

B.4.3 Contributions from B_{++}

To derive the full tree-level amplitude, we must also evaluate $B^{(\pm)}_{++}$ as well as all permuted diagrams ($k_3 \to k_1, k_2$). We defer the discussion of the permuted diagrams to section B.5.4.

Now we focus on the contribution from $B^{(\pm)}_{++}$, given by the integral eq. (B.75). We start

[9] We use the following compact notation for products of gamma functions:

$$\Gamma\begin{pmatrix}a_1,a_2,\cdots,a_m\\ b_1,b_2,\cdots,b_n\end{pmatrix} \coloneqq \frac{\Gamma(a_1)\Gamma(a_2)\cdots\Gamma(a_m)}{\Gamma(b_1)\Gamma(b_2)\cdots\Gamma(b_n)}.$$

by breaking this integral into two parts for calculation convenience:

$$B^{(\pm)}_{++}(k_1,k_2,k_3) = B^{(\pm)}_{++,\mathrm{NA}}(k_1,k_2,k_3) + B^{(\pm)}_{++,\mathrm{A}}(k_1,k_2,k_3), \tag{B.88}$$

$$B^{(\pm)}_{++,\mathrm{NA}}(k_1,k_2,k_3) \coloneqq -\frac{\lambda\,|\chi_0|^2}{4k_1^3k_2^3\Lambda^3}\int_0^\infty \mathrm{d}x_2 \int_0^\infty \mathrm{d}x_1\, \mathcal{I}^{(\mp)*}_{1,-+}(\lambda;x_1)\mathcal{I}^{(\mp)*}_{2,++}(\lambda,p,\chi;x_2), \tag{B.89}$$

$$\begin{aligned} B^{(\pm)}_{++,\mathrm{A}}(k_1,k_2,k_3) \coloneqq -\frac{\lambda\,|\chi_0|^2}{4k_1^3k_2^3\Lambda^3}\Bigg[&\int_0^\infty \mathrm{d}x_2 \int_0^{x_2} \mathrm{d}x_1\, \mathcal{I}^{(\pm)}_{1,++}(\lambda;x_1)\mathcal{I}^{(\pm)}_{2,-+}(\lambda,p,\chi;x_2) \\ &- \int_0^\infty \mathrm{d}x_2 \int_0^{x_2} \mathrm{d}x_1\, \mathcal{I}^{(\mp)*}_{1,-+}(\lambda;x_1)\mathcal{I}^{(\mp)*}_{2,++}(\lambda,p,\chi;x_2)\Bigg]. \end{aligned} \tag{B.90}$$

We will soon see that in the squeezed limit they correspond to the non-analytic and analytic contributions respectively. Since $B^{(\pm)}_{++,\mathrm{NA}}(k_1,k_2,k_3)$ also factorizes into two time integrals, it can be evaluated using eq. (B.177):

$$B^{(\pm)}_{++,\mathrm{NA}}(k_1,k_2,k_3) = \frac{1}{2k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda e^{\pm\pi\lambda}}{\left(\frac{3}{2}\mp i\lambda\right)^2+\mu^2}\mathbf{F}^{\frac{1}{2}\pm i\lambda}_{i\mu}(1)\sum_{n=0}^{2}(-1)^n\frac{d_n}{2^n}\mathbf{F}^{-\frac{3}{2}+n\mp i\lambda}_{i\mu}(-p+i\epsilon). \tag{B.91}$$

As shown in ref. [35], $B^{(+)}_{++,\mathrm{NA}}$ is suppressed by $e^{-\pi|\lambda-\mu|}$ and only becomes comparable to $B^{(-)}_{-+}$ when $\mu\simeq\lambda$, while one can similarly show that $B^{(-)}_{++,\mathrm{NA}}$ is suppressed by $e^{-\pi\max\{\lambda,\mu\}}$ and can be always neglected.

The second part of eq. (B.75) defined by eq. (B.90) does not factorize between x_1 and x_2, and we will leave its exact evaluation to appendix B.9.4. Here we will derive its approximate form in a different approach. As implied by table B.1, the two time integrals in eq. (B.90) never have stationary phase points simultaneously. Therefore, it is expected that the heavy propagator is far off-shell, so that the amplitude is dominated by the $\eta_1\to\eta_2$ limit and can be described

by an effective local Lagrangian of the inflaton:[10]

$$\begin{aligned}\mathcal{L}_\phi &= |\alpha|^2\, e^{i\phi/\Lambda}\frac{1}{-\Box+M^2}e^{-i\phi/\Lambda}\\ &= |\alpha|^2\left[-\left(\nabla^\mu-\frac{i\nabla^\mu\phi}{\Lambda}\right)\left(\nabla_\mu-\frac{i\nabla_\mu\phi}{\Lambda}\right)+M^2\right]^{-1}\cdot 1\\ &= -\,|\alpha|^2\sum_{k=0}^{\infty}\frac{\left[-\Box+\frac{2i\nabla^\mu\delta\phi\nabla_\mu+i\Box\delta\phi}{\Lambda}+\frac{(\nabla\delta\phi)^2}{\Lambda^2}-2\lambda\left(i\eta\partial_\eta-\frac{3i}{2}+\frac{\eta\delta\phi'}{\Lambda}\right)\right]^k}{\left(\lambda^2-M^2\right)^{k+1}}\cdot 1.\end{aligned}\tag{B.92}$$

For $|\lambda^2-M^2|\gtrsim\mathcal{O}(\lambda^2)$, this becomes an expansion in λ^{-1}. The leading-order contribution to the 3-point function is given by

$$\mathcal{L}^{(\mathrm{loc})}_{\delta\phi^3}=\frac{4\lambda\,|\alpha|^2}{\Lambda^3(\lambda^2-M^2)^3}(-\eta)\delta\phi'(\nabla\delta\phi)^2+\frac{8\lambda^3\,|\alpha|^2}{\Lambda^3(\lambda^2-M^2)^4}(-\eta)^3\delta\phi'^3+\mathcal{O}\big(\lambda^{-7}\big).\tag{B.93}$$

From this local Lagrangian we arrive at the following estimation for the analytic contribution:

$$B_{++}\simeq B_{++,\mathrm{A}}\simeq B_{\mathrm{loc}}(k_1,k_2,k_3)\coloneqq B_{+,\mathrm{loc}}(k_1,k_2,k_3)+(k_3\to k_1,k_2)+\mathrm{c.c.},\tag{B.94}$$

$$\begin{aligned}B_{+,\mathrm{loc}}(k_1,k_2,k_3)=\frac{1}{8k_1^3k_2^3k_3^3}\int_{-\infty_-}^{0}\frac{\mathrm{d}\eta}{\eta^4}e^{i(k_1+k_2+k_3)\eta}\Bigg\{&\frac{8i\lambda\,|\alpha|^2}{\Lambda^3(\lambda^2-M^2)^3}(-\eta)\big(k_3^2\eta\big)\\ &\times\eta^2\big[-k_1^2k_2^2\eta^2-(\mathbf{k}_1\cdot\mathbf{k}_2)(1-ik_1\eta)(1-ik_2\eta)\big]\\ &+\frac{1}{3}\frac{48i\lambda^3\,|\alpha|^2}{\Lambda^3(\lambda^2-M^2)^4}(-\eta)^3\big(k_1^2k_2^2k_3^2\eta^3\big)\Bigg\}\\ =-\frac{1}{k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{(\lambda^2-M^2)^3}&\left[\frac{p^2}{p+1}-\frac{Q(p+2)}{(p+1)^2}-\frac{1}{2}-\frac{4\lambda^2}{\lambda^2-M^2}\frac{Q^2}{(p+1)^3}\right],\end{aligned}\tag{B.95}$$

[10]For $M\gg\lambda$, χ decouples from the theory and this effective Lagrangian has the usual interpretation under effective field theory. For $\lambda>M$, such interpretation is obscured by the fact that χ are massively produced on-shell.

where $Q \coloneqq k_1 k_2/k_3^2 = \frac{1}{4}(p^2 - \chi^2)$. At large p, eq. (B.95) becomes

$$B_{+,\mathrm{loc}}(k_1,k_2,k_3) \xrightarrow{p\gg 1} -\frac{3}{4k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{(\lambda^2-M^2)^3}\left(1-\frac{1}{3}\frac{\lambda^2}{\lambda^2-M^2}\right)p, \tag{B.96}$$

$$F_{+,\mathrm{loc}}(k_1,k_2,k_3) \xrightarrow{p\gg 1} 10\left|\frac{\alpha}{\Lambda}\right|^2\frac{\lambda^2}{(\lambda^2-M^2)^3}\left(1-\frac{1}{3}\frac{\lambda^2}{\lambda^2-M^2}\right)p^{-2}. \tag{B.97}$$

Compared with eq. (B.51), we find that $B_{++}(k_1,k_2,k_3)$ is suppressed by λ^{-3} for $\lambda,\mu \gg 1$, except near $\mu \simeq \lambda$. It is also worth noting that the p^{-2} behavior in the squeezed limit of the normalized bispectrum is a feature applied to all models of single-field inflation [129].

Figure B.9 shows the sizes of $B^{(-)}_{-+}$, $B^{(+)}_{++,\mathrm{NA}}$ and $B^{(\pm)}_{++,\mathrm{A}}$ as functions of μ for $\lambda = 40$ and $p = 20$. It is obvious that $B^{(+)}_{++,\mathrm{NA}}$ and $B^{(\pm)}_{++,\mathrm{A}}$ are both negligible in magnitude compared to $B^{(-)}_{-+}$, except near $\mu \simeq \lambda$ where they become comparable. Since the one-loop amplitude involves integrating over, μ as in eq. (B.41), $B^{(+)}_{++,\mathrm{NA}}$ and $B^{(\pm)}_{++,\mathrm{A}}$ cannot be dropped in the full one-loop evaluation.

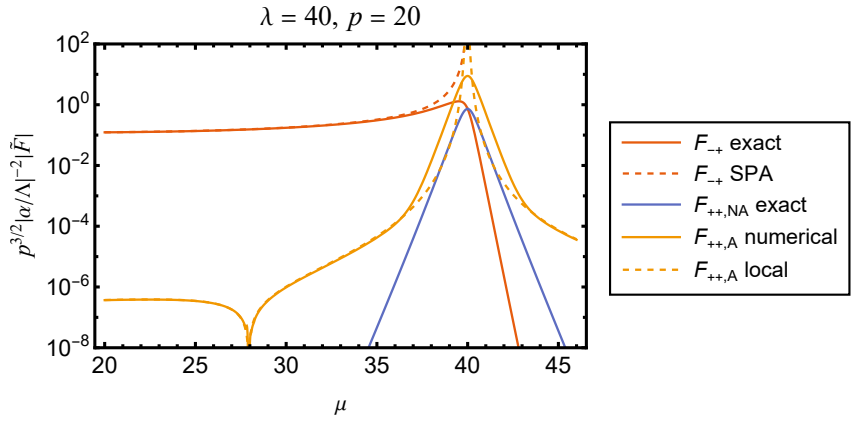


Figure B.9: The magnitudes of contributions from $F^{(-)}_{-+}$, $F^{(+)}_{++,\mathrm{NA}}$ and $F_{++,\mathrm{A}}$ versus μ for $\lambda = 40$, and $p = 20$. For $F_{++,\mathrm{A}}$ we take the average over permuted diagrams to make the numerical result and local estimation eq. (B.92) comparable. From this plot we can see all important contributions near $\mu = \lambda$, where most approximations break down. Note that there are other contributions that are exponentially suppressed by at least $e^{-\pi\lambda}$ and thus invisible on the plot.

B.5 Systematic Computation: One-loop Model

In this section we will compute the one-loop amplitude by performing the integral over the spectral representation in eq. (B.41), using the tree-level results of the previous section. We will first derive the signal and the background for general values of the squeezing parameter, p, using the same approximation as in section B.3.3. We then demonstrate how the full amplitude at $\mathcal{O}(\alpha^2)$ can be evaluated, with the full results derived in appendix B.9.5. Finally, we will include the combinatorial backgrounds from permuted diagrams, and discuss higher order corrections in the last subsection.

B.5.1 Signal from the threshold limit

For $\lambda, \mu_1, \mu_2, \lambda - \mu_1 - \mu_2 \gg 1$, we have argued in section B.3.3 that the B_{-+} amplitude is dominated by the region around $\mu = \mu_{12}$ and $\mu = \lambda$. Including both contribution gives the following approximation

$$\begin{aligned} B_{\text{1-loop}}(k_1, k_2, k_3) &= \int_0^\infty \rho^{\text{dS}}_{\mu_1\mu_2}(\mu) B_{\text{tree}}(k_1, k_2, k_3; \mu) \\ &\simeq B^{(\mu_{12})}_{\text{1-loop}}(k_1, k_2, k_3) + B^{(\lambda)}_{\text{1-loop}}(k_1, k_2, k_3). \end{aligned} \tag{B.98}$$

To extract the $\mu \to \mu_{12}$ limit we use a slightly modified form of eq. (B.55)

$$B^{(\mu_{12})}_{\text{1-loop}}(k_1, k_2, k_3) \simeq B_{\text{tree,SPA}}(k_1, k_2, k_3; \mu_{12}) \int_{-\infty}^{\infty+} \rho^{\text{dS}}_{\mu_1\mu_2}(\mu_{12} + \varepsilon) \beta^{i\varepsilon} e^{\frac{i}{2}\gamma\varepsilon^2} \, \mathrm{d}\varepsilon, \tag{B.99}$$

where

$$\begin{aligned}\beta &\coloneqq \exp\left(\frac{\partial\vartheta}{\partial\mu}\right)\Big|_{\mu=\mu_{12}} = \frac{\lambda-\mu_{12}}{\lambda+\mu_{12}}\sqrt{\frac{\lambda+\mu_{12}-px_*}{\lambda-\mu_{12}-px_*}} \xrightarrow{p\gg\lambda/\mu} \frac{2\mu_{12}p}{\lambda+\mu_{12}},\\ \gamma &\coloneqq \frac{\partial^2\vartheta}{\partial\mu^2}\Big|_{\mu=\mu_{12}} = \frac{\lambda}{\lambda^2-\mu_{12}^2}\left[\frac{\lambda p}{\sqrt{\lambda^2+\mu_{12}^2(p^2-1)}}-1\right] \xrightarrow{p\gg\lambda/\mu} \frac{\lambda}{\mu_{12}(\lambda+\mu_{12})}.\end{aligned} \tag{B.100}$$

Note that compared to eq. (B.55), we have included the second order expansion of the phase ϑ near ε in eq. (B.99) to improve the estimation for moderate p. Similarly, we extract the $\mu\to\lambda$ background with

$$B^{(\lambda)}_{\text{1-loop}}(k_1,k_2,k_3) \simeq \rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda)\int_{-\infty}^{\infty} B_{\text{tree}}(k_1,k_2,k_3;\lambda+\varepsilon)\,\mathrm{d}\varepsilon\,. \tag{B.101}$$

In other words, we can use the SPA form of B_{tree} near $\mu\sim\mu_{12}$ and the Minkowski limit of $\rho^{\text{dS}}_{\mu_1\mu_2}(\mu)$ near $\mu\sim\lambda$ respectively.

The fact that the integrand $\rho^{\text{dS}}_{\mu_1\mu_2}(\mu)B_{\text{tree}}(\mu)$ drops rapidly to zero near $\mu\lesssim\mu_{12}$ and has a peak at $\mu\simeq\lambda$ has clear physical meanings: the former corresponds to the threshold limit where the mass μ of the multi-particle state is close to or drops below the total rest mass μ_{12} of the two components; the latter corresponds to the resonance limit, which can be traced back to the resonance behavior in the homogeneous background of the χ field given by eq. (B.68) as $\mu\to\lambda$. As has been argued in section B.3.3, the threshold limit gives the non-analytic signal, while the resonance limit contributes to the analytic background.

For the signal, we can combine eq. (B.46) with eq. (B.99) to get a more accurate stationary

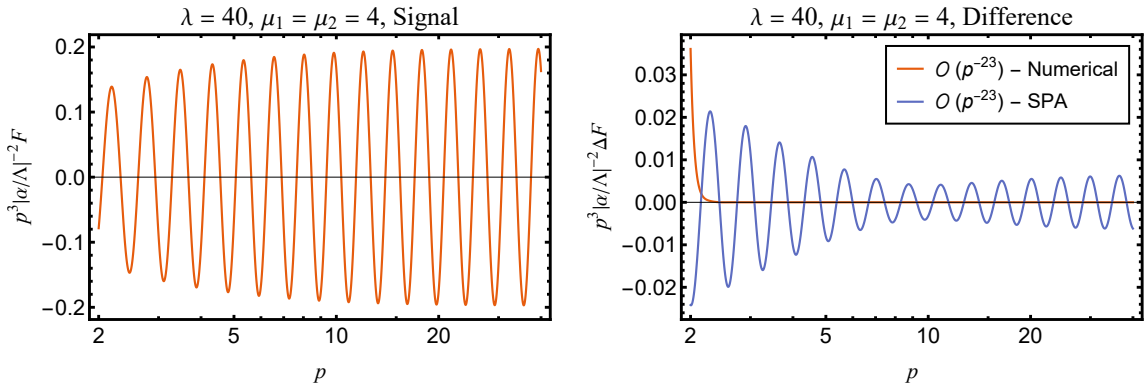


Figure B.10: (Left) The one-loop oscillatory signal coming from the threshold limit for $\lambda = 40$ and $\mu = 4$, rescaled by p^3. (Right) The difference between two approximate calculations and the numerical evaluation of eq. (B.41): (i) the analytic sum over poles up to $\mathcal{O}(p^{-23})$ (see section B.5.3 and appendix B.9.5); (ii) the SPA result eq. (B.102). In calculating the difference between the analytic sum and the numerical result, the difference in the full $F^{(-)}_{-+}$ amplitude is computed, which includes both the signal and the analytic background in $F^{(-)}_{-+}$. For p away from 2, the numerical result agrees with the analytic sum, while the error for the improved SPA result is within 5%.

phase estimation:

$$B^{(\mu_{12})}_{\text{1-loop}}(k_1,k_2,k_3) \simeq \frac{1}{2\pi^{3/2}}\sqrt{\frac{\mu_1\mu_2}{\mu_{12}}}B_{\text{tree}}(k_1,k_2,k_3;\mu_{12}) \times e^{3\pi i/4}\sum_{k=0}^{\infty}\frac{(3/2)_k}{k!}\beta^{-3/2-2k}e^{-\frac{i}{2}\gamma(3/2+2k)^2}. \qquad (\log\beta \gg \gamma) \qquad \text{(B.102)}$$

where $(a)_m \coloneqq a(a+1)\cdots(a+m-1)$ is the Pochhammer symbol. Note that we have approximated the integral as a sum over residues of the poles in eq. (B.46), which is only valid up to corrections of order

$$\rho^{\text{dS}}_{\mu_1\mu_2}(\mu_{12}+\varepsilon)\big|_{\gamma\varepsilon+\log\beta=0} \sim e^{-\pi\log\beta/\gamma}. \qquad \text{(B.103)}$$

Eq. (B.102) reduces to eq. (B.57) in the squeezed limit with $\gamma \ll 1$. Figure B.10 shows the shape of the non-analytic signal in terms of the bispectrum, F, and its comparison with the

numerical result and the analytic form we will derive later in section B.5.3. We also calculate the oscillating frequency at different p-value and the result is presented in figure B.11. As argued in [18,35], the frequency at small p and large p can be used together to infer both the total mass μ_{12}, and the chemical potential λ.

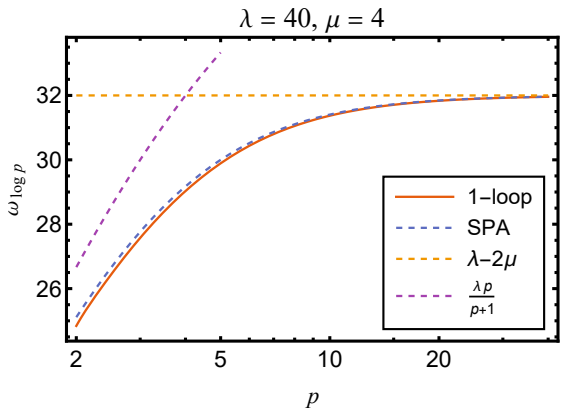


Figure B.11: The oscillation frequency of the signal for $\lambda = 40$ and $\mu_1 = \mu_2 = 4$. The frequency of the signal is defined as $\omega_{\log p} \coloneqq -\operatorname{Im} \frac{\partial \log \widetilde{F}}{\partial \log p}$. The SPA curve is given by $-\frac{\partial \vartheta}{\partial \log p}$ with ϑ defined by eq. (B.84). The other two lines are derived from the small p and the large p limits in eq. (B.84), showing that we are able to extract both the chemical potential and total mass independently from data, as emphasized in [18,35].

B.5.2 Resonance backgrounds

The one-loop amplitude also differs from the tree-level amplitude by having an analytic background coming from the resonance production. Such a background cannot be neglected as the amplitude of the one-loop signal decays as $\mathcal{O}(p^{-3})$, which is faster than a typical analytic contribution, decreasing as $\mathcal{O}(p^{-2})$ in the large-p limit. In this subsection, we will approximately evaluate this background coming from the resonance limit of B_{-+}, given by eq. (B.101), as well as a similar one from B_{++}. Starting from eq. (B.101), we need to substitute terms singular near

$\mu \simeq \lambda$ in the SPA form of eq. (B.82) by their exact expression in eq. (B.85):

$$\frac{(ix_*)^{n-2}}{\lambda^2-\mu^2} \to -\frac{1}{2\pi}\frac{\Gamma\left(\frac{1}{2}-i\lambda+i\mu\right)\Gamma\left(-\frac{3}{2}+n+i\lambda-i\mu\right)}{\left(\frac{3}{2}+i\lambda\right)^2+\mu^2}\left(\frac{x_*}{\lambda-\mu}\right)^{n-2} e^{\pi(\lambda-\mu)}. \tag{B.104}$$

Note that $x_*/(\lambda-\mu) \to 1/p$ as $\mu \to \lambda$. For $\lambda-\mu \gg 1$ one can apply the Stirling approximation on the RHS and obtain the LHS, while this does not apply when $\mu \simeq \lambda$. Besides these terms, we can approximate all other slow-varying terms in eq. (B.82) by their values at $\mu = \lambda$,

$$B^{(-)}_{\text{tree},-+}(k_1,k_2,k_3;\lambda+\varepsilon) \simeq \frac{\pi}{8k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{p^{3/2+i\varepsilon}}{2i\lambda\left(\frac{3}{2}-i\varepsilon\right)}\left(\frac{1}{\frac{3}{2}+i\varepsilon}-\frac{p^2-\chi^2}{4p^2}\right)\frac{e^{-\pi\varepsilon}}{\cosh\pi\varepsilon}. \tag{B.105}$$

One can now evaluate the μ-integral analytically, closing the contour in the upper half plane and applying using the residue theorem:

$$\begin{aligned}
&\int_{-\infty}^{\infty}\frac{p^{i\varepsilon}}{\frac{3}{2}-i\varepsilon}\left(\frac{1}{\frac{3}{2}+i\varepsilon}-\frac{p^2-\chi^2}{4p^2}\right)\frac{e^{-\pi\varepsilon}}{\cosh\pi\varepsilon}\,\mathrm{d}\varepsilon \\
&= -\frac{2i}{3}p^{-3/2}\left(p+\log p+\pi i-\sum_{k=0}^{\infty}\frac{p^{-1-k}}{k+1}+\frac{p^2+3\chi^2}{4}\sum_{k=0}^{\infty}\frac{p^{-1-k}}{k+2}\right), \\
&= -\frac{2i}{3}p^{-3/2}\left[p+\log(p-1)+\pi i-\frac{p^2+3\chi^2}{4}\left(1+p\log\frac{p-1}{p}\right)\right],
\end{aligned} \tag{B.106}$$

giving

$$\begin{aligned}
B^{(\lambda)}_{\text{1-loop},-+}(k_1,k_2,k_3) \simeq &-\frac{1}{192\pi k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\prod_{\epsilon_1,\epsilon_2=\pm}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda^2} \\
&\times\left[p+\log(p-1)+\pi i-\frac{p^2+3\chi^2}{4}\left(1+p\log\frac{p-1}{p}\right)\right].
\end{aligned} \tag{B.107}$$

In terms of the properly normalized bispectrum:

$$F^{(\lambda)}_{\text{1-loop},-+} \simeq \frac{5}{36\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_i}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda(p^3+3\chi^2 p+4)}\left[p+\log(p-1)-\frac{p^2+3\chi^2}{4}\left(1+p\log\frac{p-1}{p}\right)\right],$$
$$\xrightarrow{p\gg 1} \frac{5}{32\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_i}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda p^2}\left(1+\frac{2+24\log p}{27p}-\frac{5+16\chi^2}{6p^2}+\cdots\right), \tag{B.108}$$

where we also include the contribution from the complex conjugate.

By combining eq. (B.102) with eq. (B.108) we have approximately B_{-+}. This contribution contains an oscillatory signal on top of a smooth background. If this was the tree level exchange of [35] or [18] then we would have successfully calculated the dominant contribution to the bispectrum. However, recall from section B.4.3 that we B_{++} is not suppressed in the region $\mu \simeq \lambda$, which we must include in the loop integration. Therefore, we must take into account a further smooth background from the B_{++} diagram.

The tree level amplitude for B_{++} has been split into $B^{(\pm)}_{++,\text{NA}}$ and $B^{(\pm)}_{++,\text{A}}$ in eq. (B.88). According to section B.4.3, $B^{(-)}_{++,\text{NA}}$ is always exponentially suppressed and can be neglected. For $B^{(+)}_{++,\text{NA}}$, we use the following hypergeometric function identity [71, eq. (15.8.4)]:

$$\mathbf{F}^{\rho}_{i\mu}(-p+i\epsilon) = \Gamma\begin{pmatrix}\rho+i\mu,\rho-i\mu\\ \frac{1}{2}+\rho\end{pmatrix}\left[\Gamma\begin{pmatrix}\frac{1}{2}+\rho,\frac{1}{2}-\rho\\ \frac{1}{2}+i\mu,\frac{1}{2}-i\mu\end{pmatrix}F\begin{pmatrix}\rho+i\mu,\rho-i\mu\\ \frac{1}{2}+\rho\end{pmatrix};\frac{1-p}{2}\right)$$
$$+\Gamma\begin{pmatrix}\frac{1}{2}+\rho,-\frac{1}{2}+\rho\\ \rho+i\mu,\rho-i\mu\end{pmatrix}\left(\frac{1-p}{2}+i\epsilon\right)^{1/2-\rho}F\begin{pmatrix}\frac{1}{2}+i\mu,\frac{1}{2}-i\mu\\ \frac{3}{2}-\rho\end{pmatrix};\frac{1-p}{2}\right]. \tag{B.109}$$

For $\rho = -3/2 + n - i\lambda$, the Stirling approximation implies that

$$\begin{aligned}
&\Gamma\left(\begin{matrix}\frac{1}{2}+\rho, \frac{1}{2}-\rho \\ \frac{1}{2}+i\mu, \frac{1}{2}-i\mu\end{matrix}\right) \simeq -i(-1)^n e^{-\pi(\lambda-\mu)}, \\
&\left|\Gamma\left(\begin{matrix}\frac{1}{2}+\rho, -\frac{1}{2}+\rho \\ \frac{1}{2}+i\mu, \frac{1}{2}-i\mu\end{matrix}\right)\left(\frac{1-p}{2}+i\epsilon\right)^{1/2-\rho}\right| \sim e^{-\pi(2\lambda-\mu)}.
\end{aligned} \tag{B.110}$$

Consequently, the second term is suppressed by $e^{-\pi\lambda}$ compared to the first term, and we have

$$(-1)^n \mathbf{F}_{i\mu}^{-\frac{3}{2}+n-i\lambda}(-p+i\epsilon) \simeq -ie^{-\pi(\lambda-\mu)}\mathbf{F}_{i\mu}^{-\frac{3}{2}+n-i\lambda}(p) \tag{B.111}$$

$$B^{(+)}_{\text{tree},++,\text{NA}}(k_1,k_2,k_3;\lambda+\varepsilon) \simeq ie^{\pi\varepsilon}\overline{B}^{(-)}_{\text{tree},-+}(k_1,k_2,k_3;\lambda+\varepsilon), \tag{B.112}$$

where we changed variables to $\mu = \lambda + \varepsilon$, and in the last line we used eqs. (B.85) and (B.91) to relate the two amplitudes. The integral over ε can now be evaluated explicitly in a similar manner to give

$$\begin{aligned}
F^{(\lambda)}_{\text{1-loop},++,\text{NA}} &\simeq \frac{5}{36\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_i}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda(p^3+3\chi^2 p+4)}\left[p-\log(p+1)+\frac{p^2+3\chi^2}{4}\left(1-p\log\frac{p+1}{p}\right)\right] \\
&\xrightarrow{p\gg 1} \frac{5}{32\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_i}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda p^2}\left(1-\frac{2+24\log p}{27p}-\frac{5+16\chi^2}{6p^2}\right).
\end{aligned} \tag{B.113}$$

Finally, the tree level contribution to $B^{(\pm)}_{++,\text{A}}$ has been calculated exactly in appendix B.9.4. To obtain an approximate expression we expand eq. (B.204), with a useful identity derived in eq. (B.214), around $\mu = \lambda + \varepsilon$. Assuming $\mu, \lambda \gg 1$ we obtain

$$B^{(\pm)}_{++,\text{A}}(k_1,k_2,k_3) \simeq \frac{1}{16k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{p^{-1}}{\lambda\left(\frac{9}{4}+\varepsilon^2\right)\left(\frac{1}{2}\mp i\varepsilon\right)}\sum_{n=0}^{1}\hat{c}_n F\left(\begin{matrix}1+n, \frac{1}{2}\mp i\varepsilon \\ \frac{3}{2}\mp i\varepsilon\end{matrix}; -\frac{1}{p}\right), \tag{B.114}$$

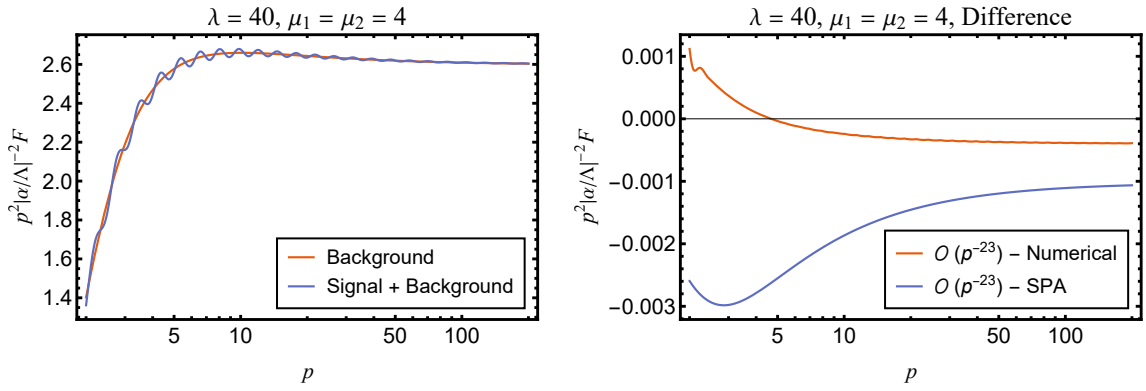


Figure B.12: (Left) The one-loop smooth background coming from the resonance limit of F_{-+}, $F_{++,\mathrm{NA}}$ and $F_{++,\mathrm{A}}$ for $\lambda = 40$ and $\mu_1 = \mu_2 = 4$, rescaled by p^2. (Right) The difference between three different calculations: (i) the analytic sum over poles up to $\mathcal{O}(p^{-23})$ (see section B.5.3 and appendix B.9.5); (ii) numerical evaluation of eq. (B.41); (iii) the SPA result eq. (B.118). For p away from 2, both the numerical result and the improved SPA result agree with the analytic sum within 0.1%. The deviation in the orange line at large p is due to the $\widetilde{F}^{(s_1,s_2\mu_{12},k)}_{++,\mathrm{A,1\text{-}loop}}$ series which, in principle, requires a full resummation even at $\mathcal{O}(p^{-2})$. In making the plot, we include terms up to $k = 40$. The details are discussed at the end of appendix B.9.5.

with $\hat{c}_0 \coloneqq p^2 - 2Q$, $\hat{c}_1 \coloneqq Q$, and $Q \coloneqq \frac{1}{4}(p^2 - \chi^2)$. It turns out that the function

$$\frac{1}{\frac{1}{2} \mp i\varepsilon} F\left(\begin{matrix} 1+n, \frac{1}{2} \mp i\varepsilon \\ \frac{3}{2} \mp i\varepsilon \end{matrix}; -\frac{1}{p}\right) \tag{B.115}$$

is meromorphic in ε with poles at $1/2 \pm i\varepsilon + n = 0$ $(n = 0, 1, 2\cdots)$, and decays as $\mathcal{O}(\varepsilon^{-1})$ as $\varepsilon \to \infty$. The integration over ε can thus be closed on either the upper or the lower half plane:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\mathrm{d}\varepsilon}{\left(\frac{9}{4} + \varepsilon^2\right)\left(\frac{1}{2} \mp i\varepsilon\right)} F\left(\begin{matrix} 1+n, \frac{1}{2} \mp i\varepsilon \\ \frac{3}{2} \mp i\varepsilon \end{matrix}; -\frac{1}{p}\right) &= \frac{\pi}{3} F\left(\begin{matrix} 1+n, 2 \\ 3 \end{matrix}; -\frac{1}{p}\right) \\ &= \frac{2\pi}{3} \begin{cases} 1 - p \log \dfrac{p+1}{p}, & n = 0, \\ -\dfrac{p^2}{p+1} + p^2 \log \dfrac{p+1}{p}, & n = 1. \end{cases} \end{aligned} \tag{B.116}$$

We thus have

$$F^{(\lambda)}_{\text{1-loop},++,\text{A}} \simeq -\frac{5}{72\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_i}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda(p^3+3\chi^2 p+4)}\left[(p^2+3\chi^2)\left(1-p\log\frac{p+1}{p}\right)+\frac{p^2-\chi^2}{p+1}\right]$$
$$\xrightarrow{p\gg 1} -\frac{5}{48\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_i}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda p^2}\left(1-\frac{8}{9p}+\frac{5-16\chi^2}{6p^2}+\cdots\right). \tag{B.117}$$

Combining eqs. (B.108), (B.113) and (B.117) gives an approximate expression for the smooth background

$$F_{\text{1-loop,bkg}} \simeq \frac{5}{18\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_1,\epsilon_2=\pm}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda(p^3+3\chi^2 p+4)}$$
$$\times\left[p-\frac{1}{2}\log\frac{p+1}{p-1}-\frac{p^2+3\chi^2}{4}\left(1-\frac{p}{2}\log\frac{p+1}{p-1}\right)-\frac{p^2-\chi^2}{4(p+1)}\right]$$
$$\xrightarrow{p\gg 1} \frac{5}{24\pi}\frac{|\alpha|^2}{\Lambda^2}\frac{\prod_{\epsilon_1,\epsilon_2=\pm}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{\lambda}p^{-2}\left(1+\frac{4}{9p}-\frac{5+8\chi^2}{3p^2}+\cdots\right). \tag{B.118}$$

The shape of the background is shown in figure B.12 and the left panel of figure B.13. One can see that the resonance background is an order of magnitude larger than the signal for $\lambda = 40$ and $\mu_1 = \mu_2 = 4$. It will thus be detected before the signal in the one-loop dominated scenario. Fortunately, since the background is a smooth function of p, the non-analytic signal can be separated from it as long as the experimental error is smaller than the amplitude of the oscillations. This detectability issue will be further discussed in section B.5.4.

B.5.3 Beyond stationary phase approximation

Besides the SPA result eqs. (B.57) and (B.102), the spectral representation also allow us to expand the exact result for the one-loop amplitude as an infinite sum over analytic terms,

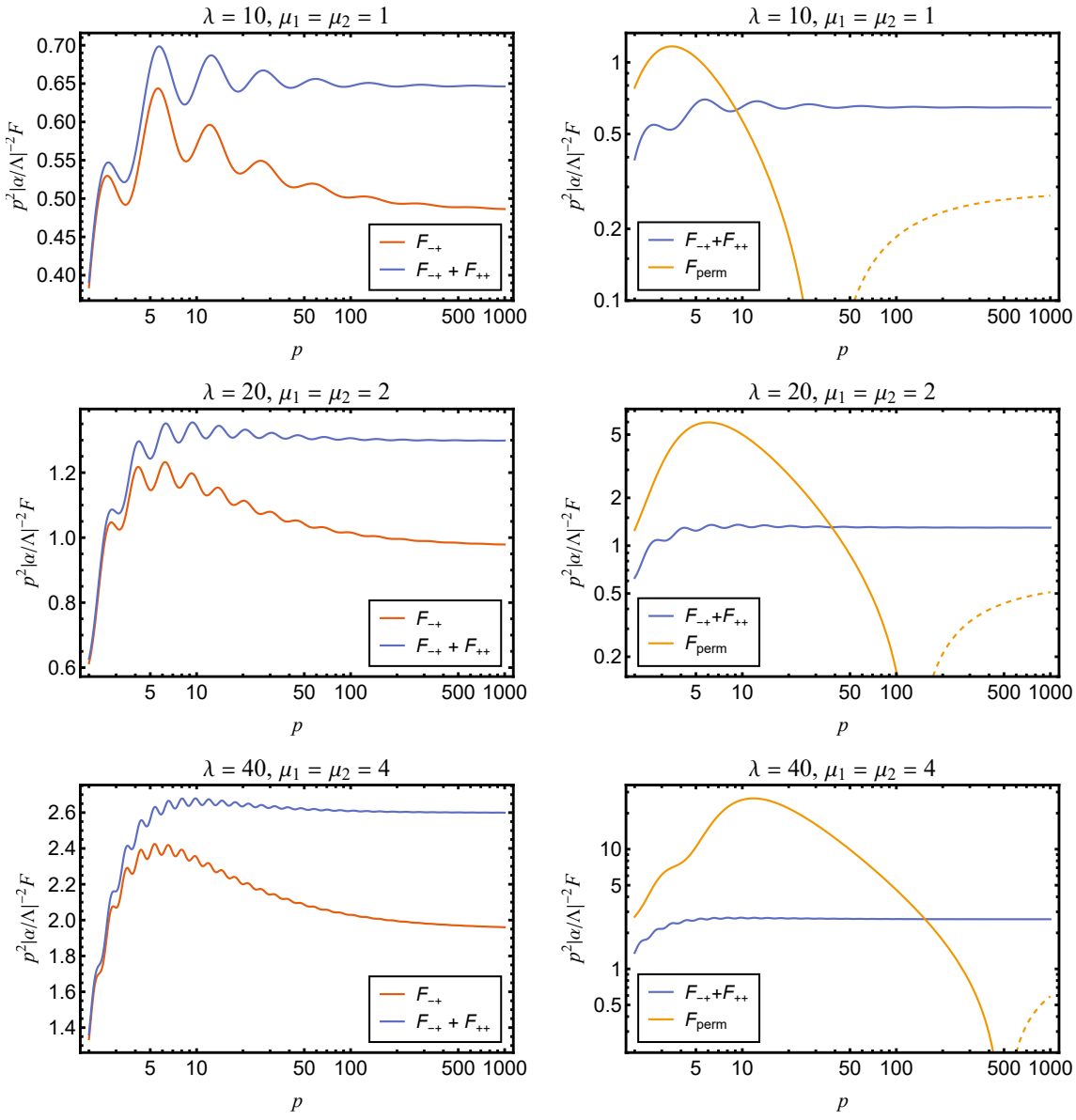


Figure B.13: Different contributions to the one-loop correlators for various λ and μ_1, μ_2, rescaled by p^2. (Left) The contribution from F_{-+} and F_{++}. (Right) The contribution from the two permuted diagrams where the soft momentum k_3 is contracted with the 3-point vertex in figure B.8. Negative values are shown in dashed lines.

which is effectively an expansion in $1/p$. To see how this works, we start with the integration for B_{-+} given by

$$B^{(-)}_{-+,\text{1-loop}}(k_1,k_2,k_3) = \int_0^\infty \rho^{\text{dS}}_{\mu_1\mu_2}(\mu) B^{(-)}_{-+,\text{tree}}(k_1,k_2,k_3;\mu)\,\mathrm{d}\mu, \tag{B.119}$$

where $B^{(-)}_{-+,\text{tree}}$ is given by eq. (B.85).

There is one difficulty in applying the residue theorem to eq. (B.119): in the complex plane the integrand blows up as $\operatorname{Im}\mu \to \pm\infty$. This is analogous to the situation in flat space where the integrand is a trigonometric function, i.e $\cos(x) \sim e^{ix} + e^{-ix}$. The integrand blows up in both the upper and lower half plane, but once we split the integrand up into complex exponentials the contour can be closed in both terms individually.

To achieve such separation for the hypergeometric function we rewrite $\mathbf{F}^\rho_{i\mu}(p)$ in eq. (B.85) using the following identity [71, eq. (15.8.2)]:

$$\mathbf{F}^\rho_{i\mu}(p) = \frac{1}{2\sqrt{\pi}}\Big[2^{2i\mu}\Gamma(i\mu)\Gamma(\rho-i\mu)\widetilde{\mathbf{F}}^{\rho-i\mu}_{i\mu}(p) + (\mu\to-\mu)\Big], \tag{B.120}$$

where

$$\widetilde{\mathbf{F}}^s_{i\mu}(p) \coloneqq \left(\frac{p-1}{2}\right)^{-s} F\left(\begin{matrix}s, \frac{1}{2}-i\mu \\ 1-2i\mu\end{matrix}; \frac{2}{1-p}\right). \tag{B.121}$$

Since the integrand in eq. (B.119) is an even function in μ, we then have

$$B^{(-)}_{-+,\text{1-loop}}(k_1,k_2,k_3) = \int_{-\infty}^{\infty} \rho^{\text{dS}}_{\mu_1\mu_2}(\mu)\widetilde{B}^{(-)}_{-+,\text{tree}}(k_1,k_2,k_3;\mu)\,\mathrm{d}\mu, \tag{B.122}$$

$$\begin{aligned}\widetilde{B}^{(-)}_{-+,\text{tree}}(k_1,k_2,k_3;\mu) \coloneqq &-\frac{1}{4\sqrt{\pi}k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{2^{2i\mu}\lambda e^{\pi\lambda}}{\left(\frac{3}{2}+i\lambda\right)^2+\mu^2}\sum_{n=0}^{2}\frac{d_n}{2^n}\widetilde{\mathbf{F}}^{-\frac{3}{2}+n+i\lambda-i\mu}_{i\mu}(p)\\ &\times\Gamma\left(\begin{matrix}i\mu,-\frac{3}{2}+n+i\lambda-i\mu,\frac{1}{2}-i\lambda+i\mu,\frac{1}{2}-i\lambda-i\mu\\ 1-i\lambda\end{matrix}\right).\end{aligned} \tag{B.123}$$

The integrand in eq. (B.122) has improved asymptotic behavior compared to the one in eq. (B.119). In particular, for $p>1$, $\operatorname{Re}\mu=\mu_*$ and $\operatorname{Im}\mu=M$, we have the following asymptotic bound as $M\to+\infty$:[11]

$$\rho^{\text{dS}}_{\mu_1\mu_2}(\mu)\widetilde{B}^{(-)}_{-+,\text{tree}}(k_1,k_2,k_3;\mu)\sim\mathcal{P}(\lambda,\mu,\mu_1,\mu_2)M^{-\frac{3}{2}-i\lambda}\left(p+\sqrt{p^2-1}\right)^{-M}, \tag{B.124}$$

where

$$\mathcal{P}(\lambda,\mu,\mu_1,\mu_2)=\frac{\sin\pi\left(\frac{3}{2}+i\mu\right)}{\sin\pi\left(\frac{1}{2}-i\lambda+i\mu\right)\prod_{\epsilon_1,\epsilon_2=\pm}\sin\pi\left[\frac{3}{4}+\frac{i}{2}(\mu+\epsilon_1\mu_1+\epsilon_2\mu_2)\right]} \tag{B.125}$$

is bounded if μ does not coincide with the zeros of any sine functions in the denominator. Therefore, we can close the contour in the upper μ-plane.

The analytic structure of the integrand in eq. (B.122) is illustrated in figure B.14 and the position of the poles are listed in table B.2, where the simplified notation $\Delta\mu\coloneqq|\mu_1-\mu_2|$ is

[11] For the gamma functions, the asymptotic property $\Gamma(a-i\mu)\sim(-i\mu)^a\Gamma(-i\mu)$ and the identity $\Gamma(a+i\mu)\Gamma(1-a-i\mu)=\pi\csc\pi(a+i\mu)$ are used. For $\widetilde{\mathbf{F}}^{\rho-i\mu}_{i\mu}$, the asymptotic expansion is given by eq. (B.189), where we leave its derivation to appendix B.9.2.

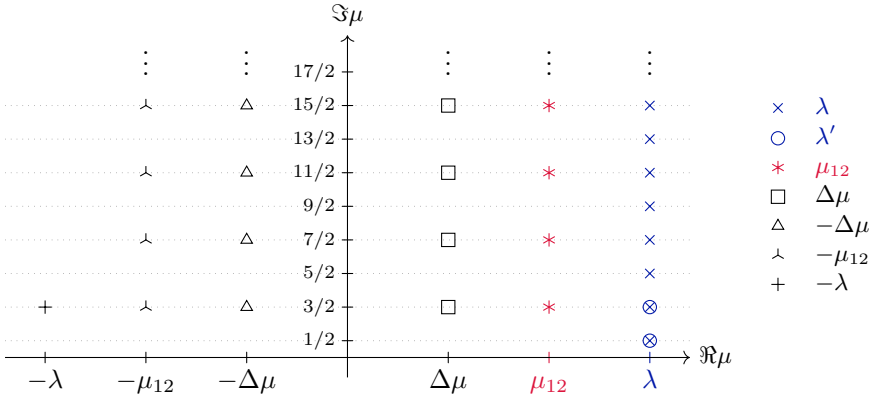


Figure B.14: The analytic structure of the integrand in eq. (B.122) in the upper half plane. The two colored columns are poles with residues that are not exponentially suppressed. Given the bound eq. (B.124), the integral can be represented as a sum over the residues of all these poles for $p > 1$.

Series ($\mathrm{Re}\,\mu$)	source		μ	residues
λ	$\widetilde{B}^{(-)}_{-+,\mathrm{tree}},$	$\Gamma\left(\frac{1}{2}-i\lambda+i\mu\right)$	$\lambda+\left(\frac{1}{2}+k\right)i$	$\mathcal{O}(1)$
λ'	$\widetilde{B}^{(-)}_{-+,\mathrm{tree}},$	$\Gamma\left(-\frac{3}{2}+n+i\lambda-i\mu\right)$	$\lambda+\frac{1}{2}i,\ \lambda+\frac{3}{2}i$	$\mathcal{O}(1)$
$-\lambda$	$\widetilde{B}^{(-)}_{-+,\mathrm{tree}},$	$\frac{1}{(3/2+i\lambda)^2+\mu^2}$	$-\lambda+\frac{3}{2}i$	$\mathcal{O}\left(e^{-\pi\lambda}\right)$
μ_{12}	$\rho^{\mathrm{dS}}_{\mu_1\mu_2},$	$\Gamma\left[\frac{3}{4}+\frac{i}{2}(\mu-\mu_{12})\right]$	$\mu_{12}+\left(\frac{3}{2}+2k\right)i$	$\mathcal{O}(1)$
$-\mu_{12}$	$\rho^{\mathrm{dS}}_{\mu_1\mu_2},$	$\Gamma\left[\frac{3}{4}+\frac{i}{2}(\mu+\mu_{12})\right]$	$-\mu_{12}+\left(\frac{3}{2}+2k\right)i$	$\mathcal{O}(e^{-\pi\mu_{12}})$
$\Delta\mu$	$\rho^{\mathrm{dS}}_{\mu_1\mu_2},$	$\Gamma\left[\frac{3}{4}+\frac{i}{2}(\mu-\Delta\mu)\right]$	$\Delta\mu+\left(\frac{3}{2}+2k\right)i$	$\mathcal{O}(e^{-2\pi\mu_2})$
$-\Delta\mu$	$\rho^{\mathrm{dS}}_{\mu_1\mu_2},$	$\Gamma\left[\frac{3}{4}+\frac{i}{2}(\mu+\Delta\mu)\right]$	$-\Delta\mu+\left(\frac{3}{2}+2k\right)i$	$\mathcal{O}(e^{-\pi\mu_{12}})$

Table B.2: The position of poles of the integrand in eq. (B.122) in the upper half plane and their sources ($n = 0, 1, 2$ and $k \in \mathbb{N}$). The residues at these poles are estimated logarithmically for $\lambda, \mu_1, \mu_2, \lambda - \mu_{12} \gg 1$, where $\mu_1 > \mu_2$ is assumed w.l.o.g. For the logarithmic estimation of all ranges of λ, see table B.5.

used. The p-dependence of each pole for large p can be read off from (B.121) directly:

$$\mathrm{Res}_{\mu_*+i\delta}\,\widetilde{B}^{(-)}_{-+,\mathrm{tree}} \xrightarrow{p\to\infty} p^{3/2-\delta-i(\lambda-\mu_*)}, \tag{B.126}$$

where the imaginary part of the exponent is given by $\lambda-\mu_*$, while poles with larger imaginary part, δ, are more suppressed in the squeezed limit.

For $\lambda,\mu_1,\mu_2,\lambda-\mu_{12}\gg 1$, one can use the Stirling approximation on all the Γ functions in eq. (B.122). By doing so, we find that all but three series of poles are exponentially suppressed, where the unsuppressed poles have either $\mu_*=\lambda$ or $\mu_*=\mu_{12}$. The former case corresponds to the resonance limit, and contributes to the bispectrum as[12]

$$F^{(-,\lambda)}_{-+,\text{1-loop}} \xrightarrow{p\to\infty} p^{-2},\quad p^{-3},\quad p^{-3}\log p,\quad p^{-4},\quad \cdots, \tag{B.127}$$

while the latter case corresponds to the threshold limit, and contributes to the bispectrum as

$$F^{(-,\mu_{12})}_{-+,\text{1-loop}} \xrightarrow{p\to\infty} p^{-3-i(\lambda-\mu_{12})},\quad p^{-5-i(\lambda-\mu_{12})},\quad p^{-7-i(\lambda-\mu_{12})},\quad \cdots. \tag{B.128}$$

In particular, when the contribution from one pole is considered, the rest of eq. (B.122) can be approximated as in eq. (B.53). This further justifies the use of eq. (B.99) as a leading-order approximation.

The same technique can also be used to evaluate $B^{(\pm)}_{++,\mathrm{NA}}$ and $B^{(\pm)}_{++,\mathrm{A}}$ for the one-loop model. For $B^{(\pm)}_{++,\mathrm{NA}}$, the following identity is used to obtain an integrand with an improved asymptotic

[12] The $\log p$ at p^{-3} comes from the overlapping pole at $\mu=\lambda+\frac{3}{2}i$, while the other apparent overlapping pole at $\mu=\lambda+\frac{1}{2}i$ does not contain $\log p$. See appendix B.9.5 for more details.

behavior:

$$\mathbf{F}^{\rho}_{i\mu}(-p+i\epsilon)=\frac{e^{-\rho\pi i}}{2\sqrt{\pi}}\Big[2^{2i\mu}e^{-\pi\mu}\Gamma(i\mu)\Gamma(\rho-i\mu)\widetilde{\mathbf{F}}^{\rho-i\mu}_{i\mu}(p)+(\mu\to-\mu)\Big], \tag{B.129}$$

while for $B^{(\pm)}_{++,\mathrm{A}}$, eq. (B.204) is already in the desired form. Eventually:

$$B^{(\pm)}_{++,\mathrm{NA/A,1\text{-}loop}}(k_1,k_2,k_3)=\int_{-\infty}^{\infty}\rho^{\mathrm{dS}}_{\mu_1\mu_2}(\mu)\widetilde{B}^{(\pm)}_{++,\mathrm{NA/A,tree}}(k_1,k_2,k_3;\mu)\,\mathrm{d}\mu, \tag{B.130}$$

$$\begin{aligned}\widetilde{B}^{(\pm)}_{++,\mathrm{NA,tree}}(k_1,k_2,k_3;\mu)\coloneqq&-\frac{i}{4\sqrt{\pi}k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{2^{2i\mu}\lambda e^{-\pi\mu}}{\left(\frac{3}{2}\mp i\lambda\right)^2+\mu^2}\sum_{n=0}^{2}\frac{d_n}{2^n}\widetilde{\mathbf{F}}^{-\frac{3}{2}+n\mp i\lambda-i\mu}_{i\mu}(p)\\&\times\Gamma\left(\begin{matrix}i\mu,-\frac{3}{2}+n\mp i\lambda-i\mu,\frac{1}{2}\pm i\lambda+i\mu,\frac{1}{2}\pm i\lambda-i\mu\\1\pm i\lambda\end{matrix}\right),\end{aligned} \tag{B.131}$$

$$\begin{aligned}\widetilde{B}^{(\pm)}_{++,\mathrm{A,tree}}(k_1,k_2,k_3;\mu)\coloneqq&\frac{i}{8k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{\mu\left[\left(\lambda^2-\mu^2-\frac{9}{4}\right)^2+9\lambda^2\right]}\sum_{n=0}^{2}\frac{c_n^{\pm}}{2^n}\mathbf{H}^{\pm i\lambda}_{n,i\mu}(p)\\=&\frac{i}{8k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{\mu\left[\left(\lambda^2-\mu^2-\frac{9}{4}\right)^2+9\lambda^2\right]}\sum_{n=0}^{2}\frac{c_n^{\pm}}{2^n}\\&\times\sum_{m=0}^{\infty}\frac{(-)^m(m+1)_n\left(\frac{1}{2}+i\mu\right)_m}{\left(\frac{1}{2}+m\pm i\lambda+i\mu\right)(1+2i\mu)_m}\widetilde{\mathbf{F}}^{1+n+m}_{i\mu}(p).\qquad(p>3)\end{aligned} \tag{B.132}$$

The full results for both the signal and all the backgrounds have an intricate form and can be found in appendix B.9.5.

B.5.4 Combinatorial backgrounds from permuted diagrams

Finally, there are also combinatorial backgrounds from permuted diagrams, where the soft momentum k_3 is contracted with the 3-point vertex in figure B.8. These contributions can be

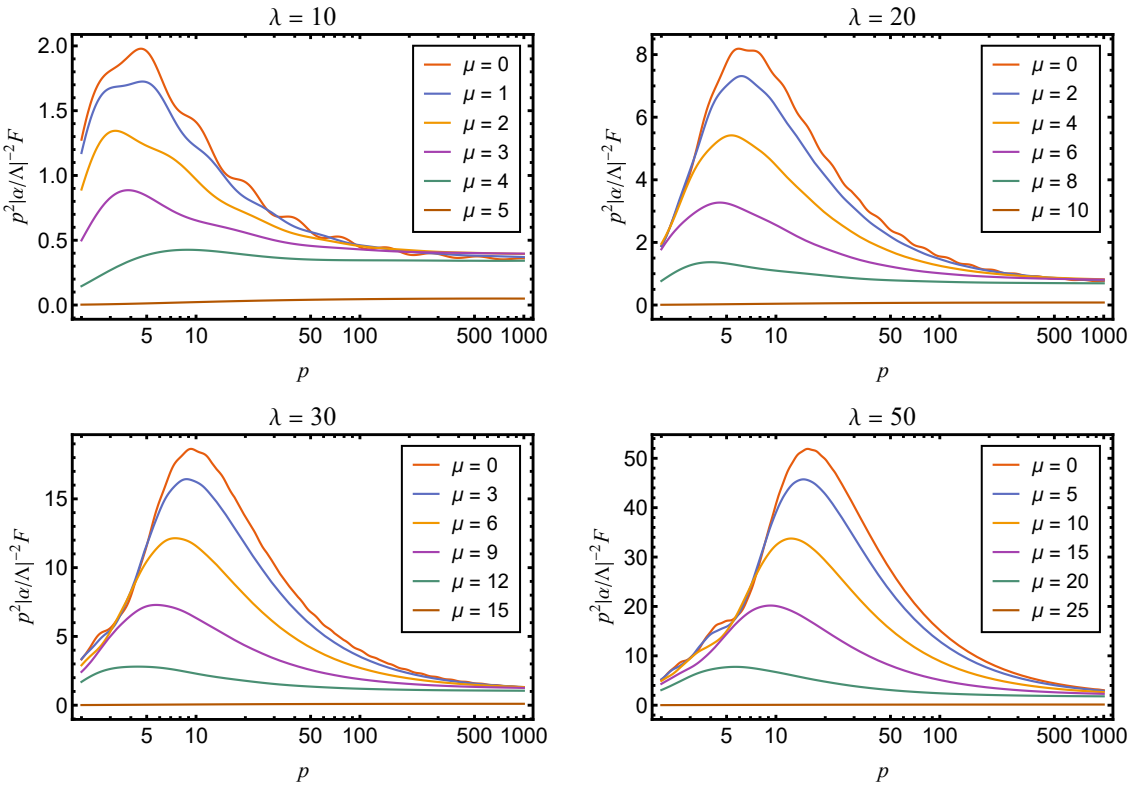


Figure B.15: The full one-loop bispectrum including the combinatorial background for various λ and $\mu_1 = \mu_2 = \mu$ rescaled by p^2. The bumps in the diagrams are artifacts of the p^2 rescaling.

obtained from B_{-+} and B_{++} by permuting $k_3 \to k_1, k_2$. In terms of p and χ:

$$\begin{cases} p' = \dfrac{p-\chi+2}{p+\chi}, \\ \chi' = -\dfrac{p-\chi-2}{p+\chi} \end{cases} \quad \text{for } k_3 \to k_1; \qquad \begin{cases} p' = \dfrac{p+\chi+2}{p-\chi}, \\ \chi' = \dfrac{p+\chi-2}{p-\chi} \end{cases} \quad \text{for } k_3 \to k_2. \tag{B.133}$$

Unfortunately, since large p corresponds to $p' \to 1$, both SPA and the analytic method described in previous sections do not apply. Instead, one needs to expand the tree-level amplitudes around $p' \to 1$. According to the analysis in section B.4.2, in this limit, all stationary phase points either disappear or approach fixed values. In particular, the creation and annihilation time become coincident for B_{-+}. Consequently, $F(k_2, k_1, k_3) + F(k_3, k_1, k_2)$ does not oscillate but falls off as p^{-2} in the squeezed limit as local interactions, i.e. the permuted diagrams only contribute to the analytic part of the amplitude at large p.

Nevertheless, this expansion is only of theoretical interest, and we must resort to numerically integrating the exact tree level amplitudes over the spectral density. This is because unless $p \gg \lambda, \mu$, the number of terms needed to be included in the expansion is $\mathcal{O}(\lambda/p, \mu/p)$. For $\lambda, \mu \sim \mathcal{O}(10)$, this would require either $\mathcal{O}(10)$ numbers of higher order terms or $p \gg \mathcal{O}(10)$ as illustrated in the right panel of figure B.13.

It is difficult in practice to detect the NG, and one has to identify the non-analyticity using data from moderate values of p, where the theory predicts a more complicated p-dependence as shown in figure B.15.

Given the large analytic backgrounds from both the resonance limit and permuted diagrams, it is natural to ask whether the non-analytic signal can be separated from them. This turns out not to be a serious problem as both backgrounds are mostly smooth functions in p

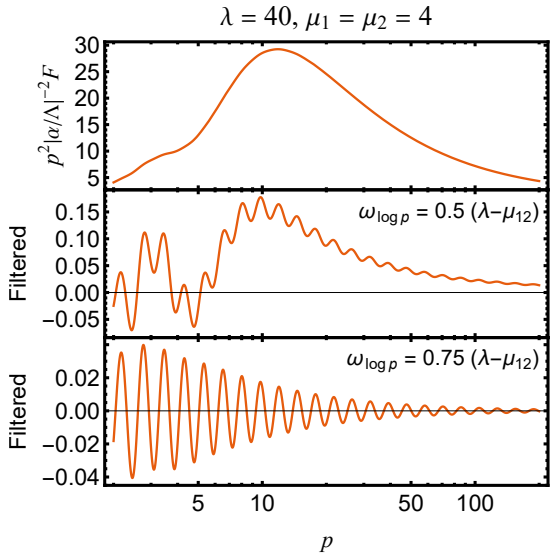


Figure B.16: An example of applying high-pass filter to the full one-loop bispectrum. The Blackman-Harris window has been used in the filter. For details of how to implement the high pass filter, the reader may consult [56]. With a cutoff frequency at $0.5(\lambda-\mu_{12})$, oscillatory patterns from both the signal and the permuted diagrams can pass the filter, while a slightly higher cutoff frequency will only allow the oscillatory signal to pass.

compared to the highly-oscillatory signal. As shown in ref. [56], one can apply a high-pass filter to extract the signal from the curve (figure B.16). To observe the signal we only require that the experimental sensitivity, $F_{\rm min}$, be sufficient to resolve the oscillations, i.e.

$$|\widetilde{F}_{\rm osc}| \gtrsim F_{\rm min}, \tag{B.134}$$

which is not affected by the existence of backgrounds, except in the sense that the large signal to background ratio, combined with the fact that we are yet to identify any NGs, can set an upper bound on the signal.

B.5.5 Higher order corrections at one-loop

At one-loop there are also higher order contributions at each $\mathcal{O}\big(\alpha^{2n}\big)$ from inserting equal numbers of $\chi_1\chi_2$ and $\bar{\chi}_1\bar{\chi}_2$ vertices. Resumming these contributions is equivalent to solving the full linearized equations of motion for $\chi_{1,2}$:

$$\left[-\Box + \mathbf{M}^2(\eta) + \frac{9}{4}\right]\begin{pmatrix}\chi_1\\ \bar{\chi}_2\end{pmatrix} = 0, \qquad \mathbf{M}^2(\eta) = \begin{bmatrix} \mu_1^2 & \bar{\alpha}(-\eta)^{i\lambda} \\ \alpha(-\eta)^{-i\lambda} & \mu_2^2 \end{bmatrix}. \tag{B.135}$$

In momentum space:

$$\left[\partial_\eta^2 - \frac{2}{\eta}\partial_\eta + k^2 + \frac{\mathbf{M}^2(\eta) + 9/4}{\eta^2}\right]\begin{pmatrix}\chi_1\\ \bar{\chi}_2\end{pmatrix} = 0. \tag{B.136}$$

It turns out that it is more intuitive to work with the proper time $t \coloneqq -\log(-\eta)$, so that the equation becomes

$$\left[\partial_t^2 + k^2\eta^2(t) + \mathbf{M}^2(t)\right]\begin{pmatrix}\hat{\chi}_1\\ \hat{\bar{\chi}}_2\end{pmatrix} = 0, \qquad \hat{\chi}_{1,2} \coloneqq (-\eta)^{-3/2}\chi_{1,2}. \tag{B.137}$$

Consider the case when $\lambda, \mu_1, \mu_2 \gg 1$, so that the expansion of the universe can be neglected near a pivot time η_*. In this case, the general solution is given by

$$\begin{bmatrix}\hat{\chi}_1(t)\\ \hat{\bar{\chi}}_2(t)\end{bmatrix} \simeq \sum_{n=1}^{4} c_n \begin{bmatrix}\hat{\chi}_1^{(n)} e^{-i(\omega_n + \lambda/2)t}\\ \hat{\bar{\chi}}_2^{(n)} e^{-i(\omega_n - \lambda/2)t}\end{bmatrix}, \tag{B.138}$$

where ω_n are the four roots of the quartic equation

$$\left[\left(\omega+\frac{\lambda}{2}\right)^2-\omega_{01}^2\right]\left[\left(\omega-\frac{\lambda}{2}\right)^2-\omega_{02}^2\right]=|\alpha|^2\,,\qquad \omega_{01,02}\coloneqq\sqrt{k^2\eta_*^2+\mu_{1,2}^2}, \tag{B.139}$$

and $\hat{\chi}_1^{(n)}$, $\hat{\bar{\chi}}_2^{(n)}$ satisfy

$$\begin{bmatrix}\left(\omega_n+\frac{\lambda}{2}\right)^2-\omega_{01}^2 & -\bar{\alpha}\\ -\alpha & \left(\omega_n-\frac{\lambda}{2}\right)^2-\omega_{02}^2\end{bmatrix}\begin{bmatrix}\hat{\chi}_1^{(n)}\\ \hat{\bar{\chi}}_2^{(n)}\end{bmatrix}=0. \tag{B.140}$$

The structure of the solutions to eq. (B.139) for different values of λ, ω_{01} and ω_{02} are illustrated in figure B.17.

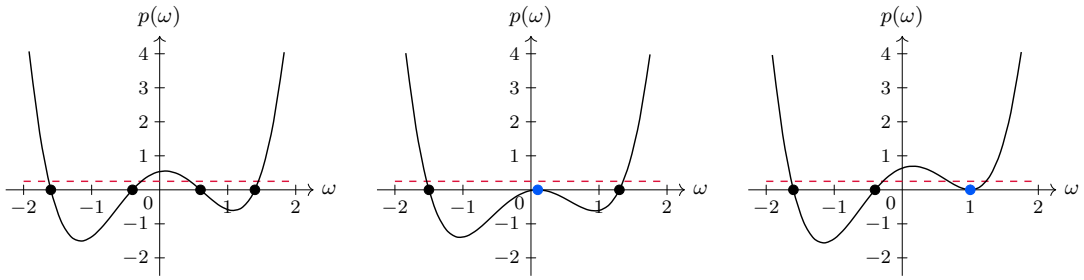


Figure B.17: The condition for parametric resonance can be understood by the structure of the solutions to the quartic equation eq. (B.139). The solutions occur when the quartic crosses the dashed red line, and parametric resonance occurs when two of the solutions are complex. (a) For small α and general values of λ, ω_1, ω_2, there are 4 real solutions and the corrections to ω_k are $\mathcal{O}(\alpha^2)$; (b) when $\lambda=\omega_1+\omega_2$, there are only two real solutions as the degenerate ω_k (blue) receives an $\mathcal{O}(\alpha)$ imaginary part (parametric resonance); (c) when $\lambda=|\omega_1-\omega_2|$ or one of ω_{01}, ω_{02} is zero, there are 4 real solutions where the degenerate ω_k (blue) receives an $\mathcal{O}(\alpha)$ real part.

Eq. (B.139) implies Parametric Resonance (PR) at some time during inflation [35]. This

happens when any of ω_k has an imaginary part, i.e. $\lambda \simeq \omega_{01} + \omega_{02}$ or

$$-k\eta \simeq -k\eta_{\rm R} \coloneqq \frac{\prod_{\epsilon_1,\epsilon_2=\pm}\sqrt{\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2}}{2\lambda}. \tag{B.141}$$

Physically, this corresponds to the time when λ can produce a $\chi_1\chi_2$ pair on-shell. Around this time, ω_2 and ω_3 become nearly degenerate and are approximately given by

$$\omega_{2,3} \simeq -\frac{\omega_{01}-\omega_{02}}{2} \pm \frac{1}{2}\sqrt{\left(\lambda-\omega_{01}-\omega_{02}\right)^2 - \frac{|\alpha|^2}{\omega_{01}\omega_{02}}}. \tag{B.142}$$

Therefore, the parametric oscillator is in resonance when

$$|\alpha|^2 \gtrsim \omega_{01}\omega_{02}\,|\lambda-\omega_{01}-\omega_{02}| \simeq \frac{\prod_{\epsilon_1,\epsilon_2=\pm}(\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2)^2}{4[\lambda^2-(\mu_1^2-\mu_2^2)^2]}\delta t^2. \tag{B.143}$$

where we have plugged in $\eta = \eta_{\rm R}e^{-\delta t}$. For α not too large, PR only persists over a finite period of time. During this period, the mode function is amplified by a factor $\mathcal{A}_{\rm R}^{1/2}$, where:

$$\log\mathcal{A}_{\rm R} \simeq 2\int_{t_{\rm min}}^{t_{\rm max}}\omega_2(t)\,{\rm d}t \simeq \frac{2\pi\lambda\,|\alpha|^2}{\prod_{\epsilon_1,\epsilon_2=\pm}(\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2)}. \tag{B.144}$$

Consequently, for our calculations to be valid we should require this amplification to be small:

$$\log\mathcal{A}_{\rm R} \ll 1, \qquad \text{i.e. } |\alpha| \ll \alpha_{\rm PR} \coloneqq \sqrt{\frac{\prod_{\epsilon_1,\epsilon_2=\pm}(\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2)}{2\pi\lambda}}. \tag{B.145}$$

For even larger α, PR may persist till the end of inflation, giving rise to a late-time behavior

$$|f_k(\eta)| \sim (-\eta)^{3/2-\mathrm{Im}\,\omega_n(0)}. \tag{B.146}$$

A simple power counting shows that, if any of $\mathrm{Im}\,\omega_n(0) \geq 1$, i.e.:

$$|\alpha| \gtrsim \alpha_{\max} \coloneqq \sqrt{\max_{-\frac{\lambda}{2}<\omega<\frac{\lambda}{2}} \left[\left(\omega+\frac{\lambda}{2}\right)^2-\mu_1^2\right]\left[\left(\omega-\frac{\lambda}{2}\right)^2-\mu_2^2\right]} \simeq \frac{\lambda^2}{4}, \quad (\mu_1,\mu_2 \ll \lambda), \tag{B.147}$$

there will be a late-time divergence in the perturbative calculation. This divergence will eventually be cut off by the decrease of $\dot{\phi}_0$ over time until it can no longer sustain the resonance.

So far, we have been treating the universe expansion in the adiabatic approximation. This cannot predict the solutions near the beginning and the end of PR, where eq. (B.139) has degenerate roots and different modes evolve into each other. Here we also present the numerical solutions to eq. (B.136) in figure B.18 for various α. This shows that our previous analysis applies qualitatively even beyond the adiabatic approximation.

To see how higher order corrections at one-loop affect the bispectrum, one has to evaluate eq. (B.35) with the corrected mode functions. Since de Sitter symmetry is badly broken by these corrections, the spectral representation method no longer applies. We can still understand its effects in the following two directions:

1. The energy as a function of time is corrected according to eq. (B.139). For $\mu_1 = \mu_2 = \mu$,

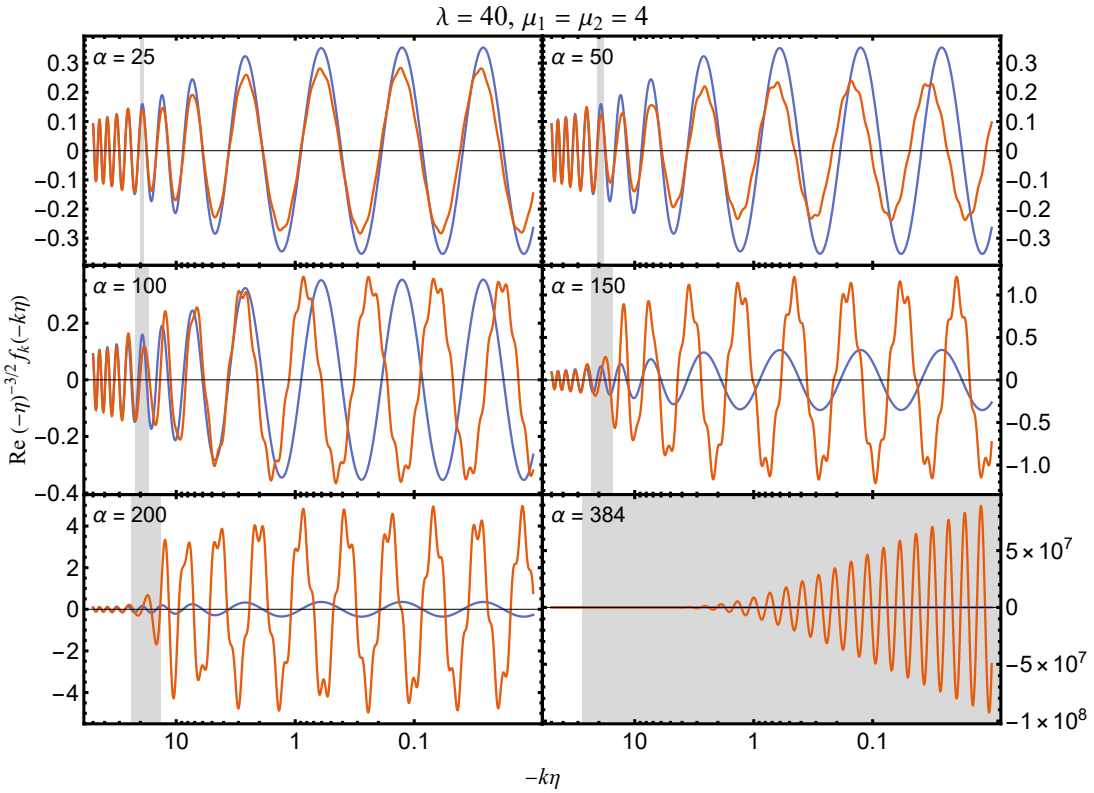


Figure B.18: The numerical mode functions from the resummed equation of motion eq. (B.136) with $\lambda = 40$ and $\mu_1 = \mu_2 = 4$ for various α ($\alpha_{\rm PR} \approx 99$, $\alpha_{\rm max} = 384$). The blue line represents eq. (B.21) ($\alpha = 0$) for comparison. The regions of parametric resonance where eq. (B.139) has complex roots are marked gray. For this particular parameter choice the late-time amplitude first decreases and then increases exponentially as α increases. This may not happen for other choices or parameters, e.g. $\lambda = 40$ and $\mu_1 = \mu_2 = 8$, where the amplitude always increases as α increases.

the corrected energies are explicitly given by

$$\omega_{\pm}^2 = \frac{\lambda^2}{4} + \omega_0^2 \pm \sqrt{\lambda^2\omega_0^2 + |\alpha|^2} \simeq \begin{cases} \left(\frac{\lambda}{2} \pm \omega_0\right)^2 \pm \frac{|\alpha|^2}{2\lambda\omega_0}, & |\alpha| \ll \lambda\omega_0, \\ \frac{\lambda^2}{4} + \omega_0^2 \pm |\alpha|, & |\alpha| \gg \lambda\omega_0. \end{cases} \tag{B.148}$$

This changes the shape of the oscillatory signal as it depends on the phase accumulation during propagation. In particular, the corrected frequency at large p can be evaluated using the late-time limit, where we can set $k = 0$ for both particles and get

$$\begin{aligned} 2\omega_- &= \sqrt{\lambda^2 + 4\mu^2 - 4\sqrt{\lambda^2\mu^2 + |\alpha|^2}} \\ &\simeq \begin{cases} \lambda - 2\mu - \frac{|\alpha|^2}{\lambda\mu(\lambda - 2\mu)}, & |\alpha| \ll \lambda\mu, \\ \lambda - \frac{2|\alpha|}{\lambda}, & \lambda\mu \ll |\alpha| \ll \lambda^2. \end{cases} \end{aligned} \tag{B.149}$$

Such corrections can be neglected if we insist that

$$|\alpha| \ll \lambda\mu. \tag{B.150}$$

2. PR increases the magnitude for both the non-analytic signal and the resonance background. Comparing $\eta_{\rm R}$ in eq. (B.141) with the particle production time from table B.1, we have

$$\frac{\lambda^2 - \mu^2}{2\lambda} < -k\eta_c \coloneqq \frac{\lambda^2 - \mu_{12}^2}{2\lambda}, \qquad \frac{\eta_c}{\eta_{\rm R}} = \sqrt{\frac{\lambda^2 - \mu_{12}^2}{\lambda^2 - \Delta\mu^2}} < 1, \tag{B.151}$$

i.e. particle production always happens after PR. Since there are two propagators with each of them containing two mode functions evaluated at the creation and annihilation

times respectively, both the signal and the background are enhanced by a factor $\mathcal{A}_{\rm R}^2$.

Given the smallness of the one-loop signal, an enhancement from PR may be an attractive phenomenological possibility. Unfortunately a quantitative answer cannot be obtained within the spectral decomposition framework, so we leave this direction to future work.

B.6 Constraints and Signal Size

There are three parameters in the chemical potential model: the chemical potential, λ, the cut-off scale, Λ, and the symmetry breaking coupling, α. The chemical potential gives the maximal energy available for particle creation, and the other two are directly related to the signal strength. Note that value of λ is related to Λ by eq. (B.27) with $\dot{\phi}_0 \approx (60H)^2$ fixed by measurements of the CMB power spectrum. In this section, we will discuss the constraints on λ and α separately and combine them to predict the signal size.

B.6.1 Constraints on λ and beyond de Sitter

For λ, A consistent effective field theory treatment requires that

$$\Lambda > \sqrt{\dot{\phi}_0} \approx 60H. \tag{B.152}$$

This is because for any operator $\mathcal{O}$, there are higher dimension operators of the form

$$\frac{(\nabla\phi)^{2n}}{\Lambda^{4n}}\mathcal{O} \tag{B.153}$$

which give extra contributions suppressed by $\dot{\phi}_0^2/\Lambda^4$ after substituting $\phi = \dot{\phi}_0 t$. To ensure that such higher-dimensional contributions are suppressed, we require that $\Lambda^2 > \dot{\phi}_0$, or equivalently:

$$\lambda := \frac{\dot{\phi}_0}{\Lambda} < \sqrt{\dot{\phi}_0} \approx 60. \tag{B.154}$$

For example, with $\lambda \approx 40$, $\dot{\phi}_0^2/\Lambda^4 = \lambda^4/\dot{\phi}_0^2$ is around 0.2.

It is worth commenting that, with even lower Λ, it is still possible to switch to a new effective field theory where one discards the de Sitter isometry completely and writes down every interactions consistent with spatial Euclidean isometry, spacetime scale invariance and the shift symmetry of $\delta\phi$, as has been shown in [130]. In this case, it turns out that the spectral decomposition method can still be applied provided that there is an emergent de Sitter isometry at leading order, i.e. the sound speeds of the two scalar fields χ_1, χ_2 are equal and all U(1)-breaking terms can be treated perturbatively. Nonetheless, since the energy scale of the kinetic energy, $\sqrt{\dot{\phi}_0} \approx 60H$, is outside the range of such an effective field theory, such theories are losing the power in predicting the size of the chemical potential, as can be seen already by noting that higher-order operators of the form

$$\frac{(\nabla\phi)^{2n}}{\Lambda^{1+4n}} \nabla_\mu \phi J^\mu \tag{B.155}$$

changes the size of the chemical potential by $\mathcal{O}\left(\dot{\phi}_0^{2n}/\Lambda^{4n}\right)$, and becomes increasingly important when $\Lambda < \sqrt{\dot{\phi}_0}$. Because of this, we will not consider this case in details.

B.6.2 Constraints on α

For the symmetry breaking coupling, α, the constraints are more subtle. Firstly, there is an experimental bound on the equilateral NG from CMB data. Depending on the shape, it is given by $f_{\rm NL} < \mathcal{O}(10)$ [117]. Given that the non-analytic signal is small compared to the big analytic background, the $f_{\rm NL}$ constraint on the background will also effectively constrain the size of the signal. Besides this, there are also several constraints from the theory side:

1. Higher order corrections at one-loop: These corrections are similar to figure B.2 but with more insertions of $\exp(\pm i\delta\phi/\Lambda)$. As has been discussed in section B.5.5, the smallness of these corrections require both eqs. (B.150) and (B.145). In later analysis, we will assume that these two conditions are satisfied so that the de Sitter symmetry is not seriously broken.

2. Corrections to the power spectrum: Another constraint comes from the correction to the power spectrum:

$$\langle \delta\phi_{\mathbf{k}} \delta\phi'_{\mathbf{k}} \rangle = \mathcal{P}_\varphi(k) \cdot (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}'). \tag{B.156}$$

   Such corrections can also be evaluated using the spectral representation, with the tree-level contributions shown in figure B.19.

   Here we quote the crude estimation eq. (B.252) for the corrections to the power spectrum,

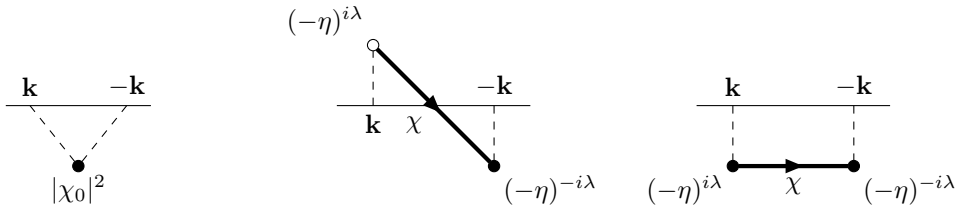


Figure B.19: Corrections to the power spectrum at $\mathcal{O}(\alpha^2)$. For each diagram, there is another diagram with opposite vertex types which is omitted here.

and leave their calculation to appendix B.9.6:

$$
\begin{aligned}
2k^3\delta\mathcal{P}_\phi^{(1,1\text{-loop})}(\mu_1,\mu_2) &\simeq -\frac{|\alpha|^2}{\Lambda^2}\cdot\frac{\pi}{3\lambda^2}\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda),\\
2k^3\delta\mathcal{P}_\phi^{(2,1\text{-loop})}(\mu_1,\mu_2) &\simeq \frac{|\alpha|^2}{\Lambda^2}\cdot\frac{\pi^2}{3\lambda}\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda),\\
2k^3\delta\mathcal{P}_\phi^{(3,1\text{-loop})}(\mu_1,\mu_2) &\simeq \frac{|\alpha|^2}{\Lambda^2}\cdot\frac{1}{\lambda^2}\left\{\frac{\pi}{12}\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda)+0.482\left[\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda)-\lambda\rho'^{\text{Mink}}_{\mu_1\mu_2}(\lambda)\right]\right\}.
\end{aligned}
\tag{B.157}
$$

Since these corrections are proportional to the zeroth-order $1/2k^3$, as required by scale invariance, they can be regarded as a pure renormalization from an observational point of view. However, there are some other physical effects associated with these corrections:

- Figure B.19 (2) produces $\chi_1\chi_2$-pairs that survive until the end of inflation. If it dominates over the power spectrum from the free field theory, the number density of χ_1, χ_2 will be as large as the one of $\delta\phi$, while the energy density will be $\mathcal{O}(\mu/k\eta)$ larger for each k mode.

- Figure B.19 (3) corrects the sound speed of the inflaton fluctuations. This follows from the observation that $\delta\chi$ is coupled only to $\delta\phi'$ at the 2-point vertex in eq. (B.70).

To avoid these complications we require that

$$
2k^3\delta\mathcal{P}_\phi^{(1\text{-loop})} \ll 1. \tag{B.158}
$$

B.6.3 Signal size

In figures B.20 and B.21 we take into account all the theoretical and experimental constraints discussed above. In these plots we compare $F_{\rm min}$ at $p = 2$ with the maximal signal size from various constraints. We did not consider the signal at larger p since for most values of λ and μ_1, μ_2 we consider, the oscillation is at frequency $\lambda - \mu_{12} \sim \mathcal{O}(10)$ and only requires data from a small range of p to be identified. From the diagrams, we see that with $\lambda \gtrsim 40$, a signal observable with $F_{\rm min} = 0.01$ is possible within all the constraints. This parameter space could be probed in future 21 cm experiments.

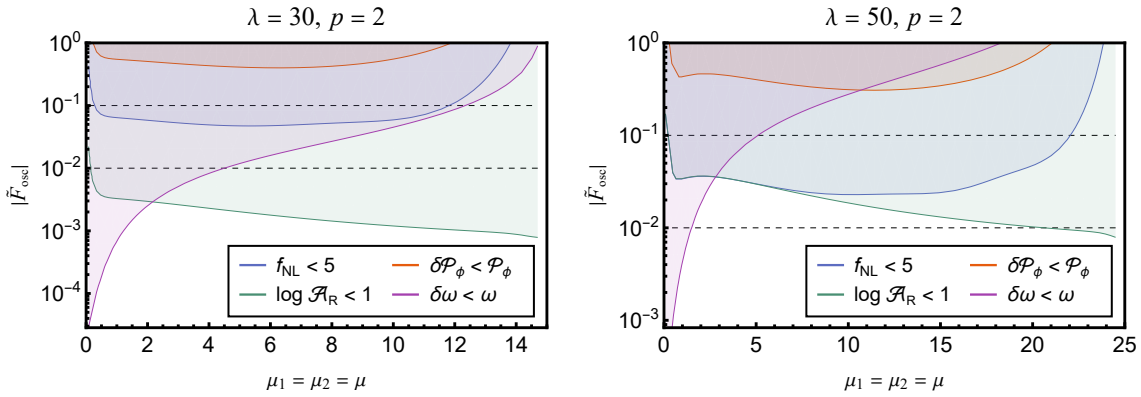


Figure B.20: The maximal signal size at $p = 2$ with various constraints. The 4 constraints shown by the shaded regions are (i) $f_{\rm NL} < 5$ from CMB data [117], (ii) the power spectrum constraint eq. (B.158), (iii) the parametric resonance constraint eq. (B.145) and (iv) the frequency correction constraint eq. (B.150). Note that the last 3 constraints are only to ensure the validity of our calculations and do not imply the breakdown of the theory on their own.

Some further features of these constraints worth noting:

- It turns out that the most stringent constraint comes from the parametric resonance

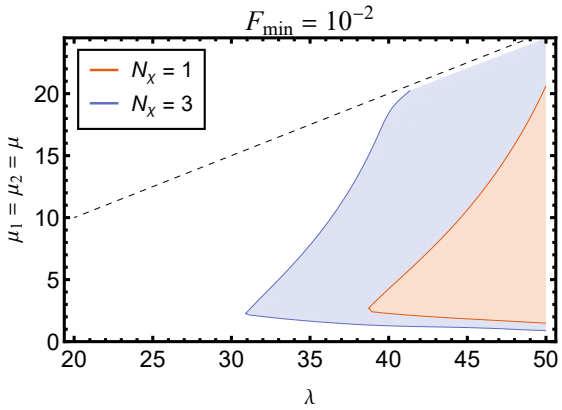


Figure B.21: The parameter space for λ and $\mu_1 = \mu_2 = \mu$ assuming the largest α. The shaded regions are allowed by all the 4 constraints. The dashed line is $\lambda = 2\mu$, above which the signal is exponentially suppressed. N_χ is the multiplicity of χ_1 and χ_2.

constraint eq. (B.145), which combined with $\Lambda = \dot{\phi}/\lambda$ gives

$$\frac{|\alpha|^2}{\Lambda^2} \lesssim \frac{\lambda^5}{2\pi(60)^2} \approx 4.4 \times 10^{-5}\, \lambda^5. \tag{B.159}$$

The maximum allowed value of the interaction strength, α, is thus proportional to λ^5. This explains why the signal size is small for small λ.

- Another important constraint comes from $f_{\rm NL}$, which becomes increasingly important for large λ. Since the background is larger than the signal, the $f_{\rm NL}$ constraint curve is roughly proportional to the signal-to-background ratio

$$r_{\rm sbr} \coloneqq \lim_{p\to\infty} \frac{p|\widetilde{F}_{\rm osc}|}{F_{\rm bkg}} \simeq \sqrt{\frac{\pi}{2}} \frac{\lambda^{5/2}(\mu_1\mu_2)^{1/2}}{\mu_{12}^{5/2}(\lambda-\mu_{12})^{3/2}(\lambda^2-\Delta\mu^2)^{1/2}}, \tag{B.160}$$

following from eqs. (B.59) and (B.118). For $\mu_1, \mu_2 \sim \mathcal{O}(\lambda)$, $r_{\rm sbr} \sim \mathcal{O}(\lambda^{-1})$. This explains why the $f_{\rm NL}$ constraint is more stringent at large λ. Note that for a typical point in the

parameter region, $r_{\rm sbr} \sim \mathcal{O}(0.1)$.

Up to this point, we have been considering χ_1, χ_2 as single complex scalars charged under some $\mathrm{U}(1)_V$ group. This can easily be promoted to the case where χ_1, χ_2 are $\mathbf{N}_\chi$ and $\overline{\mathbf{N}}_\chi$ multiplets under some non-abelian group. In this case, the bispectrum will be larger by a factor of N_χ, i.e. the constraint curves from parametric resonance eq. (B.145) will be multiplied by N_χ, while the curves from $f_{\rm NL}$ and the power spectrum constraint eq. (B.158) will stay unchanged as both the smooth background and the contribution to the power spectrum are also enhanced by N_χ. In figure B.21 we plot the allowed parameter space for $N_\chi = 1, 3$ to illustrate how the added multiplicity can open up new regions of parameter space.

B.7 Application: Colored Higgs in SUSY Orbifold GUTs

Supersymmetric (SUSY) Grand Unified Theories (GUTs) (for a review, see [72]) provide natural candidates for the pair of charged scalars appearing in the loop. It is well known that SUSY requires the existence of a pair of Higgs doublet chiral superfields (H_u, H_d) with conjugate quantum numbers to give masses to up-type and down-type quarks respectively. At energies above the unification scale, $M_{\rm U}$, these Higgses must be embedded into a larger multiplet $(\mathbf{h}_u, \mathbf{h}_d)$ to form a representation of the GUT gauge group.

For example, in the simplest SU(5) GUTs, the $\mathbf{h}_{u,d}$ multiplets are in the $\mathbf{5}$ and $\bar{\mathbf{5}}$ representation respectively, which include both the electroweak Higgs and the colored Higgs bosons:

$$\mathbf{h}_u = \begin{pmatrix} \chi_u \\ \\ h_u \end{pmatrix}, \qquad \mathbf{h}_d = \begin{pmatrix} \chi_d & h_d \end{pmatrix}. \tag{B.161}$$

While being far too heavy for direct detection at colliders, these particles may have been produced on-shell during inflation through the chemical potential mechanism and manifest themselves in non-Gaussianities. In particular, since they carry color charge and hypercharge, they can only be pair-produced by the singlet inflaton, and contribute to non-Gaussianities at the one-loop level.

In the minimal supersymmetric SU(5) GUT the unification of gauge couplings predicts $M_{\rm U} \sim 10^{16}\,{\rm GeV}$, and the masses of the colored Higgs masses are around this scale, which is somewhat higher than the reach of the chemical potential mechanism. Another problem with SUSY GUTs emerges once proton decay is considered: if the colored Higgs bosons obtain their masses through a $\mathbf{h}_u\mathbf{h}_d$ term in the superpotential, they will lead to too rapid proton decay unless their masses are $\mathcal{O}(10)$ times larger than $M_{\rm U}$ [72].

There is however a well-known variant called orbifold SUSY GUTs [46–48] that can naturally avoid the proton decay problem. In these models, the GUT partners are Kaluza-Kelin (KK) excitations from a compact extra dimension. The colored Higgs bosons then appear at the lower compactification scale $M_{\rm C} \sim 10^{14-16}\,{\rm GeV}$ [48, 75–77] and become suitable targets for the cosmological collider.

B.7.1 SUSY inflation and the fifth dimension

In this section we will briefly review the basic setup of SUSY inflation, as well as the extra dimension. For our purposes we will follow the construction in ref. [73, 74] for realizing the inflaton potential in SUSY. As a phenomenological discussion, we will not attempt to address any deeper origin behind the inflaton potential.

Inflation necessarily breaks SUSY as it is incompatible with the de Sitter isometries. Meanwhile, since the electroweak hierarchy problem suggests a SUSY breaking scale $M_{\rm SUSY} \gtrsim$ TeV today, much lower than the Hubble constant $H \sim 10^{14}\,$GeV during a high-scale inflation, we can think of SUSY as approximately restored at the end of the inflation, i.e. SUSY is spontaneously broken by the classical background of the inflaton.

To realize such a scenario, consider an inflaton chiral superfield, Φ, with its scalar component given by

$$\Phi = \frac{\sigma + i\phi}{\sqrt{2}}, \tag{B.162}$$

where ϕ will be the inflaton scalar field, and the other components of Φ are its various superpartners. To main the approximate shift symmetry in ϕ, Φ must appear in the Kähler potential in the combination $\Phi + \bar{\Phi}$. This ensures that ϕ has no potential, even with (spontaneously) broken SUSY. In comparison, the real component σ, called the sinflaton, is only massless under when SUSY is unbroken and acquires a mass typically of $\mathcal{O}(H)$ after spontaneous SUSY breaking [74].

We introduce the inflaton slow-roll potential by adding an extra chiral superfield, S, and coupling it to the inflaton through the following superpotential:

$$\mathcal{L} \supset -\int \mathrm{d}^2\theta\, S f\Big(-i\sqrt{2}\Phi\Big) + \text{c.c.} = -F_S f(\phi) + \cdots + \text{c.c.}, \tag{B.163}$$

where $f(\phi)$ is some model-dependent holomorphic function. After integrating out F_S, the scalar potential, the expectation value of F_S, and the Hubble constant are related by

$$V_{\rm inf}(\phi_0) = |\langle F_S \rangle|^2 = |f(\phi_0)|^2 \simeq 3M_{\rm pl}^2 H^2, \tag{B.164}$$

with $M_{\rm pl} \gtrsim 4 \times 10^4 H$ from eq. (B.1). Finally, to avoid extra complications due to a light sinflaton, or the scalar component of S (denoted as s), one can make them heavy by including the following term in the Kähler potential:

$$\begin{aligned}\mathcal{L} &\supset -\int \mathrm{d}^4\theta \Big[\frac{\kappa_\sigma}{2\Lambda^2}\,|S|^2\,(\Phi+\bar{\Phi})^2 + \frac{\kappa_S}{4\Lambda^2}\,|S|^4\Big] \\ &\supset -\frac{\kappa_\sigma}{\Lambda^2}\,|F_S|^2\,\sigma^2 - \frac{\kappa_S}{\Lambda^2}\,|F_S|^2\,|s|^2\,. \end{aligned} \tag{B.165}$$

This can gives them masses as large as $\mathcal{O}(M_{\rm pl}H/\Lambda) \gtrsim \lambda = (60H)^2/\Lambda$.

Note that for simplification we have considered global supersymmetry only and ignored supergravity corrections. This can be justified because supergravity effects are suppressed by $1/M_{\rm pl}^2$, so they can only shift particle masses by [74]

$$\Delta m^2_{\rm SUGRA} \sim \frac{V_{\rm inf}(\phi_0)}{M_{\rm pl}^2} = \mathcal{O}\big(H^2\big). \tag{B.166}$$

This is negligible if all particles we considered are assumed to be heavy compared to H, except for the inflaton, for which the approximate shift symmetry eliminates such contributions.

Particles	Boundary conditions	KK spectrum	$\mathrm{U}(1)_R$
h_u, h_d	$(+,+)$	$nM_{\rm C}$	0
χ_u, χ_d	$(-,+)$	$(n+1/2)M_{\rm C}$	0
h_u^c, h_d^c	$(+,-)$	$(n+1/2)M_{\rm C}$	2
χ_u^c, χ_d^c	$(-,-)$	$(n+1)M_{\rm C}$	2

Table B.3: The boundary conditions, masses and $\mathrm{U}(1)_R$ charges for the Higgs bosons in the orbifold SUSY GUT ($n = 0, 1, 2, \cdots$). For boundary conditions, plus signs represent essentially Neumann boundary conditions while negative signs represent essentially Dirichlet boundary conditions. The KK spectrum is obtained with a flat 5D geometry.

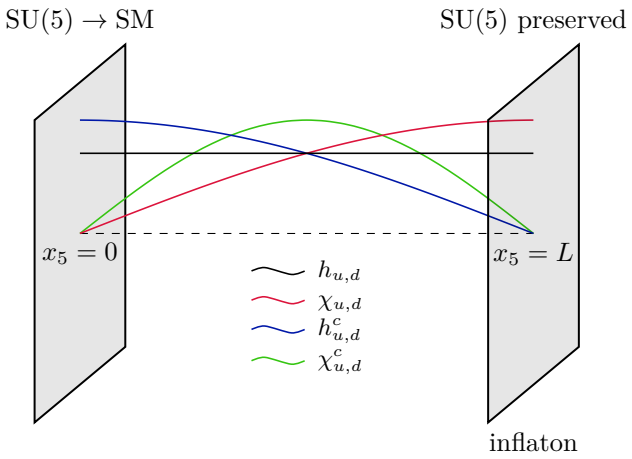


Figure B.22: The 5D spacetime with two boundaries in orbifold SUSY GUTs and the 5D profiles of the lightest KK excitations of the Higgs bosons. The left boundary at $x_5 = 0$ breaks SU(5) down to the standard model gauge group, while the right boundary at $x_5 = L$ preserves SU(5) but reduces the 5D $\mathcal{N} = 1$ SUSY down to 4D $\mathcal{N} = 1$ SUSY. The inflaton is chosen to live on the second boundary where $\chi_{u,d}$ have the largest 5D profiles.

In order to couple the inflaton to the colored Higgs bosons in orbifold SUSY GUTs, the extra dimension should also be taken into account. We consider the simplest model in (4+1)D with the extra dimension being an interval of length $L = \pi/M_C$, suggested by [48, 75–77]. In this model, $\mathbf{h}_u$ and $\mathbf{h}_d$ live in the bulk and, together with $\mathbf{h}^c_u$ and $\mathbf{h}^c_u$, form two hypermultiplets of the 5D $\mathcal{N} = 1$ SUSY. The boundary conditions at the left endpoint are chosen to break SU(5) down to the standard model gauge symmetry, while the boundary conditions at the right endpoint are chosen to respect the full GUT symmetry, but reduce the 5D $\mathcal{N} = 1$ SUSY down to 4D $\mathcal{N} = 1$ SUSY. The boundary conditions are tabulated in table B.3, and illustrated in figure B.22. Note that we assume an approximately flat 5D geometry for simplicity, which can easily be generalized to warped 5D geometries if necessary.

While the profile of the inflaton in the extra dimension is unclear, here we discuss a simple possibility that it lives on one of the 4D boundaries as in ref. [49]. To maximize the couplings,

we put the inflaton on the right boundary where the 5D profiles of χ_u, and χ_d are their largest. As has been discussed in [49], the inflationary 4D boundary has significant gravitational back-reaction on the 5D geometry if $M_{\rm C} \lesssim H$, the case one would consider without a chemical potential. Fortunately, the chemical potential allows us to work in the $M_{\rm C} \gg H$ region, where the back-reaction can be negligible.

B.7.2 Realizing the chemical potential

In the setup described above it is more natural to write down the chemical potential model in the basis in eq. (B.32). In particular, the U(1) breaking $\chi_u\chi_d$-term can be implemented via an interaction localized on the boundary

$$S_{\cancel{\rm U(1)}} = -\int_{x_5=L} {\rm d}^4x \left(\alpha_{\rm 5D}\chi_u\chi_d e^{i\phi/\Lambda} + {\rm c.c.}\right), \tag{B.167}$$

where $\alpha_{\rm 5D}$ is a dimension one coupling. After plugging in the KK decomposition of χ_u and χ_d:

$$\chi_{u,d}(x,x_5) = \sqrt{\frac{2}{L}}\sum_{n=0}^{\infty}\chi_{u,d}^{(n)}(x)\sin\left[\left(n+\frac{1}{2}\right)\frac{\pi x_5}{L}\right], \tag{B.168}$$

we have

$$\mathcal{L}_{\cancel{\rm U(1)}} = -\frac{2\alpha_{\rm 5D}M_{\rm C}}{\pi}\chi_u^{(0)}\chi_d^{(0)}e^{i\phi/\Lambda} + {\rm c.c.} + \cdots . \tag{B.169}$$

When SUSY is taken into account, we consider the supermultiplet $X_{u,d}$ containing $\chi_{u,d}$ as scalar components. It turns out that a term like eq. (B.167) cannot be obtained without breaking SUSY. Combined with the requirement that SUSY breaking be spontaneous during inflation, the naive way to implement the coupling between $\chi_u\chi_d$ and F_S would be to include

the following term in the superpotential:

$$\begin{aligned}\mathcal{L}_{\cancel{\mathrm{U}(1)}} &= -\int \mathrm{d}^2\theta\, \frac{\kappa_0}{\Lambda_{5\mathrm{D}}} X_u X_d S \exp\left(\frac{\sqrt{2}\Phi}{\Lambda}\right) + \mathrm{c.c.} \\ &= -\frac{2\kappa_0}{\pi}\frac{M_{\mathrm{C}}}{\Lambda_{5\mathrm{D}}} F_S \chi_u^{(0)} \chi_d^{(0)} e^{i\phi/\Lambda} + \cdots + \mathrm{c.c.}\end{aligned} \tag{B.170}$$

This implementation is flawed as it breaks the $\mathrm{U}(1)_R$ symmetry of the SUSY GUT theory explicitly, under which θ has $+1$ charge and the charges of Higgs bosons are listed in table B.3. This places the model in danger as the $\mathrm{U}(1)_R$ symmetry has been utilized in [75–77] to eliminate dangerous dimension-5 operators that can mediate proton decay.

To fix this, one can consider the Giudice-Masiero mechanism [131], where the interaction now comes from the following term in the Kähler potential:

$$\begin{aligned}\mathcal{L}_{\cancel{\mathrm{U}(1)}} &= -\int \mathrm{d}^4\theta \left[\frac{\kappa_1}{\Lambda_{5\mathrm{D}}^3} X_u X_d \bar{S} S \exp\left(\frac{\sqrt{2}\Phi}{\Lambda}\right) + \mathrm{c.c.}\right] \\ &= -\frac{2\kappa_1}{\pi}\frac{M_{\mathrm{C}}}{\Lambda_{5\mathrm{D}}^3} |F_S|^2 \chi_u^{(0)} \chi_d^{(0)} e^{i\phi/\Lambda} + \cdots + \mathrm{c.c.}\end{aligned} \tag{B.171}$$

This new Lagrangian renders the $\mathrm{U}(1)_R$ breaking spontaneous by $\langle F_S \rangle$ and restores it at the end of the inflation. We now have the following expression for α:

$$\alpha = \frac{2\kappa_1}{\pi}\frac{M_{\mathrm{C}} |\langle F_S \rangle|^2}{\Lambda_{5\mathrm{D}}^3} \simeq \frac{6\kappa_1}{\pi}\frac{M_{\mathrm{C}} M_{\mathrm{pl}}^2}{\Lambda_{5\mathrm{D}}^3} H^2, \tag{B.172}$$

where $M_{\mathrm{pl}} \gtrsim 4 \times 10^4 H$.

There is another way to obtain the chemical potential coupling. We can couple $\chi_u \chi_d$ to

$-(\nabla_\mu\phi)^2 \supset \dot{\phi}_0^2 = (60H)^4$ through the following Kähler potential term:

$$\begin{aligned}\mathcal{L}_{\cancel{\mathrm{U}(1)}} &= -\int \mathrm{d}^4\theta \left[\frac{\kappa_2}{\Lambda_{5\mathrm{D}}^3} X_u X_d (\Phi+\bar{\Phi})^2 \exp\left(\frac{\sqrt{2}\Phi}{\Lambda}\right) + \mathrm{c.c.}\right] \\ &= \frac{2\kappa_2}{\pi}\frac{M_\mathrm{C}}{\Lambda_{5\mathrm{D}}^3}(\nabla_\mu\phi)^2 \chi_u^{(0)}\chi_d^{(0)} e^{i\phi/\Lambda} + \cdots + \mathrm{c.c.}\end{aligned} \tag{B.173}$$

After plugging in $\sigma = 0$ and $\phi = \dot{\phi}_0 t$, we have the following relation

$$\alpha = \frac{2\kappa_2}{\pi}\frac{M_C\dot{\phi}_0^2}{\Lambda_{5\mathrm{D}}^3} \simeq \frac{2\kappa_2}{\pi}\frac{M_\mathrm{C}}{\Lambda_{5\mathrm{D}}^3}(60H)^4. \tag{B.174}$$

Note that in this realization, there are extra couplings between $\chi_u\chi_d$ (or the effective χ_μ) and $\delta\phi$ in addition to those described by eq. (B.70):

$$\delta\mathcal{L}_{\delta\chi}^{(\mathrm{int})} = -\frac{\alpha}{\lambda^2}(-\eta)^{-i\lambda}\left[2(-\eta)\frac{\lambda}{\Lambda}\delta\phi' - \frac{1}{\Lambda^2}(\nabla\delta\phi)^2\right] e^{i\delta\phi/\Lambda}\delta\chi_\mu + \mathrm{c.c.} \tag{B.175}$$

This will make the parametric dependence on λ and μ somewhat different in the bispectrum.

If we set $\Lambda_{5\mathrm{D}} \simeq \Lambda$, $M_\mathrm{C} = 2\mu H$, the upper-bounds on α in both scenarios are higher than the two constraints from eqs. (B.145) and (B.150), sufficient to cover all the parameter region in figure B.21 with $N_\chi = 3$. Consequently, a non-analytic signal of order $f_\mathrm{NL} \gtrsim \mathcal{O}(0.01)$ can be used to probe these colored Higgs bosons with $M_u = M_d = M_\mathrm{C}/2$ lighter than $30H \lesssim 1.5\times10^{15}\,\mathrm{GeV}$.

In the generic model eq. (B.31), it is hard to estimate the natural size of the α parameter as it has positive mass dimensions, while the system has enhanced symmetry when it vanishes, just like the fermion mass. With supersymmetry, we have seen that the size α is controlled by the SUSY-breaking effect during the inflation. In this case, we can further measure the naturalness

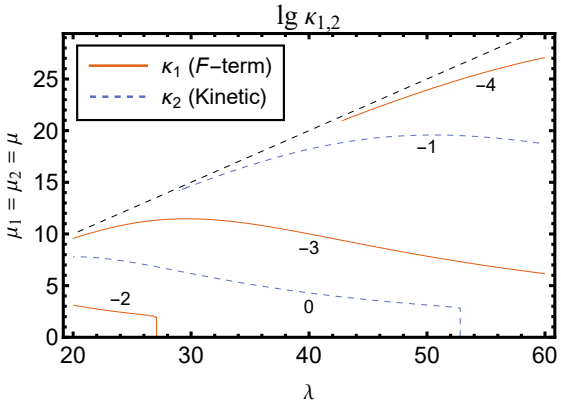


Figure B.23: The maximal sizes ($\log_{10}$) of the Wilson coefficients of the F-term-induced symmetry-breaking operator in eq. (B.171) and the kinetic-induced symmetry-breaking operator in eq. (B.173) under all the constraints on α discussed in section B.6.2, assuming $\Lambda_{5D} = \Lambda$ and maximal H from eq. (B.1). Compared with figure B.21, a typical choice of α with $F_{\min} > 10^{-2}$ while also compatible with all constraints requires $\kappa_1 \sim 10^{-3}$ or $\kappa_2 \sim 10^{-1}$. Note that with different choice of the ratio Λ_{5D}/Λ, $\kappa_{1,2}$ scale as $(\Lambda_{5D}/\Lambda)^3$.

of the parameter α by assuming that both terms contributing to α arises and see whether extremely small Wilson coefficients are required to satisfy the constraints in section B.6.2 and obtain an observable signal. Under the assumption $\Lambda_{5D} = \Lambda$ and maximal H from eq. (B.1), the result is summarized in figure B.23. Combined with figure B.21, we see that a typical parameter set that can satisfy all the constraints with an observable signal requires $\kappa_1 \sim 10^{-3}$ or $\kappa_2 \sim 10^{-1}$. While the required value for κ_1 may seem to be too small to be natural, this can be solved by introducing a small hierarchy of $\mathcal{O}(10)$ between Λ_{5D} and Λ, which will then predict $\kappa_1 \sim 1$ with contribution from κ_2 being suppressed.

B.8 Discussion and Conclusions

In this paper we have studied the chemical potential mechanism for a pair of charged scalars at the one-loop level. We have shown that the leading-order contribution to the bispectrum

can be computed using the de Sitter spectral representation, where one integrates a fictitious tree-level amplitude against the spectral density function. We have demonstrated that the non-analytic signal comes from the threshold limit of the integral with the invariant mass $\mu \simeq \mu_1+\mu_2$, while there is also a smooth background coming from the resonance limit when $\mu \simeq \lambda$.

In the limit of a large chemical potential, λ, and masses μ_1, μ_2, both the non-analytic signal and the analytic background have simple forms given by eq. (B.59) and eq. (B.118) respectively. We have that a model with only the lowest dimensional symmetry-breaking operator can satisfy all constraints, and generate an oscillatory signal in the bispectrum of order $f_{\rm NL} \gtrsim \mathcal{O}(0.01)$, which could be within the sensitivity of future 21 cm experiments. This mechanism may apply to the colored Higgs bosons in orbifold SUSY GUTs up to $M_u = M_d \lesssim 10^{15}\,\text{GeV}$.

In this work, we have been considering the exchange of two scalars as a spin-0 multi-particle state. It may be interesting to considering the generalization to nonzero spin exchange. One can consider two heavy particles with spin s_1, s_2 exchanged as a spin-s multi-particle state, where:

$$\mathbf{s} \in \mathbf{s}_1 \otimes \mathbf{s}_2 \otimes \mathbf{j} \tag{B.176}$$

for some $j = 0, 1, 2, \cdots$ (the orbital angular momentum). The amplitude is then given by integrating the tree-level amplitude of a spin-s particle times the corresponding spectral density.

Another direction for future work could be relaxing the parametric resonance constraints. The spectral decomposition approach requires treating the breaking of de Sitter isometry down to scale invariance perturbatively. This put very stringent constraints on the coupling α in the scalar model. However, parametric resonance might be desirable as it can induce a provide a

large enhancement of the non-analytic signal. Thus, it will be helpful to compute the bispectrum in this regime, and study its phenomenology.

Finally, while our model predicts a striking oscillatory signal, it is accompanied by a large smooth background. It would be interesting to see if some variant of the minimal model can improve the signal-to-background ratio.

Acknowledgments

AB is supported by DOE grant DE-SC-0013642 at the University of Chicago and the Fermi Forward Discovery Group, LLC, under Contract No. 89243024CSC000002 with the DOE Office of Science, Office of High Energy Physics. EB, RS, and ZX are supported by the NSF grant PHY-2210361 and the Maryland Center for Fundamental Physics.

B.9 Appendix

B.9.1 Integrals and Identities Involving Special Functions

In this appendix, we collect some integrals involving special functions used in the main text and subsequent sections. For $\operatorname{Re}\rho > 0$:

$$e^{-\pi\mu/2}\sqrt{\frac{\pi}{2}}\int_0^{\infty_+} x^{\rho-1}e^{ipx}H^{(1)}_{i\mu}(x)\,\mathrm{d}x = \frac{e^{(\rho-1)\pi i/2}}{2^{\rho-1/2}}\mathbf{F}^{\rho}_{i\mu}(p), \qquad p>-1, \quad \text{(B.177)}$$

$$e^{-\pi\mu/2}\sqrt{\frac{\pi}{2}}\int_0^{\infty_-} x^{\rho-1}e^{ipx}H^{(1)}_{i\mu}(x)\,\mathrm{d}x = \frac{e^{(\rho-1)\pi i/2}}{2^{\rho-1/2}}\mathbf{F}^{\rho}_{i\mu}(p+i\epsilon), \qquad p<-1, \quad \text{(B.178)}$$

$$e^{-\pi\mu/2}\sqrt{\frac{\pi}{2}}\int_0^{x} x^{\rho-1}e^{\pm ix}H^{(1)}_{i\mu}(x)\,\mathrm{d}x = \frac{e^{(\rho-1)\pi i/2}}{2^{\rho-1/2}}\big[\mathbf{G}^{\rho,\pm}_{i\mu}(x)+\mathbf{G}^{\rho,\pm}_{-i\mu}(x)\big]. \quad \text{(B.179)}$$

where the function on the RHS of the first two equations is given by

$$\mathbf{F}^{\rho}_{i\mu}(p) \coloneqq \Gamma\left(\begin{matrix}\rho+i\mu,\rho-i\mu\\ \frac{1}{2}+\rho\end{matrix}\right)F\left(\begin{matrix}\rho+i\mu,\rho-i\mu\\ \frac{1}{2}+\rho\end{matrix};\frac{1-p}{2}\right), \tag{B.180}$$

with Γ, F defined as in footnote 9. In the final equation

$$\begin{aligned}\mathbf{G}^{\rho,\pm}_{i\mu}(x) &\coloneqq \frac{1}{\rho+i\mu}\Gamma\left(\begin{matrix}-2i\mu\\ \frac{1}{2}-i\mu\end{matrix}\right)(-2ix)^{\rho+i\mu}{}_2F_2\left(\begin{matrix}\frac{1}{2}+i\mu,\rho+i\mu\\ 1+2i\mu,1+\rho+i\mu\end{matrix};\pm 2ix\right)\\ &= \frac{(-2ix)^{\rho+i\mu}}{4\pi\sinh\pi\mu}\int_L \frac{(\mp 2ix)^t}{\rho+i\mu+t}\Gamma\left(\begin{matrix}-t,\frac{1}{2}+i\mu+t\\ 1+2i\mu+t\end{matrix}\right)\mathrm{d}t,\end{aligned} \tag{B.181}$$

where $L : -i\infty \to i\infty$ is some path separating poles of $\Gamma(-t)$ from other poles of the integrand. The first two equations (B.177) and (B.178) can be derived from eq. (10.43.22) in [71] after Wick rotations (also see the appendices of ref. [35] and [132]). For the second equation, we can first use the following power expansion of $H^{(1)}_{i\mu}(x)$:

$$e^{-\pi\mu/2\pm ix}H^{(1)}_{i\mu}(x) = -\frac{2i}{\sqrt{\pi}}\left[\Gamma\left(\begin{matrix}-2i\mu\\ \frac{1}{2}-i\mu\end{matrix}\right)(-2ix)^{i\mu}\sum_{n=0}^{\infty}\frac{\left(\frac{1}{2}+i\mu\right)_n}{(1+2i\mu)_n}\frac{(\pm 2ix)^n}{n!}+(\mu\to-\mu)\right] \tag{B.182}$$

and then integrate term-by-term.

The $\mathbf{F}^{\rho}_{i\mu}(p)$ in eq. (B.180) also satisfies the following recurrence relation [71, eq. (15.5.19)]:

$$\frac{1}{4}\left(p^2-1\right)\mathbf{F}^{2+\rho}_{i\mu}(p) - \left(\frac{1}{2}+\rho\right)p\mathbf{F}^{1+\rho}_{i\mu}(p) + \left(\rho^2+\mu^2\right)\mathbf{F}^{\rho}_{i\mu}(p) = 0. \tag{B.183}$$

This has been used to prove the equivalence of the two forms in eq. (B.85), where the latter form is equivalent to the one in ref. [35].

Finally, to extract the non-analytic piece of the integral over the spectral density, we made

use of the identity [71, eq. (5.13.1)]

$$\int_L \Gamma(a+s)\Gamma(b-s)z^{-s}\,\mathrm{d}s = \Gamma(a+b)\cdot\frac{z^a}{(1+z)^{a+b}}, \tag{B.184}$$

where $L : -i\infty \to i\infty$ is some path separating the poles of $\Gamma(a+s)$ from the poles of $\Gamma(b-s)$.

B.9.2 Asymptotic Expansions of $\widetilde{\mathbf{F}}^s_{i\mu}(p)$

In eq. (B.120) we split the integrand into two pieces to facilitate use of the residue theorem. In this appendix, we collect identities needed to justify that step. Firstly, the $\widetilde{\mathbf{F}}^s_{i\mu}(p)$ defined in eq. (B.121) has the following integral representation [71, eq. (15.6.1)]:

$$\widetilde{\mathbf{F}}^s_{i\mu}(p) = \frac{2^{-2i\mu}}{\sqrt{\pi}}\Gamma\begin{pmatrix}1-i\mu\\ \frac{1}{2}-i\mu\end{pmatrix}\int_0^1 [t(1-t)]^{-1/2-i\mu}\left(\frac{p-1}{2}+t\right)^{-s}\mathrm{d}t\,. \tag{B.185}$$

In this appendix, we will use this to derive the asymptotic approximation of $\widetilde{\mathbf{F}}^s_{i\mu}$ in three limits assuming $p > 1$.

The first limit is when $s = \rho - i\mu$ and $\mu \to \infty$ with $\operatorname{Im}\mu \geq 0$:

$$2^{2i\mu}\widetilde{\mathbf{F}}^{\rho-i\mu}_{i\mu}(p) = \frac{1}{\sqrt{\pi}}\Gamma\begin{pmatrix}1-i\mu\\ \frac{1}{2}-i\mu\end{pmatrix}\int_0^1 [t(1-t)]^{-1/2}\left(\frac{p-1}{2}+t\right)^{-\rho}e^{-i\mu g(t;p)}\,\mathrm{d}t, \tag{B.186}$$

where

$$g(t;p) = \log t(1-t) - \log\left(\frac{p-1}{2}+t\right). \tag{B.187}$$

For $p > 1$, $g'(t;p) = 0$ has a unique solution between $(0,1)$ at

$$t_* = \frac{1}{2}\Big(1 - p + \sqrt{p^2-1}\Big), \qquad g''(t_*;p) = -\frac{4}{\sqrt{p^2-1}\Big(p - \sqrt{p^2-1}\Big)} < 0. \tag{B.188}$$

Therefore, as $\mu \to \infty$, the integral above can be approximated as

$$\begin{aligned} 2^{2i\mu}\widetilde{\mathbf{F}}^{\rho-i\mu}_{i\mu}(p) &\simeq \frac{e^{\pi i/4}}{\sqrt{\pi}}\Gamma\begin{pmatrix} 1-i\mu \\ \frac{1}{2}-i\mu \end{pmatrix}\sqrt{\frac{2\pi}{-\mu g''(t_*;p)}}[t_*(1-t_*)]^{-1/2-i\mu}\left(\frac{p-1}{2}+t_*\right)^{-\rho+i\mu} \\ &\simeq \frac{1}{\sqrt{2}}\left(\frac{\sqrt{p^2-1}}{2}\right)^{-\rho}\Big(p-\sqrt{p^2-1}\Big)^{-i\mu}. \end{aligned} \tag{B.189}$$

In the main text we use the asymptotic expansion of (B.189) to demonstrate that we can close the contour in the upper half plane.

There are other two limits that will be used to calculate B_{++} in appendix B.9.4. The second limit is when $s \to \infty$, with $\operatorname{Re} s > 0$. In this limit, the integration is dominated by the $t = 0$ boundary:

$$\begin{aligned} 2^{2i\mu}\widetilde{\mathbf{F}}^{s}_{i\mu}(p) &\simeq \frac{1}{\sqrt{\pi}}\Gamma\begin{pmatrix} 1-i\mu \\ \frac{1}{2}-i\mu \end{pmatrix}\int_0^\infty t^{-1/2-i\mu}\left(\frac{p-1}{2}+t\right)^{-s}\mathrm{d}t \\ &= \frac{1}{\sqrt{\pi}}\Gamma\begin{pmatrix} 1-i\mu, -\frac{1}{2}+i\mu+s \\ s \end{pmatrix}\left(\frac{p-1}{2}\right)^{1/2-i\mu-s} \\ &\simeq \frac{\Gamma(1-i\mu)}{\sqrt{\pi}}s^{-1/2+i\mu}\left(\frac{p-1}{2}\right)^{1/2-i\mu-s}. \end{aligned} \tag{B.190}$$

The third limit is when $|\mu| \to \infty$ with $\operatorname{Im}\mu \geq 0$:

$$\widetilde{\mathbf{F}}^{s}_{i\mu}(p) = \frac{2^{-2i\mu}}{\sqrt{\pi}}\Gamma\begin{pmatrix} 1-i\mu \\ \frac{1}{2}-i\mu \end{pmatrix}\int_0^1 [t(1-t)]^{-1/2}\left(\frac{p-1}{2}+t\right)^{-s}e^{-i\mu h(t)}\,\mathrm{d}t, \tag{B.191}$$

where $h(t) \coloneqq \log t(1-t)$. In this case:

$$\widetilde{\mathbf{F}}^{s}_{i\mu}(p) \simeq \frac{2^{-2i\mu}e^{\pi i/4}}{\sqrt{\pi}}\Gamma\left(\begin{matrix}1-i\mu\\ \frac{1}{2}-i\mu\end{matrix}\right)\sqrt{\frac{2\pi}{-\mu h''(t)}}[t(1-t)]^{-1/2-i\mu}\left(\frac{p-1}{2}+t\right)^{-s}\Bigg|_{t=1/2}$$
$$\simeq \left(\frac{p}{2}\right)^{-s}. \tag{B.192}$$

B.9.3 Spectral Density in Non-relativistic Limit of de Sitter Spacetime

In this section, we will derive eq. (B.46) directly using the non-relativistic limit of de Sitter spacetime without referring to eq. (B.42). To begin with, we first show that the non-relativistic limit of de Sitter spacetime is described by the quantum inverted harmonic oscillator with Hamiltonian

$$\mathscr{H} = \frac{\mathbf{P}^2}{2M} - \frac{1}{2}M\mathbf{X}^2. \tag{B.193}$$

The equivalence can either be established at the level of isometry groups or metrics. For a derivation using group contractions readers may refer to ref. [133]. We will take the second approach. We start from the expression for the de Sitter metric in the static coordinates with the speed-of-light c restored[13]

$$\mathrm{d}s^2 = -\left(1-\frac{R^2}{c^2}\right)c^2\,\mathrm{d}T^2 + \left(1-\frac{R^2}{c^2}\right)^{-1}\mathrm{d}R^2 + R^2\,\mathrm{d}\Omega^2\,. \tag{B.194}$$

[13]The coordinate transformation from the static coordinate to the conformal coordinates eq. (B.6) we have been using so far is given by

$$X^i = -\frac{x^i}{\eta}, \qquad e^{-T} = \sqrt{\eta^2 - \frac{x^2}{c^2}}.$$

The nontrivial Christoffel symbols are as follows:

$$\Gamma^R_{RR} = -\Gamma^T_{RT} = \frac{R}{c^2}\left(1 - \frac{R^2}{c^2}\right)^{-1}, \qquad \Gamma^R_{TT} = -R\left(1 - \frac{R^2}{c^2}\right). \tag{B.195}$$

In the limit $c \to \infty$, only Γ^R_{TT} survives and the geodesic equation becomes

$$\frac{\mathrm{d}^2 X^i}{\mathrm{d}T^2} + \Gamma^R_{TT}\hat{R}^i = \frac{\mathrm{d}^2 X^i}{\mathrm{d}T^2} - X^i = 0, \tag{B.196}$$

which is exactly the Newtonian equation of motion from eq. (B.193).

In the non-relativistic limit, the scalar propagator can be approximated as

$$G_{-+}(X_1, X_2) \simeq \frac{1}{2M} G_{\mathrm{NR}}(\mathbf{X}_1, \mathbf{X}_2, T_1 - T_2) e^{-iMc^2(T_1 - T_2)}, \tag{B.197}$$

where

$$G_{\mathrm{NR}}(\mathbf{X}_1, \mathbf{X}_2, T) \coloneqq \left(\frac{M}{2\pi i \sinh T}\right)^{3/2} \exp\left\{\frac{iM}{2\sinh T}\left[\left(\mathbf{X}_1^2 + \mathbf{X}_2^2\right)\cosh T - 2\mathbf{X}_1 \cdot \mathbf{X}_2\right]\right\} \tag{B.198}$$

is the propagator of the inverted harmonic oscillator eq. (B.193)[14]. Now, we rewrite the spectral decomposition eq. (B.40) using $M \simeq \mu$ as

$$G_{-+}(X_1, X_2; M_1) G_{-+}(X_1, X_2; M_2) \simeq \int_0^\infty \rho^{\mathrm{dS}}_{M_1 M_2}(M) G_{-+}(X_1, X_2; M)\, \mathrm{d}M\,. \tag{B.199}$$

[14]The readers may be more familiar with the propagator of the quantum harmonic oscillator given by

$$G(\mathbf{X}_1, \mathbf{X}_2, T) \coloneqq \left(\frac{M\omega}{2\pi i \sin \omega T}\right)^{3/2} \exp\left\{\frac{iM\omega}{2\sin \omega T}\left[\left(\mathbf{X}_1^2 + \mathbf{X}_2^2\right)\cos \omega T - 2\mathbf{X}_1 \cdot \mathbf{X}_2\right]\right\},$$

following from path integral. The propagator of the inverted harmonic oscillator can be obtained in the same manner, or by analytic continuation $\omega \to iH$, where we have then set $H = 1$.

Given that $G_{-+}(X_1, X_2; M)$ contains the highly oscillatory $e^{-iMc^2(T_1-T_2)}$ term and $\rho^{\rm dS}_{M_1M_2}(M)$ changes rapidly around $M = M_{12} \coloneqq M_1 + M_2$, we can approximate the integration as

$$\text{RHS} \simeq G_{-+}(X_1, X_2; M_{12}) \int_{-M_{12}}^{\infty} \rho^{\rm dS}_{M_1M_2}(M_{12}+\varepsilon) e^{-i\varepsilon c^2(T_1-T_2)}\, \mathrm{d}\varepsilon, \tag{B.200}$$

i.e.

$$\begin{aligned}
\int_{-M_{12}}^{\infty} \rho^{\rm dS}_{M_1M_2}(M_{12}+\varepsilon) e^{-i\varepsilon c^2 T}\, \mathrm{d}\varepsilon &\simeq \left.\frac{G_{-+}(X_1, X_2; M_1) G_{-+}(X_1, X_2; M_2)}{G_{-+}(X_1, X_2; M_{12})}\right|_{T_1-T_2=T} \\
&= \frac{(2M_{12})}{(2M_1)(2M_2)} \frac{G^{\rm NR}_{-+}(\mathbf{X}_1, \mathbf{X}_2, T; M_1) G^{\rm NR}_{-+}(\mathbf{X}_1, \mathbf{X}_2, T; M_2)}{G^{\rm NR}_{-+}(\mathbf{X}_1, \mathbf{X}_2, T; M_{12})} \\
&= \frac{e^{-3\pi i/4}}{2^{5/2}\pi^{3/2}} \sqrt{\frac{M_1M_2}{M_{12}}} (\sinh T)^{-3/2}.
\end{aligned} \tag{B.201}$$

The spectral density near $M = M_{12}$ can thus be recovered using the inverse Fourier transformation ($c = 1$):

$$\begin{aligned}
\rho^{\rm dS}_{M_1M_2}(M_{12}+\varepsilon) &\simeq \frac{1}{2^{5/2}\pi^{3/2}} \sqrt{\frac{M_1M_2}{M_{12}}} \left[e^{-3\pi i/4} \int_0^{\infty} \frac{\mathrm{d}T}{2\pi} (\sinh T)^{-3/2} e^{i\varepsilon T} \right. \\
&\qquad \left. + e^{3\pi i/4} \int_{-\infty}^{0} \frac{\mathrm{d}T}{2\pi} (-\sinh T)^{-3/2} e^{i\varepsilon T} \right] \\
&= \frac{1}{4\pi^{5/2}} \sqrt{\frac{M_1M_2}{M_{12}}} \left[e^{-3\pi i/4} \int_0^1 \frac{t^{1/2-i\varepsilon}\, \mathrm{d}t}{(1-t^2)^{3/2}} + \text{c.c.} \right] \\
&= \frac{1}{4\pi^3} \sqrt{\frac{M_1M_2}{M_{12}}} \Gamma\left(\frac{3}{4} + \frac{i\varepsilon}{2}\right) \Gamma\left(\frac{3}{4} - \frac{i\varepsilon}{2}\right) e^{\pi\varepsilon/2},
\end{aligned} \tag{B.202}$$

where the divergence at $T = 0$ cancels between the $T > 0$ and $T < 0$ integral. This agrees with eq. (B.46).

The quantum inverted harmonic oscillator naturally explains why $\rho^{\rm dS}_{\mu_1\mu_2}$ does not drop to

zero, but is only exponentially suppressed, when $M < M_1 + M_2$. This is because a particle can decay into two particles with a higher total mass by extracting energy from the expanding universe and going through a quantum tunneling process.

B.9.4 Exact Expression of $B_{++,\mathrm{A}}$

In this subsection, we provide an analytic evaluation of the integral eq. (B.90) and derive some of its properties. We first complete the x_1-integral using eq. (B.179):

$$\begin{aligned}\int_0^{x_2} \mathrm{d}x_1\, \mathcal{I}^{(\pm)}_{1,++}(\lambda;x_1) &= \frac{e^{-\pi i/4}e^{\mp\pi\lambda/2}}{2^{\pm i\lambda}}\Big[\mathbf{G}^{\frac{1}{2}\pm i\lambda,-}_{i\mu}(x_2)+\mathbf{G}^{\frac{1}{2}\pm i\lambda,-}_{-i\mu}(x_2)\Big],\\ \int_0^{x_2} \mathrm{d}x_1\, \mathcal{I}^{(\mp)*}_{1,-+}(\lambda;x_1) &= \frac{e^{\pi i/4}e^{\pm\pi\lambda/2}}{2^{\pm i\lambda}}\Big[\mathbf{G}^{\frac{1}{2}\mp i\lambda,+*}_{i\mu}(x_2)+\mathbf{G}^{\frac{1}{2}\mp i\lambda,+*}_{-i\mu}(x_2)\Big],\end{aligned} \tag{B.203}$$

where $\mathbf{G}^{\rho,\pm}_{i\mu}(x)$ is defined in eq. (B.181) with the contour L shown by the contour L_1 in figure B.24. Using the integral representation from eq. (B.181), the x_2-integration can be evaluated first using eqs. (B.177) and (B.178), leaving

$$B^{(\pm)}_{++,\mathrm{A}}(k_1,k_2,k_3) = \frac{i}{8k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{\mu\Big[(\lambda^2-M^2)^2+9\lambda^2\Big]}\sum_{n=0}^{2}\frac{c_n^{\pm}}{2^n}\big[\mathbf{H}^{\pm i\lambda}_{n,i\mu}(p)-\mathbf{H}^{\pm i\lambda}_{n,-i\mu}(p)\big], \tag{B.204}$$

where we have used eq. (B.120) and the function $\widetilde{\mathbf{F}}^{s}_{i\mu}(p)$ defined in eq. (B.121) to write

$$\mathbf{H}^{i\lambda}_{n,i\mu}(p) \coloneqq \frac{1}{2\pi i}\Gamma\begin{pmatrix}1+2i\mu\\ \frac{1}{2}+i\mu\end{pmatrix}\int_L \frac{\widetilde{\mathbf{F}}^{1+n+t}_{i\mu}(p)\,\mathrm{d}t}{\frac{1}{2}+i\lambda+i\mu+t}\Gamma\begin{pmatrix}-t, 1+n+t, \frac{1}{2}+i\mu+t\\ 1+2i\mu+t\end{pmatrix}. \tag{B.205}$$

Although the t-integral in eq. (B.205) cannot be evaluated analytically, it can be represented as an infinite sum over the residues of $\Gamma(-t)$ if the path L can be deformed to start and end at $+\infty$, as shown by the contour L_2 in figure B.24. To see when such a deformation is

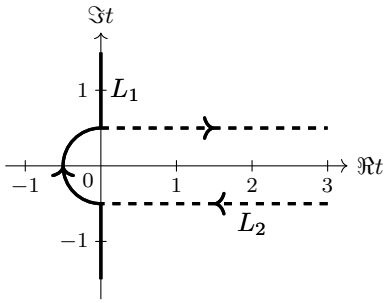


Figure B.24: Two contours used to perform the integration in eq. (B.205).

allowed, we need the following asymptotic bound on the integrand as $t \to \infty$ with $\operatorname{Re} t \geq 0$:[15]

$$\begin{aligned}\frac{\widetilde{\mathbf{F}}^{1+n+t}_{i\mu}(p)}{\frac{1}{2}+i\lambda+i\mu+t}\Gamma\left(\begin{matrix}-t, 1+n+t, \frac{1}{2}+i\mu+t\\ 1+2i\mu+t\end{matrix}\right) &\sim t^{-2+n}\left(\frac{p-1}{2}\right)^{-t}\csc\pi t \\ &\lesssim \begin{cases} t^{-2+n}\left(\dfrac{p-1}{2}\right)^{-t}, & t\to\pm i\tau+\infty, \\ t^{-2+n}e^{-\pi|t|}, & t\to\pm i\infty. \end{cases}\end{aligned} \tag{B.206}$$

From this it follows that, for $p > 3$, the integral converges exponentially for L_2 in figure B.24 and can be expressed as a sum over the residues of $\Gamma(-t)$:

$$\mathbf{H}^{i\lambda}_{n,i\mu}(p) = \sum_{m=0}^{\infty} \frac{(-)^m (m+1)_n \left(\frac{1}{2}+i\mu\right)_m}{\left(\frac{1}{2}+m+i\lambda+i\mu\right)(1+2i\mu)_m}\widetilde{\mathbf{F}}^{1+n+m}_{i\mu}(p), \qquad (p>3). \tag{B.207}$$

[15] For the gamma functions, the asymptotic property $\Gamma(a+t) \sim t^a\Gamma(t)$ and the identity $\Gamma(-t)\Gamma(1+t) = -\pi\csc\pi t$ are used. For $\widetilde{\mathbf{F}}^{1+n+t}_{i\mu}(p)$, the asymptotic expansion is given by eq. (B.190) in appendix B.9.2.

In particular, this shows that $B^{(\pm)}_{++,\mathrm{A}}$ are analytic functions at $p \to \infty$, and in the squeezed limit:

$$B^{(\pm)}_{++,\mathrm{A}}(k_1,k_2,k_3) \xrightarrow{p\to\infty} \frac{3}{8k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda(1\pm i\lambda)}{\left[\left(\frac{1}{2}\pm i\lambda\right)^2+\mu^2\right]\left[\left(\lambda^2-\mu^2-\frac{9}{4}\right)^2+9\lambda^2\right]}p, \tag{B.208}$$

$$\begin{aligned} B_{++,\mathrm{A}}(k_1,k_2,k_3) &:= B^{(+)}_{++,\mathrm{A}}(k_1,k_2,k_3)+B^{(-)}_{++,\mathrm{A}}(k_1,k_2,k_3) \\ &\xrightarrow{p\to\infty} \frac{3}{4k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda\left(\mu^2+\frac{1}{4}\right)}{\left[\left(\lambda^2-\mu^2-\frac{1}{4}\right)^2+\lambda^2\right]\left[\left(\lambda^2-\mu^2-\frac{9}{4}\right)^2+9\lambda^2\right]}p. \end{aligned} \tag{B.209}$$

On the other hand, eq. (B.206) also shows that the integral always converges exponentially for L_1 in figure B.24. Therefore, $\mathbf{H}^{i\lambda}_{n,i\mu}(p)$ can always be evaluated numerically along L_1. In practice, the contour should be far from all poles of the integrand for numerical stability. We choose the radius of the semicircle in figure B.24 to be $1/4$ in our numerical evaluation.

Returning to eq. (B.204), in the limit when both $|\lambda\pm\mu|, \mu \gg 1$, we can approximate $\mathbf{H}^{i\lambda}_{n,i\mu}(p)$ using eq. (B.192) as

$$\begin{aligned} \mathbf{H}^{i\lambda}_{n,i\mu}(p) &\simeq \frac{1}{\frac{1}{2}+i\lambda+i\mu}\cdot\frac{1}{2\pi i}\int_L\left(1-\frac{t}{\frac{1}{2}+i\lambda+i\mu}\right)\Gamma(-t)\Gamma(1+n+t)2^{-t}\left(\frac{p}{2}\right)^{-1-n-t}\mathrm{d}t \\ &= \frac{1}{\frac{1}{2}+i\lambda+i\mu}\left(\frac{p}{2}\right)^{-1-n}\sum_{k=0}^{\infty}\left(1-\frac{k}{\frac{1}{2}+i\lambda+i\mu}\right)\frac{(n+k)!}{k!}(-p)^{-k} \\ &= \frac{1}{\frac{1}{2}+i\lambda+i\mu}\left(\frac{p+1}{2}\right)^{-1-n}\left[n!+\frac{1}{\frac{1}{2}+i\lambda+i\mu}\frac{(n+1)!}{p+1}\right]. \end{aligned} \tag{B.210}$$

Therefore:

$$
\begin{aligned}
B^{(\pm)}_{++,\mathrm{A}}(k_1,k_2,k_3) &\simeq \frac{1}{2k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{(\lambda^2-\mu^2)^2\left[\left(\frac{1}{2}\pm i\lambda\right)^2+\mu^2\right]} \\
&\quad\times\sum_{n=0}^{2}\left[n!\mp\frac{2i\lambda}{\lambda^2-\mu^2}\frac{(n+1)!}{p+1}\right]c_n^{\pm}(p+1)^{-1-n} \\
&\simeq \frac{1}{2k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{(\lambda^2-\mu^2)^2\left[\left(\frac{1}{2}\pm i\lambda\right)^2+\mu^2\right]}\left[\frac{(1\pm i\lambda)p^2}{p+1}\right. \\
&\quad\left.-\frac{(1\pm i\lambda)(p+2)Q}{(p+1)^2}-\frac{1}{2}+\frac{2\lambda^2}{\lambda^2-\mu^2}\frac{1}{(p+1)^2}\left(p^2-\frac{2Q}{p+1}\right)\right],
\end{aligned}
\tag{B.211}
$$

$$
\begin{aligned}
B_{++,\mathrm{A}}(k_1,k_2,k_3) &\coloneqq B^{(+)}_{++,\mathrm{A}}(k_1,k_2,k_3)+B^{(-)}_{++,\mathrm{A}}(k_1,k_2,k_3) \\
&\simeq \frac{1}{k_1^3k_2^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{(\lambda^2-\mu^2)^4}\left\{\frac{\mu^2}{p+1}\left[p^2-\frac{(p+2)Q}{p+1}\right]+\frac{1}{2}(\lambda^2-\mu^2)\right. \\
&\quad\left.-\frac{2\lambda^2}{(p+1)^2}\left(p^2-\frac{2Q}{p+1}\right)\right\}.
\end{aligned}
\tag{B.212}
$$

This does not agree with eq. (B.95) in its naïve form. However, after summing over permutations of external momenta and using $\mu \coloneqq \sqrt{M^2-9/4} \simeq M$, both expressions are equal to

$$
\begin{aligned}
&B_{++,\mathrm{A}}(k_1,k_2,k_3)+(k_3\to k_1,k_2) \\
&\simeq \frac{1}{k_1^3k_2^3k_3^3}\frac{|\alpha|^2}{\Lambda^3}\frac{\lambda}{(\lambda^2-\mu^2)^3}\left[k_1k_2k_3-\frac{k_1^3+k_2^3+k_3^3}{2}\right. \\
&\quad\left.+\frac{2(k_1^2k_2^2+k_2^2k_3^2+k_1^2k_3^2)}{k_{123}}-\frac{k_1k_2k_3(k_1^2+k_2^2+k_3^2)}{k_{123}^2}-\frac{12\lambda^2}{\lambda^2-\mu^2}\frac{k_1^2k_2^2k_3^2}{k_{123}^3}\right].
\end{aligned}
\tag{B.213}
$$

This confirms eq. (B.95) as the leading-order approximation of eq. (B.204).

Finally, even for $|\lambda \pm \mu| \sim \mathcal{O}(1)$ with $\mu \gg 1$, eqs. (B.192) and (B.205) still imply that

$$\begin{aligned}\mathbf{H}^{i\lambda}_{n,i\mu}(p) &\simeq \frac{1}{2\pi i}\int_L \frac{\Gamma(-t)\Gamma(1+n+t)}{\frac{1}{2}+i\lambda+i\mu+t} 2^{-t}\left(\frac{p}{2}\right)^{-1-n-t} \mathrm{d}t \\ &= \frac{n!}{\frac{1}{2}+i\lambda+i\mu}\left(\frac{p}{2}\right)^{-1-n} F\left(\begin{matrix}1+n, \frac{1}{2}+i\lambda+i\mu \\ \frac{3}{2}+i\lambda+i\mu\end{matrix}; -\frac{1}{p}\right),\end{aligned} \tag{B.214}$$

this allows the corresponding one-loop amplitude being calculated approximately.

B.9.5 Full Evaluation of the One-loop Model

In this section, we present the full evaluation of the one-loop bispectrum. Section B.5.3 has shown that all contributions to the bispectrum can be written as integrations of the following form:

$$F^{(\pm)}_{X,\text{1-loop}}(k_1,k_2,k_3) = \int_{-\infty}^{\infty} \rho^{\text{dS}}_{\mu_1\mu_2}(\mu)\widetilde{F}^{(\pm)}_X(k_1,k_2,k_3;\mu)\,\mathrm{d}\mu \ + \text{c.c.}, \tag{B.215}$$

where $F_X \in \{F_{-+}, F_{++,\mathrm{NA}}, F_{++,\mathrm{A}}\}$ and

$$\begin{aligned}\widetilde{F}^{(\pm)}_{-+}(k_1,k_2,k_3;\mu) &\coloneqq \frac{\mathcal{F}}{\sqrt{\pi}}\frac{2^{2i\mu}e^{\mp\pi\lambda}}{\left(\frac{3}{2}\mp i\lambda\right)^2+\mu^2}\sum_{n=0}^{2}\frac{d_n}{2^n}\widetilde{\mathbf{F}}^{-\frac{3}{2}+n\mp i\lambda-i\mu}_{i\mu}(p)\\ &\quad\times\Gamma\begin{pmatrix} i\mu, -\frac{3}{2}+n\mp i\lambda-i\mu, \frac{1}{2}\pm i\lambda+i\mu, \frac{1}{2}\pm i\lambda-i\mu \\ 1\pm i\lambda\end{pmatrix}, \qquad \text{(B.216)}\\ \widetilde{F}^{(\pm)}_{++,\mathrm{NA}}(k_1,k_2,k_3;\mu) &\coloneqq ie^{\pi(\pm\lambda-\mu)}\widetilde{F}^{(\pm)}_{-+}(k_1,k_2,k_3;\mu), \qquad \text{(B.217)}\\ \widetilde{F}^{(\pm)}_{++,\mathrm{A}}(k_1,k_2,k_3;\mu) &\coloneqq -\frac{i\mathcal{F}}{2}\frac{1}{\mu\left[\left(\lambda^2-\mu^2-\frac{9}{4}\right)^2+9\lambda^2\right]}\sum_{n=0}^{2}\frac{c_n^{\pm}}{2^n}\mathbf{H}^{\pm i\lambda}_{n,i\mu}(p)\\ &= -\frac{i\mathcal{F}}{4}\frac{1}{\mu\left[\left(\lambda^2-\mu^2-\frac{9}{4}\right)^2+9\lambda^2\right]}\sum_{n=0}^{2}\frac{c_n^{\pm}}{2^n}\\ &\quad\times\sum_{m=0}^{\infty}\frac{(-)^m(m+1)_n\left(\frac{1}{2}+i\mu\right)_m}{\left(\frac{1}{2}+m\pm i\lambda+i\mu\right)(1+2i\mu)_m}\widetilde{\mathbf{F}}^{1+n+m}_{i\mu}(p). \qquad (p>3)\\ &\qquad\qquad \text{(B.218)}\end{aligned}$$

Here we have extracted the common factor:

$$\mathcal{F} \coloneqq \frac{5\lambda^2}{6}\left|\frac{\alpha}{\Lambda}\right|^2\frac{4}{p^3+3\chi^2p+4} \qquad \text{(B.219)}$$

and introduce the following notations:

$$\begin{cases} d_0 \coloneqq 1, \\ d_1 \coloneqq p, \\ d_2 \coloneqq Q, \end{cases} \qquad \begin{cases} c_0^{\pm} \coloneqq \frac{1}{2}\left[(1\pm 2i\lambda)\left(p^2-2Q\right)-1\right], \\ c_1^{\pm} \coloneqq \frac{1}{2}p\left[p^2-2(1\mp i\lambda)Q-1\right], \\ c_2^{\pm} \coloneqq \frac{1}{2}Q\left(p^2-1\right), \end{cases} \qquad \text{(B.220)}$$

where $Q \coloneqq \frac{1}{4}(p^2-\chi^2)$. All of these integrations can be expressed as sums over residues at poles with $\operatorname{Im}\mu > 0$. Such poles are summarized in table B.2 and table B.4. Consequently,

$$F_{\text{1-loop}} = \sum_{\mu_*, s_1=\pm} \left[\widetilde{F}^{(s_1,\mu_*)}_{-+,\text{1-loop}} + \widetilde{F}^{(s_1,\mu_*)}_{++,\text{NA,1-loop}} + \widetilde{F}^{(s_1,\mu_*)}_{++,\text{A,1-loop}} \right] + (k_1 \to k_2, k_3) + \text{c.c.}, \tag{B.221}$$

where $\mu_* = \pm\lambda, \pm\mu_{12}, \pm\Delta\mu$ are the real parts of the poles and $\widetilde{F}^{(s_1,\mu_*)}_{X,\text{1-loop}}$ are the some of residues at these poles.

$\widetilde{F}$	$\operatorname{Re}\mu = \lambda$		$\operatorname{Re}\mu = -\lambda$	
	source	μ	source	μ
$\widetilde{F}^{(-)}_{-+}$	$\Gamma\left(\frac{1}{2}-i\lambda+i\mu\right)$, $\Gamma\left(-\frac{3}{2}+n+i\lambda-i\mu\right)$	$\lambda+\left(\frac{1}{2}+k\right)i$	$\frac{1}{(3/2+i\lambda)^2+\mu^2}$	$-\lambda+\frac{3}{2}i$
$\widetilde{F}^{(+)}_{++,\text{NA}}$	$\frac{1}{(3/2-i\lambda)^2+\mu^2}$	$\lambda+\frac{3}{2}i$	$\Gamma\left(\frac{1}{2}+i\lambda+i\mu\right)$, $\Gamma\left(-\frac{3}{2}+n-i\lambda-i\mu\right)$	$-\lambda+\left(\frac{1}{2}+k\right)i$
$\widetilde{F}^{(+)}_{++,\text{A}}$	$\frac{1}{(\lambda^2-\mu^2-9/4)^2+9\lambda^2}$	$\lambda+\frac{3}{2}i$	$\frac{1}{\frac{1}{2}+m+i\lambda+i\mu}$, $\frac{1}{(\lambda^2-\mu^2-9/4)^2+9\lambda^2}$	$-\lambda+\left(\frac{1}{2}+k\right)i$

Table B.4: The position of poles of the integrand in eqs. (B.122) and (B.130) over the upper complex-μ plane and their sources ($n = 0, 1, 2$ and $m,\ k \in \mathbb{N}$). The positions of poles at $\operatorname{Re}\mu = \pm\mu_{12},\ \Delta\mu$ are from the spectral density and are the same for all amplitudes, given in table B.2. For other amplitudes, e.g. $\widetilde{F}^{(-)}_{++,\text{A}}$, the position of poles can be obtained using eq. (B.73).

Assuming $\mu_1 \geq \mu_2$ w.l.o.g., we have the estimation of orders for all 36 terms in table B.5 with the Stirling approximation. For $\lambda, \lambda-\mu_{12}, \mu_1, \mu_2$ all much larger than 1, only the following terms are not exponentially suppressed ($s_1, s_2 = \pm$):

$$\widetilde{F}^{(-,\mu_{12})}_{-+,\text{1-loop}},\quad \widetilde{F}^{(-,\lambda)}_{-+,\text{1-loop}},\quad \widetilde{F}^{(+,-\lambda)}_{++,\text{NA,1-loop}},\quad \widetilde{F}^{(s_1,s_2\lambda)}_{++,\text{A,1-loop}},\quad \widetilde{F}^{(s_1,s_2\mu_{12})}_{++,\text{A,1-loop}}. \tag{B.222}$$

$\widetilde{F}$	s_1	$\mathrm{Re}\,\mu$	$\lambda > \mu_{12}$	$\Delta\mu < \lambda < \mu_{12}$	$\lambda < \Delta\mu$
$-+$	$-$	λ	$\mathbf{0}$	$-\pi(\mu_{12}-\lambda)$	$-2\pi(\mu_1-\lambda)$
		μ_{12}	$\mathbf{0}$	$-2\pi(\mu_{12}-\lambda)$	
		$\Delta\mu$	$-2\pi\mu_2$		$-2\pi(\mu_1-\lambda)$
		$-\Delta\mu$	$-\pi\mu_{12}$		$-\pi(2\mu_1-\lambda)$
		$-\mu_{12}$	$-\pi\mu_{12}$	$-\pi(2\mu_{12}-\lambda)$	
		$-\lambda$	$-\pi\lambda$	$-\pi\mu_{12}$	$-\pi(2\mu_1-\lambda)$
	$+$	any μ_*	$-2\pi\lambda$ + (the one with $-\mu_*$ above)		
++, NA	$+$	$-\lambda$	$\mathbf{0}$	$-\pi(\mu_{12}-\lambda)$	$-2\pi(\mu_1-\lambda)$
		$-\mu_{12}$	$-\pi(\lambda-\mu_{12})$	$-\pi(\mu_{12}-\lambda)$	
		$-\Delta\mu$	$-\pi(\lambda+3\mu_2-\mu_1)$		$-\pi(\mu_{12}-\lambda)$
		$\Delta\mu$	$-\pi(\lambda+2\mu_1)$		$-\pi(3\mu_1-\mu_2)$
		μ_{12}	$-\pi(\lambda+2\mu_{12})$	$-3\pi\mu_{12}$	
		λ	$-3\pi\lambda$	$-\pi(2\lambda+\mu_{12})$	$-\pi(\lambda+2\mu_1)$
	$-$	any μ_*	$2\pi\mu_*$ + (the one with $-\mu_*$ above)		
++, A	both	$\pm\lambda$	$\mathbf{0}$		
		$\pm\mu_{12}$	$\mathbf{0}^*$		
		$\pm\Delta\mu$	$-2\mu_2$		

Table B.5: Estimates of the exponential suppression factors for all contributions to $\widetilde{F}$, assuming $p \gg 1$ and $\mu_1 \geq \mu_2$. For brevity, we only show the exponents of the suppression factors. Note that $\widetilde{F}^{(s_1,s_2\mu_{12})}_{++,\mathrm{A}}$ is polynomially suppressed and is indicated by the * sign.

Among these terms, the non-analytic signal is given by

$$\widetilde{F}^{(-,\mu_{12})}_{-+,\text{1-loop}} = \sum_{k=0}^{\infty} \widetilde{F}^{(-,\mu_{12},k)}_{-+,\text{1-loop}}, \qquad \widetilde{F}^{(-,\mu_{12},k)}_{-+,\text{1-loop}} \coloneqq \widetilde{F}^{(-)}_{-+}\bigg(\mu_{12} + 2ik + \frac{3i}{2}\bigg)\hat{\rho}_k(\mu_1,\mu_2), \tag{B.223}$$

where

$$\begin{aligned}\hat{\rho}_k(\mu_1,\mu_2) &\coloneqq 2\pi i \,\text{Res}\bigg[\rho_{\sigma_1\sigma_2}(\mu), \mu_{12} + 2ik + \frac{3i}{2}\bigg] \\ &= \frac{(-)^k(2k+1)!!}{2^{k+2}\pi^{5/2}k!}\frac{\prod_{i=1,2,12}\Gamma(-k+i\mu_i)\Gamma\big(\frac{3}{2}+k-i\mu_i\big)}{\Gamma\big(-\frac{3}{2}-2k+i\mu_{12}\big)\Gamma\big(\frac{3}{2}+2k-i\mu_{12}\big)\Gamma(-2k+i\mu_{12})\Gamma(3+2k-i\mu_{12})}.\end{aligned} \tag{B.224}$$

The remaining four terms are all resonance backgrounds. The background terms from $\widetilde{F}_{-+}$ are given by

$$\widetilde{F}^{(-,\lambda)}_{-+,\text{1-loop}} = \sum_{k=0}^{\infty} \widetilde{F}^{(-,\lambda,k)}_{-+,\text{1-loop}} \coloneqq \sum_{k=0}^{\infty} \widetilde{F}^{(-,\lambda,k)}_{-+}\rho^{\text{dS}}_{\mu_1\mu_2}\bigg(\lambda + ik + \frac{i}{2}\bigg), \tag{B.225}$$

$$\widetilde{F}^{(-,\lambda,k)}_{-+} \coloneqq 2\pi i \,\text{Res}\bigg[\widetilde{F}^{(-,\lambda,k)}_{-+}(\mu), \lambda + ik + \frac{i}{2}\bigg]. \tag{B.226}$$

where the residues are explicitly given by[16]

$$\widetilde{F}^{(-,\lambda,0)}_{-+} := -\frac{\sqrt{\pi}\mathcal{F}}{8}\Gamma\begin{pmatrix}1-2i\lambda,-\frac{1}{2}+i\lambda\\2-i\lambda\end{pmatrix}\frac{2^{2i\lambda}e^{\pi\lambda}}{1+2i\lambda}\left[(1-2i\lambda)p+\frac{c_1^+}{p}\widetilde{\mathbf{F}}^1_{-\frac{1}{2}+i\lambda}(p)\right], \tag{B.229}$$

$$\begin{aligned}\widetilde{F}^{(-,\lambda,1)}_{-+} :=& -\frac{\sqrt{\pi}\mathcal{F}}{12}\Gamma\begin{pmatrix}2-2i\lambda,-\frac{3}{2}+i\lambda\\1-i\lambda\end{pmatrix}\frac{2^{2i\lambda}e^{\pi\lambda}}{2i\lambda}\left\{\frac{\partial}{\partial s}\widetilde{\mathbf{F}}^s_{-\frac{3}{2}+i\lambda-s}(p)\Big|_{s=0}\right.\\&+\sum_{n=1}^{2}\frac{d_n}{2^n}\widetilde{\mathbf{F}}^n_{-\frac{3}{2}+i\lambda}(p)-\frac{1}{2}\sum_{\epsilon,\epsilon_1,\epsilon_2=\pm1}\epsilon\psi\left[\frac{3(1-\epsilon)}{4}+\frac{i}{2}\epsilon(\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2)\right]\\&\left.+\psi(i\lambda)-\psi(2-2i\lambda)-\frac{1}{1-i\lambda}-\frac{1}{2-i\lambda}+\frac{1}{2i\lambda}-\frac{4}{3}\right\},\end{aligned} \tag{B.230}$$

$$\begin{aligned}\widetilde{F}^{(-,\lambda,k\geq2)}_{-+} :=& \frac{(-)^k\sqrt{\pi}\mathcal{F}}{2^{2k}k!(k+2)}\Gamma\begin{pmatrix}1+k-2i\lambda,-\frac{1}{2}-k+i\lambda\\1-i\lambda\end{pmatrix}\\&\times\frac{2^{2i\lambda}e^{\pi\lambda}}{1-k+2i\lambda}\sum_{n=0}^{2}\frac{(n+k-2)!d_n}{2^n}\widetilde{\mathbf{F}}^{-1+n+k}_{-\frac{1}{2}-k+i\lambda}(p).\end{aligned} \tag{B.231}$$

The background terms from $\widetilde{F}_{++,\text{NA}}$ can be expressed in term of those from $\widetilde{F}_{-+}$ as

$$\widetilde{F}^{(+,-\lambda)}_{++,\text{NA,1-loop}} = \sum_{k=0}^{\infty}\widetilde{F}^{(+,-\lambda,k)}_{++,\text{NA,1-loop}} := \sum_{k=0}^{\infty}\widetilde{F}^{(+,-\lambda,k)}_{++,\text{NA}}\rho^{\text{dS}}_{\mu_1\mu_2}\left(-\lambda+ik+\frac{i}{2}\right), \tag{B.232}$$

$$\widetilde{F}^{(+,-\lambda,k)}_{++,\text{NA}} = (-)^k e^{2\pi\lambda}\widetilde{F}^{(+,-\lambda,k)}_{-+} + \delta_{k1}\cdot\frac{i\pi^{3/2}\mathcal{F}}{12}\Gamma\begin{pmatrix}2+2i\lambda,-\frac{3}{2}-i\lambda\\1+i\lambda\end{pmatrix}\frac{2^{-2i\lambda}e^{\pi\lambda}}{2i\lambda}, \tag{B.233}$$

[16] $\psi(z) =: \frac{\Gamma'(z)}{\Gamma(z)}$ is the digamma function. The derivative of hypergeometric function in $\widetilde{F}^{(-,\lambda,1)}_{-+}$ can be expressed in terms of Meijer-G functions as

$$\begin{aligned}\frac{\partial}{\partial s}\widetilde{\mathbf{F}}^s_{-\frac{3}{2}+i\lambda-s}(p)\Big|_{s=0} &= -\frac{\Gamma(4-2i\lambda)}{\Gamma(2-i\lambda)}G^{22}_{33}\left(\begin{matrix}0,-1+i\lambda;1\\0,0;-3+2i\lambda\end{matrix};\frac{2}{p-1}\right)+\psi(4-2i\lambda)-\psi(2-i\lambda)\\&= -\log\frac{p}{2}+\frac{1}{2(5-2i\lambda)p^2}+\frac{3}{4(5-2i\lambda)(7-2i\lambda)p^4}+\cdots.\end{aligned} \tag{B.227}$$

To prove this, one uses the eq. (16.17.2) in [71] to convert

$$G^{22}_{33}\left(\begin{matrix}0,-1+i\lambda;1\\0,z;-3+2i\lambda\end{matrix};\frac{2}{p-1}\right) \tag{B.228}$$

into hypergeometric functions, and then take the limit $z\to0$.

where $\widetilde{F}^{(+,-\lambda,k)}_{-+}$ is equal to $\widetilde{F}^{(-,\lambda,k)}_{-+}$ with λ replaced by $-\lambda$. Finally, the background terms from $\widetilde{F}_{++,\mathrm{A}}$ are given by

$$\begin{aligned}
\widetilde{F}^{(s_1,-s_1\lambda)}_{++,\mathrm{A,1\text{-}loop}} &= \sum_{k=0,k\neq 1}^{\infty} \widetilde{F}^{(s_1,-s_1\lambda,k)}_{++,\mathrm{A}} \widetilde{F}^{(s_1,-s_1\lambda,k)}_{++,\mathrm{A,1\text{-}loop}} + \sum_{m=0}^{\infty} \widetilde{F}^{(s_1,-s_1\lambda,1,m)}_{++,\mathrm{A,1\text{-}loop}} \\
&:= \sum_{k=0,k\neq 1}^{\infty} \widetilde{F}^{(s_1,-s_1\lambda,k)}_{++,\mathrm{A}} \rho^{\mathrm{dS}}_{\mu_1\mu_2}\left(-s_1\lambda + ik + \frac{i}{2}\right) \\
&\quad + \sum_{m=0}^{\infty} \widetilde{F}^{(s_1,-s_1\lambda,1,m)}_{++,\mathrm{A}} \rho^{\mathrm{dS}}_{\mu_1\mu_2}\left(-s_1\lambda + \frac{3i}{2}\right), \qquad \text{(B.234)} \\
\widetilde{F}^{(s_1,s_1\lambda)}_{++,\mathrm{A,1\text{-}loop}} &= \sum_{m=0}^{\infty} \widetilde{F}^{(s_1,s_1\lambda,m)}_{++,\mathrm{A,1\text{-}loop}} := \sum_{m=0}^{\infty} \widetilde{F}^{(s_1,s_1\lambda,m)}_{++,\mathrm{A}} \rho^{\mathrm{dS}}_{\mu_1\mu_2}\left(s_1\lambda + \frac{3i}{2}\right), \qquad \text{(B.235)} \\
\widetilde{F}^{(s_1,s_2\mu_{12})}_{++,\mathrm{A,1\text{-}loop}} &= \sum_{k=0}^{\infty} \widetilde{F}^{(s_1,s_2\mu_{12},k)}_{++,\mathrm{A,1\text{-}loop}} := \sum_{k=0}^{\infty} \widetilde{F}^{(s_1)}_{++,\mathrm{A}}\left(s_2\mu_{12} + 2ik + \frac{3i}{2}\right)\hat{\rho}_k(s_2\mu_1, s_2\mu_2), \qquad \text{(B.236)}
\end{aligned}$$

where, for $s_1 = +$,

$$\widetilde{F}^{(+,-\lambda,k\neq 1)}_{++,\mathrm{A}} := \frac{2\pi\mathcal{F}}{(k-1)(k+2)} \frac{(-)^k}{(1-k-2i\lambda)(1+2k+2i\lambda)(2+k+2i\lambda)} \times \frac{(-k-i\lambda)_k}{(-2k-2i\lambda)_k} \sum_{n=0}^{2} \frac{(k+1)_n c_n^+}{2^n} \widetilde{\mathbf{F}}^{1+n+k}_{-\frac{1}{2}-k-i\lambda}(p), \tag{B.237}$$

$$\widetilde{F}^{(+,-\lambda,1,m\neq 1)}_{++,\mathrm{A}} := \frac{2\pi\mathcal{F}}{3(m-1)} \frac{(-)^m}{2i\lambda(3+2i\lambda)^2} \frac{(-1-i\lambda)_m}{(-2-2i\lambda)_m} \sum_{n=0}^{2} \frac{(m+1)_n c_n^+}{2^n} \widetilde{\mathbf{F}}^{1+n+m}_{-\frac{3}{2}-i\lambda}(p), \tag{B.238}$$

$$\begin{aligned}\widetilde{F}^{(+,-\lambda,1,1)}_{++,\mathrm{A}} := \frac{2\pi\mathcal{F}}{3} \frac{1}{2i\lambda(3+2i\lambda)} \Bigg\{ & \frac{-1}{3+2i\lambda} \sum_{n=0}^{1} \frac{f_n}{2^n(p-1)} \widetilde{\mathbf{F}}^{1+n}_{-\frac{3}{2}-i\lambda}(p) \\ & - \frac{1}{2} \sum_{\epsilon,\epsilon_1,\epsilon_2=\pm 1} \epsilon\psi\left[\frac{3(1-\epsilon)}{4} + \frac{i}{2}\epsilon(-\lambda+\epsilon_1\mu_1+\epsilon_2\mu_2)\right] \\ & - \frac{2\pi i}{\tanh 2\pi\lambda} + \frac{1}{3+2i\lambda} - \frac{1}{2+i\lambda} - \frac{1}{1+i\lambda} - \frac{1}{2i\lambda} - \frac{1}{3} \Bigg\},\end{aligned} \tag{B.239}$$

$$\widetilde{F}^{(+,\lambda,m)}_{++,\mathrm{A}} := \frac{2\pi\mathcal{F}}{3} \frac{(-)^{1+m}}{2i\lambda(3-2i\lambda)^2} \frac{(-1+i\lambda)_m}{(-2+2i\lambda)_m} \sum_{n=0}^{2} \frac{(m+1)_n c_n^+}{2^n(-1+m+2i\lambda)} \widetilde{\mathbf{F}}^{1+n+m}_{-\frac{3}{2}+i\lambda}(p), \tag{B.240}$$

and

$$\begin{cases} f_0 := p^2 - 2Q, \\ f_1 := 2pQ. \end{cases} \tag{B.241}$$

For $s_1 = -$, one simply replaces all λ with $-\lambda$.

The asymptotic estimation eq. (B.124) implies that the infinite sum over residues is also

an series expansion over $1/p$ with the following power-counting rules:

$$\widetilde{F}^{(-,\mu_{12},k)}_{-+,\text{1-loop}} \sim \mathcal{O}\big(p^{-3-2k}\big), \tag{B.242}$$

$$\widetilde{F}^{(-,\lambda,k)}_{-+,\text{1-loop}} \sim \widetilde{F}^{(+,-\lambda,k)}_{++,\text{NA,1-loop}} \sim \begin{cases} \mathcal{O}\big(p^{-2-k}\big), & k \neq 1, \\ \mathcal{O}\big(p^{-3}\log p\big), & k = 1, \end{cases} \tag{B.243}$$

$$\widetilde{F}^{(s_1,-s_1\lambda,k)}_{++,\text{A,1-loop}} \sim \mathcal{O}\big(p^{-2-k}\big), \tag{B.244}$$

$$\widetilde{F}^{(s_1,-s_1\lambda,1,m)}_{++,\text{A,1-loop}} \sim \widetilde{F}^{(s_1,s_1\lambda,m)}_{++,\text{A,1-loop}} \sim \mathcal{O}\big(p^{-2-m}\big), \tag{B.245}$$

$$\widetilde{F}^{(s_1,s_2\mu_{12},k)}_{++,\text{A,1-loop}} \sim \mathcal{O}\big(p^{-2}\big). \tag{B.246}$$

For F_{-+} and $F_{-+,\text{NA}}$, the series is convergent for any $p > 1$, while for $F_{++,\text{A}}$, the series is at least convergent for $p > 3$ as indicated by eq. (B.218). Note that $\widetilde{F}^{(s_1,s_2\mu_{12},k)}_{++,\text{A,1-loop}}$ are $\mathcal{O}(p^{-2})$ for all k. This explains why the difference between the numerical result and the sum over poles does not approach zero in figure B.12 at large p.

B.9.6 Corrections to the Power Spectrum

In this appendix, we evaluate the corrections to the power spectrum by our model. We first evaluate the correction for the tree-level model given by (B.70). At $\mathcal{O}\big(|\alpha|^2\big)$, there are three diagrams contributing to the power spectrum as shown in figure B.19. The first diagram is simply given by

$$2k^3\delta\mathcal{P}^{(1)}_{\phi} = -\frac{2\,|\chi_0|^2}{\Lambda^2}. \tag{B.247}$$

The second diagram is given by

$$
\begin{aligned}
2k^3\delta\mathcal{P}_\phi^{(2)} &= 2k^3\left|\chi_0\right|^2\left|e^{-\pi\mu/2}\sqrt{\frac{\pi}{2}}\int_{-\infty_+}^{0}\frac{\mathrm{d}\eta}{\eta^{5/2}}\cdot 2(-\eta)^{1-i\lambda}\frac{\lambda}{\Lambda}\partial_\eta D_-(k,\eta)H_{i\mu}^{(1)}(-k\eta)\right|^2 \\
&\quad + (\lambda\to-\lambda) + \text{c.c.} \\
&= \frac{2\lambda^2\left|\chi_0\right|^2}{\Lambda}\left|\int_0^{\infty_+}\mathcal{I}_{1,-+}^{(-)}(\lambda;x)\,\mathrm{d}x\right|^2 + (\lambda\to-\lambda) \\
&= \frac{4\lambda^2\left|\chi_0\right|^2\cosh\pi\lambda}{\Lambda^2}\left|\mathbf{F}_{i\mu}^{\frac{1}{2}-i\lambda}(1)\right|^2 \\
&= \frac{2\left|\chi_0\right|^2}{\Lambda^2}\frac{2\pi\lambda\sinh 2\pi\lambda}{\cosh 2\pi\lambda+\cosh 2\pi\mu},
\end{aligned}
\tag{B.248}
$$

where $\mathcal{I}_{1,\pm+}^{(\pm)}$ is defined by eq. (B.76). Finally, the third diagram can be evaluated in the same manner as in section B.4.3, except now we can utilize the symmetry between the two vertices and write

$$
\begin{aligned}
2k^3\delta\mathcal{P}_\phi^{(3)} &= -2k^3\left|\chi_0\right|^2\int_{-\infty_+}^{0}\frac{\mathrm{d}\eta_2}{\eta_2^4}\int_{-\infty_+}^{0}\frac{\mathrm{d}\eta_1}{\eta_1^4}\cdot 2(-\eta_1)^{1-i\lambda}\frac{\lambda}{\Lambda}\partial_{\eta_1}D_+(k,\eta_1) \\
&\quad\times 2(-\eta_2)^{1+i\lambda}\frac{\lambda}{\Lambda}\partial_{\eta_2}D_+(k,\eta_2)G_{++}(k,\eta_1,\eta_2) + (\lambda\to-\lambda) + \text{c.c.} \\
&= -\frac{\lambda^2\left|\chi_0\right|^2}{\Lambda^2}\int_0^{\infty_+}\mathrm{d}x_2\int_0^{x_2}\mathrm{d}x_1\left[\mathcal{I}_{1,++}^{(-)}(\lambda;x_1)\mathcal{I}_{1,-+}^{(-)*}(\lambda;x_2) + (\lambda\to-\lambda)\right] + \text{c.c.}
\end{aligned}
\tag{B.249}
$$

Then, by applying eqs. (B.177) and (B.179), we obtain

$$
\begin{aligned}
2k^3\delta\mathcal{P}_\phi^{(3)} &= \frac{i\lambda^2\left|\chi_0\right|^2}{4\pi\Lambda^2}\frac{e^{\pi\mu}}{\sinh\pi\mu}\int_L\frac{\mathrm{d}t}{\frac{1}{2}-i\lambda+i\mu+t}\Gamma\begin{pmatrix}-t,1+t,\frac{1}{2}+i\mu+t\\ \frac{3}{2}+i\mu+t\end{pmatrix} \\
&\quad +(\lambda\to-\lambda,\mu\to-\mu)+\text{c.c.} \\
&= -\frac{\lambda^2\left|\chi_0\right|^2}{2\Lambda^2}\frac{e^{\pi\mu}}{\sinh\pi\mu}\sum_{n=0}^{\infty}\frac{(-)^n}{\left(\frac{1}{2}+n+i\mu\right)\left(\frac{1}{2}+n-i\lambda+i\mu\right)} \\
&\quad +(\lambda\to-\lambda,\mu\to-\mu)+\text{c.c.} \\
&= \frac{i\left|\chi_0\right|^2}{\Lambda^2}\frac{\lambda e^{\pi\mu}}{4\sinh\pi\mu}\left[\psi\left(\frac{1}{4}-\frac{i\lambda-i\mu}{2}\right)-\psi\left(\frac{3}{4}-\frac{i\lambda-i\mu}{2}\right)\right] \\
&\quad +(\lambda\to-\lambda,\mu\to-\mu)+\text{c.c.}
\end{aligned}
\tag{B.250}
$$

Note that a similar calculation without chemical potential has been carried out in ref. [108]. Our results for the second and the third diagrams agree with their expressions in the limit $\lambda\to 0$.

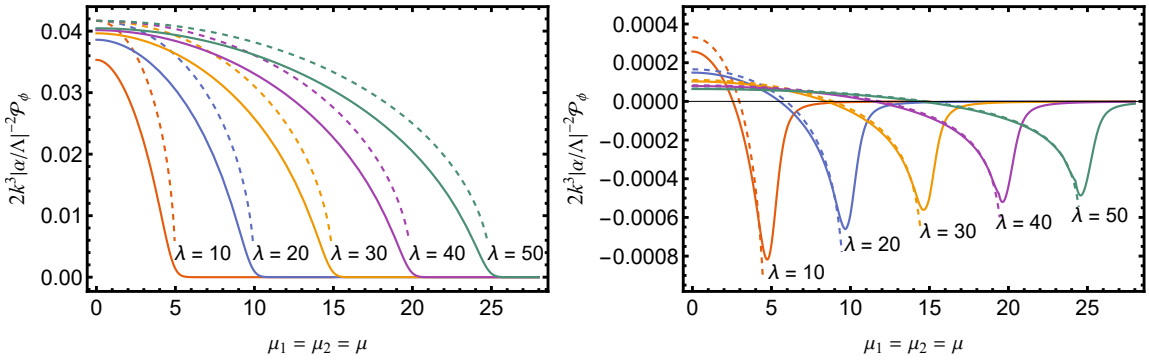


Figure B.25: The corrections to the power spectrum in the one-loop model for various λ and $\mu_1=\mu_2=\mu$ from $\delta\mathcal{P}_\phi^{(2)}$ (left) and $\delta\mathcal{P}_\phi^{(3)}$ (right). The dashed lines are estimations from eq. (B.252).

For the one-loop model, the correction can also be derived from the tree-level one using

the Källén–Lehmann representation:

$$\delta\mathcal{P}_\phi^{(\text{1-loop})}(\mu_1,\mu_2)=\int_0^\infty \delta\mathcal{P}_\phi^{(\text{tree})}(\mu)\rho^{\text{dS}}_{\mu_1\mu_2}(\mu)\,\mathrm{d}\mu\,. \tag{B.251}$$

For $\lambda,\mu_1,\mu_2,\lambda-\mu_{12}\gg 1$, all three contributions peak around $\mu\simeq\lambda$ and we can expand the integrand around this point, giving:

$$\begin{aligned}
2k^3\delta\mathcal{P}_\phi^{(\text{1,1-loop})}(\mu_1,\mu_2)&\simeq -\frac{|\alpha|^2}{\Lambda^2}\cdot\frac{\pi}{3\lambda^2}\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda),\\
2k^3\delta\mathcal{P}_\phi^{(\text{2,1-loop})}(\mu_1,\mu_2)&\simeq \frac{|\alpha|^2}{\Lambda^2}\cdot\frac{\pi^2}{3\lambda}\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda),\\
2k^3\delta\mathcal{P}_\phi^{(\text{3,1-loop})}(\mu_1,\mu_2)&\simeq \frac{|\alpha|^2}{\Lambda^2}\cdot\frac{1}{\lambda^2}\Big\{\frac{\pi}{12}\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda)+\zeta\big[\rho^{\text{Mink}}_{\mu_1\mu_2}(\lambda)-\lambda\rho'^{\text{Mink}}_{\mu_1\mu_2}(\lambda)\big]\Big\},
\end{aligned} \tag{B.252}$$

where

$$\zeta\coloneqq\frac{1}{4}\int_{-\infty}^{\infty}\frac{-ix\big[\psi\big(\frac{1}{4}+\frac{ix}{2}\big)-\psi\big(\frac{3}{4}+\frac{ix}{2}\big)\big]}{\frac{9}{4}+x^2}\,\mathrm{d}x=0.482003\cdots\,. \tag{B.253}$$

For a more accurate result, one has to carry out numerical calculations and the results are presented in figure B.25 for $\delta\mathcal{P}_\phi^{(2)}$ and $\delta\mathcal{P}_\phi^{(3)}$.

Appendix C: Warm Inflation with Pseudo-scalar Couplings

Edward Broadberry[1], Anson Hook[1], Sagnik Mondal[1]

[1]Maryland Center for Fundamental Physics, Department of Physics,

University of Maryland, College Park, MD 20742, USA

Uploaded to ArXiv https://arxiv.org/abs/2505.07943,

Submitted to JHEP

Abstract:

Inflaton couplings during warm inflation result in the production of a thermal bath. Thermal friction and fluctuations can dominate the standard de Sitter analogues, resulting in a modified slow-roll scenario with a new source of density fluctuations. Due to issues with back-reaction, it is advantageous to consider inflaton couplings with the thermal bath that are pseudo-scalar in nature, e.g., derivative interactions or topological $F\tilde{F}$ couplings. We demonstrate that every single existing model of warm inflation utilizing pseudo-scalar couplings needs to be corrected to properly account for all of the chemical potentials that the thermal bath acquires in response to the inflaton coupling. These chemical potentials are for non-conserved charges, and are non-zero only because of the applied inflaton couplings. The model-dependent chemical potentials modify the fluctuation-dissipation theorem, making the relationship between the thermal friction and thermal fluctuations model-dependent. In extreme cases, these chemical potentials can cause the friction term to vanish while thermal fluctuations remain non-zero. In the context of a simple example, we demonstrate how to calculate the chemical potentials, thermal friction, and thermal

fluctuations using both the Boltzmann equations and by calculating thermal expectation values, showing explicitly that the two approaches give the same result.

C.1 Introduction

The observable universe has several intriguing features that are all elegantly solved by the hypothesis that the universe underwent an extended period of inflation [2, 134]. At very large scales, the universe is approximately homogeneous with very low curvature. This homogeneity is extremely surprising when we realize that the many patches of space that we observe today appear to have never been in causal contact with one another. Requiring extremely homogeneous initial conditions is unappealing and begs a dynamical explanation.

Inflation solves all of the above problems by diluting away curvature and rapidly homogenizing the universe. An additional intriguing feature of inflation is that it can also explain the small 10^{-5} level in-homogeneities of the universe as well. Quantum fluctuations are blown up and expanded to cosmological scales and are the origin of all structure in the universe. The idea that our galaxy is only present due to a small quantum fluctuation of the early universe is an intriguing consequence of inflation.

Aside from all its virtues, inflation does have some flaws. The simplest and most appealing models of inflation are ruled out by measurements of the scalar-to-tensor ratio r [17]. Inflation also tends to require field excursions that are $\mathcal{O}(1)$ super-Planckian, which may or may not be an issue from a more ultraviolet point of view. Finally, slow-roll inflation requires a potential that is somewhat flat.

Warm inflation provides an interesting take on cold inflation. The idea of warm inflation

is that some coupling of the inflaton allows it to constantly lose energy populating a thermal bath [78–81] (for a review see [82]). This energy loss generates a friction term which effectively adds to Hubble friction, even dominating over it in specific cases. Additionally, the thermal bath has classical thermal fluctuations, whose interactions with the inflaton can be the dominant source of its fluctuations.

On a more conceptual level, warm inflation provides a counterpoint where the observed density fluctuations are the consequence of classical thermal fluctuations, rather than quantum fluctuations. As such, it is a natural playground in which to explore experimental signals of the classical versus quantum nature of the origin of density perturbations. Additionally, the new source of friction can mitigate some of inflation's perceived flaws, such as the super-Planckian field excursions.

For all of its attractiveness, it has been notoriously difficult to write down a theory of warm-inflation. The original approach towards warm inflation model building involved utilizing scalar couplings between the inflaton and particles making up the thermal bath. Unfortunately, it was subsequently realized that for scalar couplings, friction terms are a sub-leading effect and the leading effect is to generate a thermal potential that makes inflation harder instead of easier [84].

Over the years, there have been several different attempts to solve this problem. For example, one could endeavor to cancel the finite-temperature potential while keeping the friction term [85, 86]. Another popular approach, the one we will focus on in this article, is the use of pseudo-scalar couplings to mitigate the generation of the thermal potential [83, 87–91]. These couplings are quite natural as there are other reasons to suspect that the inflaton might be a pseudo-scalar as opposed to a scalar, a scenario sometimes called natural inflation [135].

Once the pseudo-scalar couplings have been specified, the framework of linear response theory can be used to calculate the friction and/or the fluctuations [136,137]. Because the two are connected by the fluctuation dissipation theorem, usually only one of the two is calculated and the other is obtained from the theorem. Once the fluctuations are specified, the corresponding two-point function of the inflaton can be derived and compared to experiment.

In this article, we point out that pseudo-scalar interactions to leading order act like applied chemical potentials and, as such, the thermal bath acquires opposite signed chemical potentials to counterbalance the applied chemical potentials such that the total chemical potential is zero. The chemical potentials for non-conserved charges can drastically change both the friction and fluctuations, but have been neglected previously. All warm inflation models which contain pseudo-scalar interactions need to have their friction and fluctuations recalculated.

As an example of the drastic modification due to a pseudo-scalar interaction, consider the leading shift-symmetric interaction connecting an inflaton to any other sector :

$$\delta\mathcal{L} = \frac{\partial_\mu \phi}{f} J^\mu. \tag{C.1}$$

In the presence of a slow-rolling inflaton, $\dot{\phi} \neq 0$, this term modifies the Hamiltonian by $\frac{\dot{\phi}}{f} Q$, which is clearly an applied chemical potential, we denote by μ_{applied}. The immediate response of the system is clear. It is energetically favorable to produce particles over anti-particles, and the thermal bath evolves towards a situation where there are more particles than anti-particles, which can be expressed in terms of a chemical potential μ_{dist} that appears in the distribution function. In the process, it extracts energy from the inflaton, inducing friction. However, the end result of this process is also equally clear. In thermal equilibrium, the chemical potential

for all non-conserved charges is zero. As such, the chemical potential present in the distribution functions, $\mu_{\rm dist}$, exactly balances out the applied chemical potential, $\mu_{\rm applied}$, so that in chemical equilibrium, the total chemical potential of the system, $\mu_{\rm total} = \mu_{\rm applied} + \mu_{\rm dist}$, vanishes, implying $\mu_{\rm dist} = -\mu_{\rm applied}$. Once this occurs, the bath is in an equilibrium state and no more energy is extracted from the inflaton. The friction therefore vanishes up to Hubble-induced corrections. Fluctuations will remain non-zero as the inflaton is still coupled to the bath.

What the previous example shows is that the backreaction of the thermal bath on the applied inflaton interaction can drastically change the resulting friction terms. The fact that the thermal bath will acquire chemical potentials for a generic pseudo-scalar interaction is to be expected from symmetries. Chemical potentials are CPT odd. When a pseudo-scalar is expanded around $\dot{\phi} \neq 0$, this vacuum expectation value (vev) combined with ϕ's CP breaking interactions mean that $\dot{\phi}/f$ acts as a CPT-breaking spurion. Thus it is actually expected that unless other symmetries forbid it, this background will generate chemical potentials $\mu_{\rm dist} \sim \dot{\phi}/f$.

Finally, there is a useful analogy to be made between warm inflation and stirring water with a straw. Water in a cup will evolve towards having no net velocity. Now consider a person stirring the water with a straw. Initially there will be friction on the straw. However, experience shows that eventually the water will start co-rotating with the straw and the friction will drastically decrease and in some limit vanish. This process does not change the thermal fluctuations present in the water.

In the context of warm inflation, a bath will naturally evolve towards having no net asymmetry between particles and anti-particles as long as number violating processes are present (water wants to be at rest). Now consider the situation where the inflaton interacts via an applied chemical potential (stirring the water with a straw). Initially there is friction that can

be calculated with linear response theory, γ_0, and is the standard result used for friction in the literature (the initial friction on the straw). However, eventually the bath responds with its own chemical potential, $\mu_{\rm dist}$, (water co-rotating with the straw) and the friction is suppressed and possibly vanishes, $\gamma = (1-d)\gamma_0$. This entire process does not change the amplitude of thermal fluctuations, demonstrating that it is possible to get fluctuations that are not proportional to the equlibrium value of friction.

In Sec. C.2, we work through a simple toy model to show how to calculate the chemical potentials developed by the bath in response to the inflaton interactions, as well as how to calculate the thermal fluctuations and the thermal friction on the inflaton. We do this both from the point of view of the Boltzmann equations, as well as simply by calculating a thermal expectation value of the relevant interaction in the equation of motion of ϕ. Next, in Sec. C.3, we go through other pseudo-scalar interactions showing that each of them also generates chemical potentials that need to be taken into account properly.

Our results mean that the friction terms are smaller than previously calculated. As such, in Sec. C.4 we show that warm inflation can still work with a ϕ^4 inflaton potential. Finally, we conclude in Sec. C.5.

In our appendices, we review a great deal of background for those who have not done the relevant calculations before. In App. C.6.1, we review the thermal field double that we use to calculate thermal expectation values. We calculate expectation values using the in-in formalism that we review in App. C.6.2. Since the fluctuation dissipation theorem is often used to relate the thermal friction and fluctuation terms, in App. C.6.4 we show that our results are consistent with this theorem once chemical potentials are included. In App. C.6.5, we review the calculation of the inflaton 2-point function. In the context of warm inflation, previous calculations of the

friction typically used linear response theory and the adiabatic approximation. In App. C.6.6, we show how linear response theory is connected to how we calculate the friction.

C.2 A Simple Toy Model

In this section, we present a simple toy model of warm inflation within which we will perform all of our calculations. The model will be chosen to make the calculations as simple as possible. The backbone of warm inflation is an equation of motion of the inflation that obeys

$$\ddot{\phi} + (3H + \gamma)\dot{\phi} + V'(\phi) = \xi. \tag{C.2}$$

There is a new source of friction, γ, which may or may not be larger than Hubble, and there is an extra term ξ on the right hand side. The expectation value of $\langle \xi \rangle = 0$, so it does not affect slow roll. However, its two-point function is non-zero, $\langle \xi(x)\xi(y) \rangle \propto \delta^4(x-y)$, and can dominate over the vacuum fluctuations that typically determine the inflaton two-point function. Our goal in this section is to write a model from which these two features can be easily calculated.

The thermal bath from which friction and fluctuations originate consists of a single complex scalar, χ, whose interactions are

$$\mathcal{V}_{\rm int} = m_\chi^2 |\chi|^2 + \lambda_\chi |\chi|^4 + \frac{\alpha_1}{4!}\chi^4 e^{4i\phi/f} + \frac{\alpha_2}{4!}\chi^4 e^{5i\phi/f} + h.c. \tag{C.3}$$

Thermal equilibrium of the bath is maintained by the large coupling $\lambda_\chi \sim \mathcal{O}(1)$, while the couplings to the inflaton $\alpha_{1,2}$ are chosen to be small, and $m_\chi \ll T$ is chosen so that the thermal bath acts like radiation with $w = 1/3$.

We will additionally be taking the chemical potential $\mu_{\rm dist} \ll m_\chi$, so that the induced chemical potential does not cause the formation of a Bose-Einstein condensate. For simplicity, we will take $\alpha_{1,2}$ to be real. In what follows, we will be working to leading order in $\alpha_{1,2}$ and taking $\mu_{\rm dist}, \dot{\phi}/f \ll T$.

Note also that the inflaton coupling is chosen to be such that if either $\alpha_{1,2}$ are zero, then the inflaton is derivatively coupled à la a chemical potential. To see this, let's consider the case where $\alpha_2 = 0$. In this situation, we can do the field redefinition $\chi \to e^{-i\phi/f}\chi$. ϕ reappears from the kinetic term as $\frac{\partial_\mu \phi}{f} J^\mu = \frac{\dot{\phi}}{f} J^0 \sim \mu_{\rm applied} Q$. Expanding around the slow-roll background shows that in this limit, the α_1 coupling acts like a chemical potential.

From the potential shown in Eq. C.3, we arrive at the equation of motion for the inflaton

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = -4\frac{i}{f}\frac{\alpha_1}{4!}\chi^4 e^{4i\phi/f} - 5\frac{i}{f}\frac{\alpha_2}{4!}\chi^4 e^{5i\phi/f} + h.c. \equiv O. \tag{C.4}$$

To make a connection with Eq. C.2, it will be convenient to separate the RHS into its expectation value and fluctuation as

$$\xi = O - \langle O \rangle. \tag{C.5}$$

In this section, we will use the Boltzmann equation to calculate both the chemical potential, $\mu_{\rm dist}$, of the radiation bath, and the induced friction on the inflaton. We will then use thermal perturbation theory to calculate

$$\langle O \rangle = -\gamma \dot{\phi}, \tag{C.6}$$

and find agreement with the result obtained through the Boltzmann equation.

Next, we calculate the two-point function $\langle \xi(x)\xi(y)\rangle$, which will ultimately determine the contribution of thermal fluctuations to the power spectrum. We will find that it agrees with the fluctuation-dissipation theorem, but only after chemical potentials have been included.

C.2.1 Response of the Thermal Bath

We first calculate how the thermal bath responds to the inflaton coupling in the limit where the Hubble time is large compared to all time scales in the problem. As is well known, there are only a limited number of ways in which a thermal bath can respond to its external environment, such as by adjusting its temperature and chemical potential. In the simple toy model we consider, there is only a single complex scalar present, so that the only chemical potential possible is χ's chemical potential, μ_{applied}.

To find how μ_{dist} and T respond to a constant $\dot{\phi}$ background, we utilize the Boltzmann equations. First focus on the evolution of the number density due to α_1. During slow-roll inflation, the inflaton coupling can inject or remove energy using

$$\mathcal{L}_{\text{int}} = -\frac{\alpha_1}{4!} e^{4i\frac{\dot{\phi}}{f}t}\chi^4, \tag{C.7}$$

as the $e^{4i\frac{\dot{\phi}}{f}t}$ factor modifies the energy-conserving delta function in the Boltzmann equation. At leading order in $\dot{\phi}/Tf$ and μ_{dist}/T, we only need to consider two-to-two scattering.

$$\chi(p_1) + \chi(p_2) \rightleftharpoons \chi^\dagger(p_3) + \chi^\dagger(p_4), \tag{C.8}$$

for which the Boltzmann equation is

$$\left(\frac{d\Delta n_\chi}{dt}\right)_{\alpha_1} = 4\alpha_1^2 \prod_{i,f} \int d\Pi_i d\Pi_f (2\pi)^4 \delta^3(\vec{p}_i - \vec{p}_f)\delta\left(E_i - E_f + 4\frac{\dot{\phi}}{f}\right) \times\left(\bar{f}_3\bar{f}_4(1+f_1)(1+f_2) - f_1 f_2(1+\bar{f}_3)(1+\bar{f}_4)\right), \quad \text{(C.9)}$$

where $\Delta n_\chi = n_\chi - n_{\chi^\dagger}$ and the distribution functions are given by

$$f_i = \frac{1}{e^{\beta(\omega_i - \mu_{\rm dist})} - 1} \implies \frac{1+f_i}{f_i} = e^{\beta(\omega_i - \mu_{\rm dist})}, \quad \text{(C.10)}$$

$$\bar{f}_i = \frac{1}{e^{\beta(\omega_i + \mu_{\rm dist})} - 1} \implies \frac{1+\bar{f}_i}{\bar{f}_i} = e^{\beta(\omega_i + \mu_{\rm dist})}. \quad \text{(C.11)}$$

Because $\alpha_{1,2} \ll \lambda_\chi$, kinetic equilibrium is maintained faster than chemical equilibrium and the distribution functions are the standard ones allowed by the Boltzmann H theorem. These identities allow us to expand to first order in $\dot{\phi}/Tf$ and $\mu_{\rm dist}/T$:

$$\left(\frac{d\Delta n_\chi}{dt}\right)_{\alpha_1} \approx 4\alpha_1^2\beta\left(4\frac{\dot{\phi}}{f} - 4\mu_{\rm dist}\right)\int \prod_{i=1}^{4} d\Pi_i f_1 f_2 \bar{f}_3 \bar{f}_4 e^{\beta(\omega_3+\omega_4)}(2\pi)^4\delta^4(\sum_i p_i) \quad \text{(C.12)}$$

$$\equiv 4c_1\alpha_1^2 T^3\left(4\frac{\dot{\phi}}{f} - 4\mu_{\rm dist}\right). \quad \text{(C.13)}$$

The phase space integral, c_1, can be calculated numerically. When $m_\chi \lesssim T$, we find that $c_1 \approx (0.0589 - 0.0632\log(m/T) - 0.143e^{-m/T})/(2\pi)^4$ is a good numerical fit; namely, c_1 is an $\mathcal{O}(1)$ number divided by a phase space factor of $(2\pi)^4$. A similar computation for α_2 gives the total asymmetry of χ evolving as

$$\frac{d\Delta n_\chi}{dt} = 4c_1\alpha_1^2 T^3\left(4\frac{\dot{\phi}}{f} - 4\mu_{\rm dist}\right) + 4c_1\alpha_2^2 T^3\left(5\frac{\dot{\phi}}{f} - 4\mu_{\rm dist}\right). \quad \text{(C.14)}$$

The evolution in time of the energy density is

$$\frac{d\rho}{dt} = 4\frac{\dot{\phi}}{f}c_1\alpha_1^2 T^3\left(4\frac{\dot{\phi}}{f} - 4\mu_{\rm dist}\right) + 5\frac{\dot{\phi}}{f}c_1\alpha_2^2 T^3\left(5\frac{\dot{\phi}}{f} - 4\mu_{\rm dist}\right), \tag{C.15}$$

which can be simply understood as the α_1 (α_2) processes depositing/removing $4\dot{\phi}/f$ ($5\dot{\phi}/f$) energy respectively. Because $\Delta n_\chi \sim \mu_{\rm dist}T^2$ and $\rho \sim T^4$, Eqs. C.14 and C.15 imply that $\dot{T}/T \ll \dot{\mu}_{\rm dist}/\mu_{\rm dist}$. Thus, while the thermal bath will slowly heat up, $\dot{T} > 0$, it will quickly reach chemical equilibrium with $\dot{\mu}_{\rm dist} = 0$, giving

$$\mu_{\rm dist} = \frac{\alpha_1^2 + \frac{5}{4}\alpha_2^2}{\alpha_1^2 + \alpha_2^2}\frac{\dot{\phi}}{f}, \tag{C.16}$$

$$\frac{d\rho}{dt} = \frac{\alpha_1^2\alpha_2^2}{\alpha_1^2 + \alpha_2^2}c_1 T^3\left(\frac{\dot{\phi}}{f}\right)^2. \tag{C.17}$$

In the limit that α_1 or α_2 are zero, the inflaton coupling is a chemical potential coupling, and we find the expected result that the system applies an equal and opposite dynamically generated chemical potential and the system reaches an equilibrium where $\dot{T} = \dot{\mu}_{\rm dist} = 0$.

Amusingly, at this point we have already calculated both the friction and fluctuations. Eq. C.17 describes how energy is deposited into the radiation bath as a function of time. Conservation of energy dictates that this energy comes from the kinetic energy of the inflaton ϕ, and thus we have already obtained the friction that the inflaton experiences. By the fluctuation dissipation theorem, discussed in more detail in App. C.6.4, this friction can be transformed into the inflaton fluctuations.

C.2.2 Thermal Friction

In this subsection, we discuss how to calculate the friction term. In the previous subsection, we calculated the energy deposition in the process of calculating the chemical potential using the Boltzmann equation. However, it is always useful to calculate the same quantity in multiple ways. Thus, in this subsection, we directly calculate $\langle \mathcal{O} \rangle$ from Eq. C.6 using thermal perturbation theory, reviewed in App. C.6.1. We start with the in-in correlator [1]

$$\langle O \rangle = \langle -4 \frac{i}{f} \frac{\alpha_1}{4!} \chi^4 e^{4i\phi/f} - 5 \frac{i}{f} \frac{\alpha_2}{4!} \chi^4 e^{5i\phi/f} + h.c. \rangle. \tag{C.18}$$

Let us focus on the α_1 term, as the α_2 calculation will be analogous. To leading order in the interactions, we have

$$\langle O_{\alpha_1} \rangle \;\; = \;\; -4 \frac{i}{f} \frac{\alpha_1}{4!} \langle i \int d^4y \, [\chi(x)^4 e^{4i\frac{\dot{\phi}}{f} t_x}, \mathcal{L}_{\rm int}(y)] \rangle + h.c. \tag{C.19}$$

Note that due to the explicit time dependence of the Lagrangian, many of the terms will vanish as one averages over times scales longer than $1/(\dot{\phi}/f)$. Meanwhile, other terms can be absorbed into the definition of the (thermal) potential of ϕ. The only remaining term in this expectation value is the term coming from $\mathcal{L}_{\rm int}(y)$ whose time dependence is exactly $e^{-4i\frac{\dot{\phi}}{f} t_y}$. Using this, we

[1] Since we are studying an equilibrium configuration, we could also use the in-out correlator and get the exact same result. The in-in correlator happens to be easier to calculate. See App. C.6.2 for more details.

find that

$$\begin{aligned}\langle O_{\alpha_1}\rangle &= -\frac{4}{f}\frac{\alpha_1^2}{4!}\int d^3y\int_{-\infty}^{t_x} dt_y e^{4i\frac{\dot\phi}{f}(t_x-t_y)}\left(\langle\chi(x)\chi^\dagger(y)\rangle^4-\langle\chi^\dagger(y)\chi(x)\rangle^4\right)+h.c.\\ &= -\frac{4}{f}\frac{\alpha_1^2}{4!}\int d^3y\int_{-\infty}^{t_x} dt_y\int\prod_j\frac{d^3k_j}{(2\pi)^3 2\omega_i}e^{4i\frac{\dot\phi}{f}(t_x-t_y)}\Big\{\left((1+f_j)e^{-ik_j\cdot(x-y)}+\bar f_j e^{ik_j\cdot(x-y)}\right)^4\\ &\qquad -\left((1+\bar f_j)e^{ik_j\cdot(x-y)}+f_j e^{-ik_j\cdot(x-y)}\right)^4\Big\}+h.c.,\end{aligned}\tag{C.20}$$

where in the second line we have used the thermal field propagators in Eq. C.89 and Eq. C.90.

There are several facts that can be used to simplify this expression. Firstly, the d^3y integral just gives momentum-conserving delta functions. While not as obvious, something similar happens for the time integral over t_y. t_y is only integrated over the half plane so that instead of giving a delta function in energy, its imaginary part is a principle value while its real part is half the typical delta function. The effect of adding the hermitian conjugate is to remove the principle value piece and just give the energy-conserving delta functions modified to include the energy deposition of the inflaton.

The rest of Eq. C.20 has interpretations coming from different terms in the Boltzmann equations. For example, the piece containing $(1+f_j)^4$ describes the process $0\to\chi^4$, a process energetically allowed only at higher orders in $\dot\phi/f$, but is nevertheless still present. At leading order, however, we obtain the Boltzmann equation for two-to-two scattering from the previous subsection. Putting it all together, including the symmetry factors for the phase space of

identical particles, we find

$$\begin{aligned}\langle O\rangle &= -\frac{4c_1}{f}\alpha_1^2 T^3\left(4\frac{\dot\phi}{f}-4\mu_{\rm dist}\right)-\frac{5c_1}{f}\alpha_2^2 T^3\left(5\frac{\dot\phi}{f}-4\mu_{\rm dist}\right)=-\frac{\alpha_1^2\alpha_2^2}{\alpha_1^2+\alpha_2^2}c_1\frac{T^3}{f^2}\dot\phi,\\ \gamma &= c_1\frac{\alpha_1^2\alpha_2^2}{\alpha_1^2+\alpha_2^2}\frac{T^3}{f^2}.\end{aligned} \tag{C.21}$$

In our model, $\alpha_{1,2}$ are dimensionless so that $\gamma\sim\alpha_{1,2}^2T^3$ by dimensional analysis. If instead χ^3 (χ^2) interactions were being considered, then again dimensional analysis would give $\gamma\sim\alpha_{1,2}^2T$ ($\gamma\sim\alpha_{1,2}^2/T$). This shows that simple modifications of our model will be able to accommodate different temperature scalings of the friction.

A good double check of our results is that

$$\langle O\rangle = -\frac{d\rho/dt}{\dot\phi} \tag{C.22}$$

holds even when the chemical potential is not at its equilibrium value and without needing to take the $\mu_{\rm dist},\dot\phi/f\ll T$ limit. It is refreshing that while our friction coefficient was calculated using thermal perturbation theory, and energy deposition was calculated using the Boltzmann equations, the results agree.

C.2.3 Two Point Function

We now use thermal perturbation theory to calculate the two point function $\langle\xi(x)\xi(y)\rangle$. The first thing to note is that it involves a propagator from x to y. Famously, position space propagators will fall off exponentially with correlation length and time. As such, as long as we are

interested in distances much larger than the correlation length, it is reasonable to approximate

$$\langle \xi(x)\xi(y)\rangle \approx A\delta^4(x-y). \tag{C.23}$$

In the rest of the subsection, we will calculate the constant A. Selecting out the pieces that do not oscillate quickly and average away, we have

$$\begin{aligned} A &= \int d^4y \langle \xi(x)\xi(y)\rangle & \text{(C.24)}\\ &= \int d^4y \frac{1}{4!f^2}\big(\langle \chi(x)\chi^\dagger(y)\rangle^4 \left(16\alpha_1^2 e^{4i\frac{\dot\phi}{f}(t_x-t_y)} + 25\alpha_2^2 e^{5i\frac{\dot\phi}{f}(t_x-t_y)}\right) \\ &+ \langle \chi^\dagger(x)\chi(y)\rangle^4 \left(16\alpha_1^2 e^{-4i\frac{\dot\phi}{f}(t_x-t_y)} + 25\alpha_2^2 e^{-5i\frac{\dot\phi}{f}(t_x-t_y)}\right)\big). & \text{(C.25)} \end{aligned}$$

The only difference between this equation and our previous Eq. C.20 comes down to a minus sign. Due to the forward and backward process depositing or removing energy, Eq. C.20 has a piece $e^{-\beta(4\dot\phi/f-2\mu_{\rm dist})} - e^{-\beta 2\mu_{\rm dist}} \approx \beta(4\mu_{\rm dist} - 4\dot\phi/f)$. Due to factors of i difference, the current two point function, Eq. C.25, has $e^{-\beta(4\dot\phi/f-2\mu_{\rm dist})} + e^{-\beta 2\mu_{\rm dist}} \approx 2$.

As a result, after a bit of algebra we find that

$$\langle \xi(x)\xi(y)\rangle \approx \frac{2c_1T^4}{f^2}\left(16\alpha_1^2+25\alpha_2^2\right)\delta^4(x-y); \qquad A = \frac{2c_1T^4}{f^2}\left(16\alpha_1^2+25\alpha_2^2\right). \tag{C.26}$$

We can combine this result with the friction term from the previous subsection to find

$$\begin{aligned} A &= 2\left(\gamma\frac{16\alpha_1^2+25\alpha_2^2}{16\alpha_1^2\left(1-\frac{\mu_{\rm dist}}{\dot\phi/f}\right)+25\alpha_2^2\left(1-\frac{4\mu_{\rm dist}}{5\dot\phi/f}\right)}\right)T \\ &= 2\gamma T\frac{(\alpha_1^2+\alpha_2^2)(16\alpha_1^2+25\alpha_2^2)}{\alpha_1^2\alpha_2^2}. \end{aligned} \tag{C.27}$$

While not immediately obvious, as discussed in App. C.6.4, this result for the fluctuation is what we expect from the fluctuation dissipation theorem. As can also be seen from Eq. C.27, if one were to neglect the chemical potential, $\mu_{\rm dist} = 0$, then the standard result of $A = 2\gamma T$ would be obtained. From our results, we see that zero chemical potentials is an important implicit assumption when it was originally claimed that the fluctuation dissipation theorem implies $A = 2\gamma T$ [96, 97].

We have seen in this section that the chemical potential will always act to reduce the friction coefficient, which fits with the analogy of stirring a water cup, where the water develops a current that travels along with the stirrer. We also found that at leading order the chemical potential has no effect on the size of the fluctuations, which is also consistent with this analogy. It is therefore clear that in Eq. C.27, the fluctuation-dissipation theorem is modified not by enhancing the fluctuations, but by reducing the dissipation. This is clearly a general statement that is applicable to any model of warm inflation with a chemical potential.

C.3 Other Examples of Chemical Potentials

In this section, we discuss examples other than our toy model in order to demonstrate the ubiquity of chemical potentials. Because of the number of models being discussed, we will consider each example only qualitatively, though one can obtain the $\mathcal{O}(1)$ numbers in a manner similar to the previous section. As before, our analyses will be done in the $\dot{\phi}/f \ll T$ and $H \approx 0$ limit.

In the language of the water stirring analogy in the introduction, we find three main take-aways from these examples:

1. The friction in chemical equilibrium is γ. Neglecting chemical potentials, would lead one to calculate γ_0 and miss the factor of $1 - d$.

2. If one calculates fluctuations and uses the fluctuation-dissipation theorem to find the friction, one has to account for chemical potentials or else one will find γ_0 instead of γ.

3. Finally, axion-like couplings lead to both perturbative and non-perturbative contributions to the friction.

While many of the couplings below will have $\gamma \approx 0$ due to a chemical potential when $\dot{\phi}/f \ll T$, this does not mean that it is impossible to do warm inflation with these couplings. Some of the couplings have equilibrium configuration where $\dot{\phi}/f \sim \mu_{\rm dist} \sim T$. Taking into account a non-zero Hubble, these couplings can support $\gamma \propto H$, possibly supporting warm inflation in the weak ($\gamma < H$) regime.

C.3.1 Chemical Potential Couplings

The first example we will illustrate is a chemical potential coupling of a different form. In particular, we will take the inflaton to be coupled to fermions $\psi_{1,2}$ with

$$\mathcal{L}_{\rm int} = \frac{\partial_\mu \phi}{f}(J^\mu_{\psi_1} + J^\mu_{\psi_2}), \tag{C.28}$$

which amounts to a chemical potential favoring particles over anti-particles. Additionally, we include a symmetry-breaking coupling of the form $\delta\mathcal{L} = \alpha\psi_1^2\psi_2^2$.

There are two ways in which we can use this coupling. Because this coupling is bilinear in the fermion fields, we can include it directly when decomposing the field in terms of creation

and annihilation operators:

$$\psi_{1,2}(x) \;=\; \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}} \sum_s \left(u_s(p)a_p e^{-ipx-i\mu_{\rm dist}t} + v_s(p)b_p^\dagger e^{ipx-i\mu_{\rm dist}t}\right). \tag{C.29}$$

This amounts to just changing the energy of particles and anti-particles to $E = \sqrt{p^2+m^2} \pm \dot{\phi}/f$. Note that in Eq. C.9, the effect of the coupling is purely to modify the energy conservation delta function, so it is clear that the results are the same whether one works in this basis or the other basis. Symmetry-breaking operators will cause a deposition of energy, and the induced thermal chemical potential will be such to exactly cancel the $\dot{\phi}/f$-induced energy differences.

A different approach to this coupling uses perturbation theory. Like when doing perturbation theory with a mass term, the previous answer can be obtained by doing a resummation of diagrams. When doing this calculation, one needs to go to higher order in perturbation theory to see a non-trivial answer. Schematically,

$${}_{\rm in}\langle \dot{J}^0_\psi(0)\rangle_{{\rm in},\beta} \sim \langle \dot{J}^0_\psi(0)\frac{\dot{\phi}}{f}J^0_\psi(x)\rangle_\beta + \alpha^2\frac{\dot{\phi}}{f}\langle \dot{J}^0_\psi(0)J^0_\psi(x)\psi_1^2(y)\psi_2^2(y)\psi_1^2(z)\psi_2^2(z)\rangle_\beta + \cdots \tag{C.30}$$

The effect of $\frac{\dot{\phi}}{f}J^0_\psi$ is to include the energy change $\pm\dot{\phi}/f$ for particles versus anti-particles at leading order in perturbation theory. However, notice that the leading term is zero. One needs interactions to create a connected diagram. This is equivalent to the statement that if $\alpha \to 0$, the chemical potential can be removed by a field redefinition. In particular, one needs three different insertions, $\frac{\dot{\phi}}{f}J^0_\psi$ and two $\alpha\psi_1^2\psi_2^2$ insertions, until a non-zero result is obtained. The expression obtained will be the fermionic analogue of Eq. C.13.

C.3.2 Symmetry Breaking Couplings

In this subsection, we will consider the interaction between the inflaton and scalar field, χ, to be

$$\mathcal{V}_{\rm int} = \frac{\alpha_1}{4!}\chi^4(1 + 4i\frac{\phi}{f}) + h.c. \tag{C.31}$$

Note that this interaction is similar to the previous Eq. C.3 with $\alpha_2 = 0$, but only the leading order term in $1/f$ is included. Because of the lack of higher-order terms, this is not a shift-symmetric chemical potential term. However, it will still induce a chemical potential response from the bath. The reason is that when we did all of our calculations in Sec. C.2, we eventually had to Taylor series our result in $\dot{\phi}/(fT)$. Because of this, the entirety of our calculation is completely insensitive to the exact form of the coupling $e^{4i\phi/f}$ and only to the first-order term in the Taylor series.

The most important aspect of the coupling in Eq. C.31 is the fact that the symmetry-breaking term is a pseudo-scalar interaction and has a factor of i. This i is critical in making the particle versus anti-particle interactions different by a minus sign and thus gives rise to a chemical potential.

C.3.3 Topological Axion Couplings

The final couplings that we discuss are axion couplings to abelian (F) and non-abelian (G) gauge groups of the form

$$\mathcal{V}_{\rm int} = \frac{\alpha\phi}{4f} F\tilde{F} + \frac{\alpha\phi}{4f} G\tilde{G}. \tag{C.32}$$

Both of these couplings have been used for warm inflation model building [83,89–91,101,102,138]. We will discuss the case where the gauge bosons are abelian or non-abelian separately. As mentioned before, we will be qualitative about $\mathcal{O}(1)$ numbers, but we will keep track of α as α counting can help clarify issues when light fermions become involved.

C.3.3.1 Abelian Gauge Groups - No Fermions:

As is well known, the axion coupling changes the dispersion relation of the gauge boson to

$$E^2 = p^2 \pm \alpha\, p \frac{\dot{\phi}}{f} + m^2, \tag{C.33}$$

where we have allowed for a mass term to stabilize the tachyon if need be. The $\pm$ depends on if the gauge boson is left- or right-circularly polarized. In the absence of a thermal bath, one of the two polarizations can have a tachyonic mode that grows exponentially quickly. The time scale of the exponential growth is $t \sim f/\dot{\phi}$. This removal of energy is exponential in nature and thus too fast to be accounted for with a simple friction term $\gamma\dot{\phi}$. Even without a thermal bath, it has been proposed to use this coupling to support inflation [139, 140].

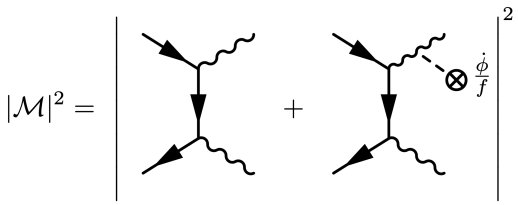


Figure C.1: The Feynman diagrams for the scattering process $e + e^- \to A_{L,R} + A_{L,R}$. The leading contribution in $\dot{\phi}/f$ comes from the cross term.

As it will be useful later on, observe that the dispersion relation in the limit of $p \sim T \gg \mu_{\rm dist}, m$ is

$$E = p \pm \alpha \frac{\dot{\phi}}{2f} + \cdots . \tag{C.34}$$

This modification looks similar to a chemical potential, but also leads to a plethora of effects, such as polarization rotation, that are often used to look for axion dark matter.

C.3.3.2 Abelian Gauge Groups - Fermions:

An abelian gauge theory by itself does not thermalize [2], so let us consider what occurs when there is a light fermion, $m_e \ll T$, that we will lazily call the electron. In the presence of a thermal bath, the tachyonic instability can be cured, as a thermally-induced photon mass $m \sim eT$ will be generated that eventually removes the instability. To obtain the friction, it is easiest to consider the Boltzmann Equation. As can be clearly seen from Eq. C.34, the leading-order effect of the axion coupling is to apply an external chemical potential for helicity, much like the situation in Sec. C.3.1. To see that the system will develop a chemical potential in response, let us consider a scattering process $e^+ + e^- \to A_L + A_L$, writing the left(right)-handed gauge

[2] If the inflaton comes into equilibrium, then thermal equilibrium can be obtained without fermions. This situation was numerically studied in Ref. [89] where it was found that the equilibrium configuration contained a chemical potential for the helicity of the gauge boson.

boson as $A_L(A_R)$, with Feynman diagrams shown in Fig. C.1. This helicity-violating process requires a mass insertion. If one wishes to avoid the possible mass suppression, consider instead the process $e^+ + e^- \to 3A$. The combination of these different processes makes it clear that the chemical potential that the system develops in response is the expected chemical potential for helicity of the gauge boson, and not a chemical potential for the helicity of the electron.

The end result of these scattering processes is that

$$\frac{dn_{A_R}}{dt} - \frac{dn_{A_L}}{dt} \sim (\alpha^2 m_e^2 T + \alpha^3 T^3) \times \left(\frac{\alpha \dot{\phi}}{f} - \mu_{\rm dist} \right), \tag{C.35}$$

where the factors of T are found by dimensional analysis. The thermal bath will develop its own chemical potential $\mu_{\rm dist} \sim \alpha\dot{\phi}/f$ to counteract the applied chemical potential.

As the system approaches chemical equilibrium, there will be a friction term

$$\gamma \sim \alpha^4 m_e^2 T/f^2, \text{ or } \gamma \sim \alpha^5 T^3/f^2.$$

However, once the system reaches chemical equilibrium, the friction will vanish, $\gamma \to 0$.

The only energy deposition will now come from the fact that at higher orders in $\dot{\phi}/fT$, the coupling is not exactly a chemical potential. The appearance of more powers of $\dot{\phi}$ indicates that while there will be energy deposition, it will not take the simple form of a friction term. Interestingly enough, the thermal bath responds to the applied chemical potential even if the electron is massless.

It is useful to study what occurs when one performs a field redefinition $\psi \to e^{i\phi/f}\psi$ to

move the topological coupling in Eq. C.32 into

$$\mathcal{L} = me^{i\phi/f}\psi\psi^c - \frac{\partial_\mu \phi}{f} J^\mu. \tag{C.36}$$

In this basis, it appears as if everything is about the chirality of the electron rather than the chirality of the photon. We will now describe how to show that, even in this basis, everything is about the chirality of the photon and not the electron.

In this basis, there are two effects. Any fermion chirality violation necessarily involves both a mass insertion, which deposits energy $\dot{\phi}/f$ from the mass term, and at the same time removes an energy $\dot{\phi}/f$ from the derivative interaction. Due to this minus sign, any process that changes the chirality of the fermion has no energy injection, and no chemical potential is obtained for the fermion helicity. The two effects in Eq. C.36 cancel each other out exactly.

The only effect of the combined couplings is on photon propagation via a loop of electrons. For example, the typical Compton scattering diagrams can interfere with analogous diagrams where one of the photons is dressed by electron loops, changing its dispersion relation. This process would give the rate Eq. C.35. As before, even if the electron is massless, a chemical potential for photon helicity still develops, shutting off the friction up to effects due to non-zero Hubble.

C.3.3.3 Non-Abelian Gauge Groups - No Light Fermions:

When discussing non-abelian gauge groups, there are both perturbative effects and non-perturbative effects. Let us discuss these contributions separately.

Perturbative Effects:

As with the abelian case, to leading order in the gauge coupling, g, the effect of Eq. C.32 is a chemical potential for helicity. At higher order in g, there are other effects. As we are working perturbatively, the main difference between non-abelian and abelian gauge groups is that non-abelian groups contain interactions that automatically violate helicity conservation as well as thermalize the system. As such, we will discuss thermal equilibrium and the Boltzmann equation from the beginning.

Let us first discuss the effect of the applied chemical potential for helicity. The leading-order violation of helicity in pure Yang-Mills theory is from two-to-three scattering process such as $A_R + A_R \to A_R + A_R + A_R$. This leads to a rate

$$\frac{dn_{A_R}}{dt} - \frac{dn_{A_L}}{dt} \propto \alpha^3 T^4 \times \left(\frac{\alpha \dot{\phi}}{fT} - \frac{\mu_{\rm dist}}{T} \right). \tag{C.37}$$

Aside from a chemical potential, there are also $A^3\dot{\phi}/f$ and $A^4\dot{\phi}/f$ couplings. These also contribute to helicity-violating processes at the same order as Eq. C.37. Due to the different kinematic structure of the vertices caused by the ϵ tensor, the interference that generates Eq. C.37 is a bit more subtle, but the overall expectation is something of nearly the same order.

Unlike the abelian case, even at leading order in $\dot{\phi}/f$, the coupling is not just an applied chemical potential. The situation is similar to our toy model, and we expect that after the system develops a chemical potential $\mu_{\rm dist} \propto \alpha\dot{\phi}/f$, there are still heating effects proportional

to $\dot{\phi}/f$. Namely, we expect that

$$\frac{d\rho_{\rm bath}}{dt} \propto \alpha^3 T^4 \times \frac{\alpha\dot{\phi}}{fT} \times \frac{\alpha\dot{\phi}}{f}, \tag{C.38}$$

where the first term is the scattering rate, the second term is the bias, and the last term is how much energy is injected into the bath per event. However, additional work is needed to make certain of this result.

Non-Perturbative Effects:

Sphalerons are a non-perturative effect that has similar scaling to the perturbative effects considered above, but whose order-one numbers make it less suppressed. The non-perturbative friction, assuming no chemical potentials that result from an axion coupled to a non-abelian gauge group, has been studied before in Ref. [98, 141] giving the result $\gamma \sim \alpha^5 T^3/f^2$. To better understand the origin of this friction and how a chemical potential for helicity can change it, we provide a simple understanding of the effect below.

Sphalerons are a process by which the Chern-Simons number changes by one with a rate $\sim \alpha^5 T^4$. Roughly speaking, sphalerons are CP-violating collective excitations of $1/\alpha$ particles. Thus, to leading order, we can interpret sphalerons as objects that create $1/\alpha$ more left-handed than right-handed gauge bosons. As discussed before, the effect of the $\phi G\tilde{G}$ coupling is to introduce a chemical potential for helicity. As with Eq. C.37, we find that the leading-order effect of sphalerons is

$$\frac{dn_{A_R}}{dt} - \frac{dn_{A_L}}{dt} \propto \alpha^5 T^4 \times \frac{1}{\alpha}\left(\frac{\alpha\dot{\phi}}{fT} - \frac{\mu_{\rm dist}}{T}\right). \tag{C.39}$$

The rate is suppressed compared to Eq. C.37; however, there is an enhancement that comes with each event, since it changes the helicity of $1/\alpha$ gauge bosons.

Before the chemical potential is taken into account, the energy deposited into the thermal bath mediated by sphalerons is

$$\frac{d\rho_{\rm bath}}{dt} \propto \alpha^5 T^4 \times \frac{1}{\alpha}\frac{\alpha\dot{\phi}}{fT} \times \frac{1}{\alpha}\frac{\alpha\dot{\phi}}{f}, \tag{C.40}$$

giving the friction term found in Ref. [98]. Note that this is parametrically similar to the perturbative result in Eq. C.38, but it comes about in a different way. The rate is suppressed, but since it deals with $1/\alpha$ particles, the end result is comparable.

The next step is to properly account for the effects of the chemical potential. If we only consider sphalerons, then it is possible that the chemical potential will cancel all helicity violation (up to terms higher order in α) and that friction will be suppressed. It is also equally possible that, much like the perturbative result, the chemical potential will not cancel out the effect entirely. Either way, eventually terms that are higher order in α or the perturbative effects will render the friction non-zero. However, these effects are either suppressed by α or by many phase-space factors. A dedicated study would be needed to determine the equilibrium configuration.

C.3.3.4 Non-Abelian Gauge Groups - Light Fermions:

Perturbative Effects :

The presence of light fermions allow for new helicity-violating processes similar to the Abelian case. As discussed there, Compton-type scattering will give a contribution to Eq. C.37 of order $\alpha^2 m_f^2 T \alpha \dot{\phi}/f$, while two-to-three scattering processes will give rise to terms that scale like Eq. C.37 up to Casimirs.

As before, these results can be understood in any basis. In the basis with a $\phi G\tilde{G}$ coupling, the fermion scattering allows for gauge-boson helicity violation which extracts energy from the chemical potential portion of the coupling. Additionally, A^3 and A^4 terms will contribute in a way that is not chemical potential like.

In the basis where the axion coupling has been rotated into the light fermion couplings, a similar feature to what occurs in the abelian case abounds. The mass and derivative interactions cancel each other out for almost all interactions except for processes where there is a loop of fermions that modify either the propagator, the three-point function, or the four-point function of the non-abelian gauge boson.

Regardless of basis, a chemical potential for the gauge boson helicity is formed and will drastically change the results. As in the case of no fermions, a dedicated analysis will be required.

Non-Perturbative Effects :

Non-perturbative frictional effects including fermions were first studied in Ref. [141]. Because of the anomaly, sphalerons are tied together with fermion helicity violation. This effect is

similar to how helicity changes of the photon and fermion are connected in Compton scattering. As such, it is not at all clear what the equilibrium configuration is, as chemical potentials for both gauge boson and fermion helicity will need to be considered. Non-abelian warm inflation with light fermions has also been studied in Ref. [99–102], where they included the chemical potential for fermions but not the gauge bosons and neglected the effect of the chemical potential on the fluctuation dissipation theorem when calculating the power spectrum. A dedicated calculation would be needed to elucidate what actually occurs in this situation.

C.4 A warm ϕ^4 example

In the previous sections, we have discussed how friction and fluctuations are drastically changed by the presence of chemical potentials. In this section, we demonstrate that warm inflation can still occur for a simple inflaton potential.

We first give a brief review of warm inflation before summarizing the results of a power spectrum calculation. Next, we discuss the thermal potential, as our example model was chosen for calculational simplicity as opposed to suppressing the thermal potential. Finally, we present a data point demonstrating that warm ϕ^4 inflation is viable.

C.4.1 The Background Solution

The basic idea of warm inflation is to modify the usual slow-roll inflationary dynamics by coupling the inflaton to a radiation bath that's in thermal equilibrium. The background

equations of motion are

$$\ddot{\phi} + (3H + \gamma)\dot{\phi} + V'(\phi) = 0, \tag{C.41}$$

$$H^2 = \frac{1}{3M_{\rm pl}^2}\left(V(\phi) + \frac{1}{2}\dot{\phi}^2 + \rho_r\right), \tag{C.42}$$

$$\dot{\rho}_r + 4H\rho_r = \gamma\dot{\phi}^2, \tag{C.43}$$

where the friction term, γ, is the method by which energy is constantly pumped from the inflaton into the radiation sector. The radiation energy density is $\rho_r = g_*\pi^2 T^4/30$ with g_* being the number of degrees of freedom. Typically, one defines the ratio of the friction term to the Hubble parameter

$$Q \equiv \frac{\gamma}{3H}. \tag{C.44}$$

The usual slow-roll parameters are modified to give

$$\epsilon = \frac{1}{1+Q}\frac{M_{pl}^2}{2}\left(\frac{V'}{V}\right)^2, \qquad \eta = \frac{1}{1+Q}\frac{M_{pl}^2 V''}{V}. \tag{C.45}$$

If $\epsilon, |\eta| \ll 1$, then the approximate slow-roll background solutions are

$$\dot{\phi} \approx -\frac{V'(\phi)}{3H+\gamma}, \tag{C.46}$$

$$H^2 \approx \frac{V}{3M_{\rm pl}^2}, \tag{C.47}$$

$$\rho_r \approx \frac{\gamma}{4H}\dot{\phi}^2. \tag{C.48}$$

The third equation is the condition that the energy extracted from the inflaton per unit Hubble time is sufficient to keep the radiation sector in thermal equilibrium.

In the rest of this section, we show that within our simple model it is possible to satisfy the slow-roll approximation. Upon expanding perturbatively around the background, the contributions of classical thermal fluctuations to the primordial power spectrum can overwhelm the quantum fluctuations that are usually held responsible for the inhomogeneities of our universe.

C.4.2 Power Spectrum

In this section we summarise how the thermal fluctuations turn into the super-horizon inflaton fluctuations we observe today. A derivation of the results of this section can be found in App. C.6.5. We should point out that the power spectrum of warm inflation has been calculated many times before, see Refs. [96,97,103–105], but many of the papers disagree with one another. Our results agree with Ref. [104].

During warm inflation, the finite temperature bath has classical thermal fluctuations that eventually imprint themselves on the inflaton. We expand the field about the slow-roll background

$$\phi(\vec{x},t) = \phi(t) + \delta\phi(\vec{x},t). \tag{C.49}$$

The primordial power spectrum is given by the two-point function of the gauge-invariant scalar curvature perturbation

$$\zeta \equiv -H\frac{\delta\phi}{\dot{\phi}}, \tag{C.50}$$

$$\langle\zeta_{\vec{k}}\zeta_{\vec{k}'}\rangle \equiv (2\pi)^3\delta^3(\vec{k}+\vec{k}')P(k). \tag{C.51}$$

The equations for the perturbations during warm inflation are derived and numerically solved

in App. C.6.5. Roughly speaking, we solve the equation of motion of ϕ and the conservation of the stress energy tensor while properly taking into account the thermal fluctuations. The result using Eq. C.26 and Eq. C.125 is

$$P(k) = \frac{1}{k^3}\frac{H^2}{\dot{\phi}^2}\frac{2c_1 T^4}{f^2}(16\alpha_1^2 + 25\alpha_2^2)F_2(Q), \tag{C.52}$$

where the function $F_2(Q)$ is shown in Fig. C.3 and a numerical approximation to it is given in Eq. C.127.

C.4.3 Thermal Potential

The coupling of the inflaton to the thermal bath induces both a vacuum and finite-temperature potential. The vacuum piece can typically be canceled by symmetries such as supersymmetry, while the finite-temperature piece is typically irreducible. As such, we take the standard warm inflation approach of assuming the vacuum piece has been canceled and only consider the finite-temperature contribution.

To calculate the thermal potential, we expand the field

$$\phi = \varphi_{\rm cl} + \delta\phi, \tag{C.53}$$

in Eq. C.3. The Lagrangian with this substitution is given by

$$\mathcal{L} = |\partial_\mu \chi|^2 - m_\chi^2|\chi^2| - \frac{a(\varphi_{\rm cl})}{4!}\chi^4 + \cdots \tag{C.54}$$

$$a(\varphi_{\rm cl}) \equiv \alpha_1 e^{4i\varphi_{\rm cl}/f} + \alpha_2 e^{5i\varphi_{\rm cl}/f}. \tag{C.55}$$

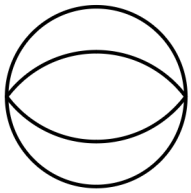
$$-\int d^4x V(\varphi_{\rm cl}) =$$ 

Figure C.2: The three-loop vacuum bubble which gives the leading-order contribution to the effective potential.

We have kept only the terms relevant for the thermal potential calculation. At leading order, we won't need any $\delta\phi$ propagators.

The potential $V(\varphi_{\rm cl})$ can be calculated from the effective action

$$\Gamma[\varphi_{\rm cl}, \tilde{\varphi}_{\rm cl}] = -\int d^4x \left(V(\varphi_{\rm cl}) - \tilde{V}(\tilde{\varphi}_{\rm cl}) + \cdots \right), \tag{C.56}$$

which is the generator of one-particle irreducible diagrams. For completeness, we have included the thermofield double sector, reviewed in App. C.6.1. The potential is given by the vacuum bubble diagrams, and at leading order, one can drop the tilde sector entirely. The leading-order vacuum bubble that gives a potential for ϕ is the 3-loop diagram shown in Fig. C.2 :

$$V(\varphi_{\rm cl}) = \frac{|a|^2}{4!} \int d^4y [D_T(x-y)]^4 + h.c, \tag{C.57}$$

where $D_T(x-y)$ is the thermal Feynman propagator, the Fourier transform of which is given in Eq. C.88. In the limit $m, \mu_{\rm dist} \to 0$, this 3-loop diagram has been calculated with dimensional regularization in the imaginary time formalism [92, 93], and gives

$$V_T(\varphi_{\rm cl}) = \frac{1.72}{4!} \frac{\alpha_1 \alpha_2}{16\pi^2} T^4 \cos(\varphi_{\rm cl}/f), \tag{C.58}$$

where we have chosen to renormalise the couplings at the scale $4\pi T$ in the $\overline{\text{MS}}$ scheme.

The Effect of the Thermal Potential:

Finite temperature potentials can have a variety of effects. For the more traditional scalar couplings these finite temperature potentials make slow roll more difficult. For example, the coupling $y\phi\overline{\psi}\psi$ would generate a temperature-dependent contribution to the inflaton potential $\delta V \sim y^2\phi^2T^2$ at high temperatures. This mass can break the η slow-roll condition $V'' \ll H^2(1+Q)$, making inflation more difficult to achieve.

For periodic scalars, such as the one we are considering, it is a bit more complicated. To understand this, we first note that the distance the inflaton travels per Hubble patch is typically much larger than f, $\dot{\phi}/fH \gg 1$. This means that one needs to consider the entire potential shown in Eq. C.58, as opposed to a perturbative expansion of it.

The effect of the periodic finite-temperature potential is to produce small oscillations on top of the slow-roll potential. In other words, instead of changing the slope or curvature of the potential and thus affecting the slow-roll parameters, the inflaton should instead be imagined as rolling down a slope with wiggles on top of it. Much like a car rolling down a bumpy road, as long as the kinetic energy of the inflaton is such that it can easily roll over the bumps, the behavior of the background solution will be unaffected [94]. Namely, as long as the condition

$$\dot{\phi}^2 \gg \frac{1.72}{4!}\frac{\alpha_1\alpha_2}{16\pi^2}T^4 \tag{C.59}$$

is satisfied, slow roll will proceed as usual. We have checked this numerically and found that as long as the inequality in Eq. C.59 is satisfied by about ~ 10, that the usual slow-roll solution is

an attractive solution stable under perturbations. More colloquially, even though at any given point $|V_T''| \gtrsim H^2(1+Q)$, the rapidly oscillating nature of V_T'' means that the effects of V_T'' average to zero.

Despite averaging to zero for the homogeneous piece, there exist instabilities at finite momentum [95] when rolling over a bumpy potential. Requiring these instabilities do not lead to a large back-reaction on the slow-roll background provides a further constraint which seems to limit us to the weak regime of warm inflation

$$\left(\frac{1.72}{4!}\frac{\alpha_1\alpha_2}{16\pi^2}\frac{T^4}{f^2}\right)^2\left(\frac{\dot{\phi}}{f}\right)^{-3} \lesssim H. \tag{C.60}$$

A derivation of this condition is given in appendix App. C.6.3

C.4.4 The Data Point

In this section we present a data point of ϕ^4 warm inflation where the thermal bath provides the source of fluctuations. As our model was selected for calculational simplicity, we will not map out all of parameter space and instead simply provide a single data point to show that warm inflation can be realized in this model satisfying all current constraints.

Our example is fairly simple. It is our previous model shown in Eq. C.3 combined with the inflaton potential

$$V_{\rm inf} = \frac{\lambda}{4}\phi^4. \tag{C.61}$$

In order to match the observed cosmological data, one must match the scalar tilt, the scalar

power spectrum, and the bound on the scalar to tensor ratio r.

Since perturbations are due to the thermal bath, matching the observed scalar power requires (see Eq. C.26, App. C.6.5, and Ref. [17]) :

$$\Delta_R = 2.09 \times 10^{-9} = \frac{c_1 T^4}{\pi^2 f^2} \left(16\alpha_1^2 + 25\alpha_2^2\right) \frac{H^2}{\dot{\phi}^2} F_2(Q). \tag{C.62}$$

In addition to this, we must match the spectral tilt [17]

$$\begin{aligned} n_s - 1 &= -0.0335 = \frac{d \log \Delta_R}{dN} \\ &= 4\frac{2-Q}{1+7Q}\epsilon - 6\frac{1+Q}{1+7Q}\eta + \frac{d\log F_2}{d\log Q}\frac{1+Q}{1+7Q}(10\epsilon - 6\eta). \end{aligned} \tag{C.63}$$

A derivation of this is given in App. C.6.5.

Finally, there are a few other constraints we must satisfy.

- The validity of Eq. C.62 requires that the thermal fluctuations from warm inflation dominate over the quantum fluctuations:

$$\frac{H^4}{\dot{\phi}^2} \ll \frac{4c_1 T^4}{f^2} \left(16\alpha_1^2 + 25\alpha_2^2\right) \frac{H^2}{\dot{\phi}^2} F_2(Q). \tag{C.64}$$

- We consistently made the approximation

$$\dot{\phi}/f \ll T, \tag{C.65}$$

which must be satisfied by our data point. Satisfying this inequality pushes us to low scale inflation. Note that in addition to this, we must make sure $\mu_{\rm dist} \sim O(\dot{\phi}/f) < m < T$.

- Ensuring the thermal backreaction is subdominant to the slow-roll potential leads to the inequality in Eq. C.59, and avoiding parametric resonance for the non-zero momentum modes gives the constraint of Eq. C.60.

- Validity of the effective field theory requires the non-zero temperature not excite mdoes that are out of EFT control

$$4\pi f > 3T. \tag{C.66}$$

Finally, for perturbation theory in α_i to be valid we require

$$\alpha_i < 16\pi^2. \tag{C.67}$$

All the constraints listed above can be satisfied with the following choices:

$$\frac{\alpha_1}{16\pi^2} = 0.443, \qquad \frac{\alpha_2}{16\pi^2} = 0.450, \qquad \lambda = 3.98 \times 10^{-18}, \quad f = 9.56 \times 10^{-4} M_{\rm pl}. \tag{C.68}$$

We also picked $m_\chi/T = 0.55$. These parameters realize warm inflation with

$$\phi = 17.07 M_{\rm pl}, \quad Q = 0.027, \quad H = 1.68 \times 10^{-7} M_{\rm pl}, \quad T = 8.21 \times 10^{-5} M_{\rm pl}. \tag{C.69}$$

Finally, the prediction for the tensor perturbations in warm inflation is unmodified from cold inflation, so the tensor to scalar ratio is suppressed

$$r = \left(\frac{\dot\phi}{M_{\rm pl}}\right)^2 \frac{2f^2}{c_1(16\alpha_1^2 + 25\alpha_2^2)T^4 F_2(Q)} = 1.36 \times 10^{-6}, \tag{C.70}$$

which is well below current bounds.

We should also mention that this choice of parameters can only support around 40 e-folds of inflation $N = \frac{\sqrt{1+Q}}{M_{\rm pl}} \int_{\phi_{\rm end}}^{\phi} \frac{d\phi}{\sqrt{2\epsilon}} \approx 37$. The number of e-folds required to solve the horizon problem depends both on the Hubble scale during inflation, and the cosmic history of the universe. Because the Hubble scale for our data point is many orders of magnitude lower than what is typically considered, a lower number of e-folds is required.

C.5 Conclusions and Further Work

In this article we have shown that warm inflation utilizing pseudo-scalar inflaton interactions develops a chemical potential. In the context of a simple model, we then explained how to calculate both the friction and thermal fluctuations. We saw that both of these quantities can be drastically altered by the presence of a chemical potential.

We then discussed models in which the inflaton has generic pseudo-scalar couplings. We demonstrated that these models develop chemical potentials and estimated the friction terms. There is still a lot that can be done with these interactions. For some interactions, such as the $\phi G\tilde{G}$ coupling, additional work is required to properly calculate the thermal friction and fluctuations. For other interactions, such as the $\phi F\tilde{F}$ coupling with light fermions, the equilibrium configuration has energy injection whose form is not as simple as $\gamma\dot{\phi}$. It would be interesting if a warm inflation model could be made where the thermal bath is the Standard Model. It should be noted that while warm inflation with the Standard Model has been studied in Ref. [102], it is unclear whether their predictions would survive taking into account the gauge boson chemical potential.

Finally, it would be interesting to explore signals of the thermal classical nature of the fluctuations that led to structure formation. Interesting avenues might be the non-Gaussian predictions of warm inflation [97, 104, 142, 143] or its gravitational wave predictions [144, 145]. Additionally, it would be interesting to explore if warm inflation can successfully address some of the typical criticisms of cold inflation without introducing problems of its own.

Acknowledgments

We would like to thank Mehrdad Mirbabayi, and Kim Berghaus for helpful discussions. The authors were supported by NSF grant PHY-2210361 and the Maryland Center for Fundamental Physics.

C.6 Appendix

C.6.1 A Lightning Fast Review of the Thermal Field Double

In this Appendix we give a brief review of how to use the thermal field double to calculate thermal expectation values. For a more comprehensive review consult [146,147]. The expectation value of an operator in a thermal state is given by

$$\langle A\rangle_\beta = \mathrm{Tr}\,(A\rho) = \frac{1}{Z}\sum_n e^{-\beta E_n}\langle n|A|n\rangle, \tag{C.71}$$

where we have expanded the Hilbert space, $\mathcal{H}$, in the energy eigenbasis, and Z is the partition function. The main difficulty in generalizing the techniques of perturbative Quantum Field Theory to finite temperature is that the thermal state is not a pure state; its density matrix is

diagonal. The thermal field double is a trick that allows one to calculate thermal expectation values like any other quantity in QFT.

The essence of the thermal field double is to have two copies of the theory in question. We will suggestively label them the system, with states n, and the environment, with states $\tilde{n}$. As long as one only asks questions about the system, the environment is traced over giving the thermal mixed state. More precisely we define the "thermal vacuum"

$$|0\rangle_\beta = \frac{1}{\sqrt{Z}} \sum_n e^{-\beta E_n/2} |n, \tilde{n}\rangle. \tag{C.72}$$

The expectation value, in this state, of any operator that acts only on the system obeys

$$\begin{aligned} {}_\beta\langle 0|A|0\rangle_\beta &= \frac{1}{Z} \sum_{n,\tilde{n},m,\tilde{m}} e^{-\beta(E_n+E_m)/2} \langle m|A|n\rangle \langle \tilde{m}|\tilde{n}\rangle, & \text{(C.73)} \\ &= \frac{1}{Z} \sum_n e^{-\beta E_n} \langle n|A|n\rangle = \langle A\rangle_\beta, & \text{(C.74)} \end{aligned}$$

so we have succeeded in writing thermal expectation values as the expectation value in a pure state.

While not necessary, it is convenient to choose the thermal vacuum, $|0\rangle_\beta$, such that it doesn't evolve in time and has zero energy. This amounts to saying that the total Hamiltonian $\hat{H}$ and Lagrangian $\hat{\mathcal{L}}$ obey [148]

$$\hat{H} = H - \tilde{H} \qquad \hat{\mathcal{L}} = \mathcal{L} - \tilde{\mathcal{L}}. \tag{C.75}$$

If the Lagrangian of the bath is $\mathcal{L}(\phi, \phi^\dagger)$, then one can obtain the Lagrangian of the environment

using $\tilde{\mathcal{L}} = \mathcal{L}(\tilde{\phi}^\dagger, \tilde{\phi})^\dagger$.

We can apply this formalism to the simple harmonic oscillator. There are creation and annihilation operators for the system and the environment. The thermal vacuum can be obtained by performing a unitary transformation on the ground state of the combined system

$$|0\rangle_\beta = U(\theta)|0,\tilde{0}\rangle, \tag{C.76}$$

$$= e^{-\theta_\beta(\tilde{a}a - \tilde{a}^\dagger a^\dagger)}|0,\tilde{0}\rangle. \tag{C.77}$$

The angle θ_β is given by

$$\cosh\theta_\beta = \frac{1}{\sqrt{1-e^{-\beta\omega}}}, \tag{C.78}$$

$$\sinh\theta_\beta = \frac{e^{-\beta\omega/2}}{\sqrt{1-e^{-\beta\omega}}}, \tag{C.79}$$

where ω is the oscillator frequency.

Since the thermal ground state is a unitary transformation acting on the vacuum, the new creation and annihilation operators are Bogoliubov transformations of the familiar creation and annihilation operators with

$$a_\beta = \cosh\theta_\beta a - \sinh\theta_\beta \tilde{a}^\dagger, \tag{C.80}$$

$$a_\beta^\dagger = \cosh\theta_\beta a^\dagger - \sinh\theta_\beta \tilde{a}, \tag{C.81}$$

$$\tilde{a}_\beta = \cosh\theta_\beta \tilde{a} - \sinh\theta_\beta a^\dagger, \tag{C.82}$$

$$\tilde{a}_\beta^\dagger = \cosh\theta_\beta \tilde{a}^\dagger - \sinh\theta_\beta a. \tag{C.83}$$

The thermal creation and annihilation operators obey $a_\beta|0\rangle_\beta = \tilde{a}_\beta|0\rangle_\beta = 0$. These expressions have an intuitive understanding. Since a_β annihilates the vacuum, destroying a particle in the system is equivalent to introducing a particle to the environment. In other words, any non-conservation of energy, momentum, and/or particles is just the system exchanging these quantities with the environment. The extra factors of $\cosh\theta_\beta$ and $\sinh\theta_\beta$ are merely the normalization factors associated with states that contain several identical particles.

One can calculate the expectation value of the particle number operator in the thermal state

$$ {}_\beta\langle 0|a^\dagger a|0\rangle_\beta \;=\; \sinh^2\theta_\beta\, {}_\beta\langle 0|a_\beta^\dagger a_\beta|0\rangle_\beta = \frac{1}{e^{\beta\omega}-1}, \tag{C.84} $$

which is the usual Bose-Einstein distribution function. It's clear that adding a chemical potential will simply amount to replacing $\omega \to \omega - \mu_{\rm dist}$.

The generalisation to field theory is as simple as at zero temperature. We know how to decompose any field in terms of the typical creation and annihilation operators, a and $a^\dagger$. We can use Eq. C.80 and its inverse to write these fields in terms of the thermal creation and annihilation operators acting on the thermal bath, a_β and $a_\beta^\dagger$. For our application we should include a chemical potential, $\mu_{\rm dist}$, for the particle and, $-\mu_{\rm dist}$, for the antiparticle. More

explicitly,

$$\chi(x) = \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}}\left(a_p e^{-ipx} + b_p^\dagger e^{ipx}\right), \tag{C.85}$$

$$= \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}}\left((u_\beta(\mu_{\rm dist})a_{p,\beta} + v_\beta(\mu_{\rm dist})\tilde{a}^\dagger_{p,\beta})e^{-ipx} + (u_\beta(-\mu_{\rm dist})b^\dagger_{p,\beta} + v_\beta(-\mu_{\rm dist})\tilde{b}_{p,\beta})e^{ipx}\right)$$

$$u_\beta(\mu_{\rm dist}) = \frac{1}{\sqrt{1-e^{-\beta(\omega_p-\mu_{\rm dist})}}} \qquad v_\beta(\mu_{\rm dist}) = \frac{e^{-\beta(\omega_p-\mu_{\rm dist})/2}}{\sqrt{1-e^{-\beta(\omega_p-\mu_{\rm dist})}}}, \tag{C.86}$$

where we have left some of the p dependence implicit. At this point, the calculation proceeds as any other standard Quantum Field Theory calculation.

For applications in the main text we will require various expressions for the thermal field propagator. Firstly, the usual Feynman propagator will be used to calculate the thermal back-reaction on the slow-roll potential

$$iD_F(p,T) \equiv \int d^4x e^{ip\cdot(x-y)}\,{}_\beta\langle 0|T\{\chi(x)\chi^\dagger(y)\}|0\rangle_\beta \tag{C.87}$$

$$= \frac{i}{p^2-m^2+i\epsilon} + (2\pi)\delta(p^2-m^2)\left(\theta(p^0)\bar{f}_p + \theta(-p^0)f_p\right), \tag{C.88}$$

where the distribution functions are defined as in Eqns. C.10 and C.11. It's interesting to note that the effect of the non-zero temperature is to simply modify the propagator by the addition of a single term that depends on the distribution function.

We use the in-in formalism to calculate the friction term in the equation of motion. This requires the use of propagators without time-ordering. Because unitarity is simpler in the time

domain [149], it is convenient not to Fourier transform in time giving

$$
\begin{aligned}
{}_\beta\langle 0|\chi^\dagger(x)\chi(y)|0\rangle_\beta &= \int \frac{d^3p}{(2\pi)^3 2\omega_p}\left((1+\bar{f}_p)e^{-ip\cdot(x-y)} + f_p e^{ip\cdot(x-y)}\right), && \text{(C.89)}\\
{}_\beta\langle 0|\chi(x)\chi^\dagger(y)|0\rangle_\beta &= \int \frac{d^3p}{(2\pi)^3 2\omega_p}\left((1+f_p)e^{-ip\cdot(x-y)} + \bar{f}_p e^{ip\cdot(x-y)}\right). && \text{(C.90)}
\end{aligned}
$$

One can easily pick out initial and final states in the Boltzmann equation from this form of the propagator.

C.6.2 A lightning Fast Review of the In-in Formalism

In this Appendix we give a lightning fast review of the in-in formalism. We will be a bit sloppy with the $i\epsilon$ prescription as it will not matter for the calculations we do in this paper. When one first learns Quantum Field Theory, we are taught in-out expectation values defined as

$$
\langle 0|T\left(e^{-i\int_t^\infty H^{\rm int}(t_2)dt_2}\right)O(t)T\left(e^{-i\int_{-\infty}^t H^{\rm int}(t_1)dt_1}\right)|0\rangle. \qquad \text{(C.91)}
$$

The ket is the vacuum in the past, time evolved to when the measurement occurs. This is called the in state. Meanwhile, the bra is the vacuum in the future, time evolved backwards to when the measurement occurs. This is called the out state. While completely un-intuitive, this object is easy to calculate and related to physical processes such as cross sections.

Even further back, when one first learns quantum mechanics, expectations values are

instead evaluated as

$$\begin{aligned}\langle\Psi(t)|O(t)|\Psi(t)\rangle &\sim \langle 0|\left(T\left(e^{-i\int_{-\infty}^{t}H^{\rm int}(t_2)dt_2}\right)\right)^{\dagger}O(t)T\left(e^{-i\int_{-\infty}^{t}H^{\rm int}(t_1)dt_1}\right)|0\rangle \\ &= \langle 0|\overline{T}\left(e^{i\int_{-\infty}^{t}H^{\rm int}(t_2)dt_2}\right)O(t)T\left(e^{-i\int_{-\infty}^{t}H^{\rm int}(t_1)dt_1}\right)|0\rangle, \end{aligned} \tag{C.92}$$

where $\overline{T}$ represents anti-time ordering. Because the bra and ket state are both obtained by time-evolution of a state from the past to the present, this is called an in-in expectation value.

The in state and the out state are related by time evolution of the vacuum from $t=-\infty$ to ∞. As such, in-in and in-out calculations are often just related to each other by a phase. For time dependent vacua and Lagrangians, the situation can become significantly more complicated.

A particularly useful way in which to write Eq. C.92 is [118, 150]

$$\langle O(t)\rangle = \sum_{V=0}^{\infty}(-i)^V\int_{t_0}^{t}dt_v\cdots\int_{t_0}^{t_3}dt_2\int_{t_0}^{t_2}dt_1\langle[[\cdots[\mathcal{O},H_{\rm int}(t_v)]\cdots,H_{\rm int}(t_2)],H_{\rm int}(t_1)]\rangle. \tag{C.93}$$

It turns out that this way of writing expectation values in terms of commutators is particularly conducive to calculations using the thermal field double. Both in-in and in-out calculations give exactly the same result, because in the thermal field double calculation the thermal vacuum was chosen to be time invariant with energy equal to zero. However, in-in calculations are computationally easier. The reason for this is that as shown in Eq. C.93, in-in calculations involve retarded propagators while in-out calculations only ever use Feynman propagators. Because the bath and environment fields commute

$$[\phi,\tilde{\phi}]=0, \tag{C.94}$$

when doing in-in calculations, the environment simply does not show up. This renders calculations much easier to do in the in-in formalism as opposed to the in-out formalism. Of course, the two approaches give the same answer.

C.6.3 Parametric Resonance

In this appendix we give a quick derivation of the condition Eq. C.60. For a more detailed discussion consult [95], where the estimates given here are checked numerically. In the presence of the thermal effective potential Eq. C.58 the equation of motion for the fluctuation is given by

$$\delta\ddot{\phi}_k + \Gamma\delta\dot{\phi}_k + \left(\frac{k^2}{a_0^2}e^{-2Ht} - A\cos(\omega t)\right)\delta\phi_k = 0, \tag{C.95}$$

where we have dropped the contribution to the fluctuations from interactions with the radiation sector to zoom in on the effect of parametric resonance. The various constants are given by

$$\Gamma = 3H + \gamma; \quad A = \frac{1.72}{4!}\frac{\alpha_1\alpha_2}{16\pi^2}\frac{T^4}{f^2}; \quad \omega = \dot{\phi}/f. \tag{C.96}$$

This equation has an instability when

$$\frac{k}{a_0}e^{-Ht_*} \approx \frac{\omega}{2}, \tag{C.97}$$

because the final term will have a component that acts like a forcing term on resonance. We have defined t_* as the time at which this particular k-mode is on resonance.

We assume that the length of time a given k mode is unstable is short relative to Hubble,

$t = t_* + \delta t$, and that the friction term can be neglected during the instability

$$\frac{k^2}{a_0^2} e^{-2Ht} \approx \frac{\omega^2}{4}(1 - 2H\delta t); \qquad \frac{\Gamma}{\omega} \ll \frac{A}{\omega^2}. \tag{C.98}$$

During the instability, the equation of motion for the fluctuations resembles the Matthieu equation, and the unstable mode grows exponentially [95]

$$\delta\phi_k(t) \sim \delta\phi_k(t_*) \exp\left(\frac{A}{2\omega}(t - t_*)\right). \tag{C.99}$$

This equation is true exactly on resonance, and therefore gives an upper bound on the exact growth rate.

To estimate the size of the enhancement we need to know how long each mode is unstable. It turns out that each mode is unstable as long as

$$\frac{\omega^2}{4} - \frac{A}{2} < \frac{\omega^2}{4}(1 - 2H\delta t) < \frac{\omega^2}{4} + \frac{A}{2}. \tag{C.100}$$

We use this condition to estimate the length of time the mode is unstable as

$$\Delta t = \frac{1}{H}\frac{2A}{\omega^2}. \tag{C.101}$$

Note here that $A \ll \omega^2$ is the condition Eq. C.59 that the inflation has enough kinetic energy to roll over the thermal potential, so this confirms our assumption that the instability is short relative to Hubble. If we plug this in we find that during the instability the non-zero k mode

grows by a factor

$$\exp\left(\frac{1}{H}\frac{A^2}{\omega^3}\right) \implies \frac{A^2}{\omega^3} \ll H, \tag{C.102}$$

which is the inequality given in Eq. C.60[3].

C.6.4 The Fluctuation Dissipation Theorem

In this appendix we present the derivation of the fluctuation dissipation theorem and show that our results obey it. While the derivation of the fluctuation dissipation theorem with a chemical potential for a conserved charge is well known, we are interested in the theorem in the presence of a chemical potential for a non-conserved charge. At leading order in the symmetry breaking coupling, there is no difference between the two situations. As this is all we need to show the validity of our results, we do not push further. It would be interesting to know what is the full form of the fluctuation dissipation theorem in the presence of chemical potentials for non-conserved charges.

C.6.4.1 Derivation of the Theorem

Let us first present a quick derivation of the fluctuation-dissipation theorem with a chemical potential. The presentation will follow the style of [98]. The thermal expectation value of an operator is given by

$$\langle \mathcal{O}(x) \rangle \equiv \frac{1}{Z} \mathrm{Tr}\left[\mathcal{O}(x) e^{-\beta(H-\mu_{\mathrm{dist}}Q)}\right], \tag{C.103}$$

[3] A slightly more sophisticated estimate, taking into account that the mode is slightly off resonance during the instability, gives a factor of $\pi/4$ as opposed to the factor of 1, and fits the numerics remarkably well [95].

where Q is the charge operator, which in our model is the χ number operator. For an operator with a definite charge under Q

$$[\mathcal{O}, Q] = q\mathcal{O}. \tag{C.104}$$

The theorem relates the following 3 thermal expectation values

$$\Pi^{>}_{\mathcal{O}}(k) \equiv \int d^4x e^{ik\cdot x}\langle \mathcal{O}(x)\mathcal{O}^\dagger(0)\rangle, \tag{C.105}$$

$$\Pi^{<}_{\mathcal{O}}(k) \equiv \int d^4x e^{ik\cdot x}\langle \mathcal{O}^\dagger(0)\mathcal{O}(x))\rangle, \tag{C.106}$$

$$\rho_{\mathcal{O}}(k) \equiv \int d^4x e^{ik\cdot x}\frac{1}{2}\langle [\mathcal{O}(x), \mathcal{O}^\dagger(0)]\rangle. \tag{C.107}$$

For the purpose of this derivation we must assume $[H, Q] = 0$. Although this isn't true because the couplings break χ number conservation, it is true in the free theory, and can be used at leading order in perturbative calculations. We can insert $1 = \sum_n |n\rangle\langle n|$ twice into the first two expectation values to obtain

$$\Pi^{>}_{\mathcal{O}}(k) = \frac{1}{Z}\sum_{m,n}\int d^3x e^{-i\vec{k}\cdot\vec{x}}(2\pi)\delta(k^0 + E_m - E_n)e^{-\beta E_m}\langle m|e^{\beta\mu_{\text{dist}}Q}\mathcal{O}(\vec{x})|n\rangle\langle n|\mathcal{O}^\dagger(0)|m\rangle \tag{C.108}$$

$$\Pi^{<}_{\mathcal{O}}(k) = \frac{1}{Z}\sum_{m,n}\int d^3x e^{-i\vec{k}\cdot\vec{x}}(2\pi)\delta(k^0 + E_m - E_n)e^{-\beta E_n}\langle m|\mathcal{O}(\vec{x})e^{\beta\mu_{\text{dist}}Q}|n\rangle\langle n|\mathcal{O}^\dagger(0)|m\rangle \tag{C.109}$$

where we used $\mathcal{O}(\vec{x}, t) = e^{iHt}\mathcal{O}(\vec{x})e^{-iHt}$ in the intermediate steps. At this point one can use the commutation relation shown in Eq. C.104 to find $\mathcal{O}(\vec{x})e^{\beta\mu_{\text{dist}}Q} = e^{\beta\mu_{\text{dist}}(Q+q)}\mathcal{O}(\vec{x})$, giving a relationship between the first two expectation values. Combining this with the fact that the third expectation value is half of the difference of the first two we obtain the fluctuation-dissipation

theorem

$$\rho_{\mathcal{O}}(k) = \frac{1}{2}\left(e^{\beta(k^0-\mu_{\rm dist}q)} - 1\right)\Pi^{<}_{\mathcal{O}}(k). \tag{C.110}$$

C.6.4.2 Consistency of our Calculation

We can check if the calculations of Sec. C.2 are consistent with this theorem. Focus on the first term of Eq. C.20

$$\begin{aligned} -\gamma\dot{\phi} &\supset -\frac{4}{f}\frac{\alpha_1^2}{(4!)^2}\int d^3y \int_{-\infty}^{t_x} dt_y e^{4i\frac{\dot{\phi}}{f}(t_x-t_y)}\langle[\chi^4(x),\chi^{\dagger 4}(y)]\rangle & \text{(C.111)}\\ &= \frac{4}{f}\frac{\alpha_1^2}{(4!)^2}\int d^4y e^{-4i\frac{\dot{\phi}}{f}(t_y-t_x)}\frac{1}{2}\langle[\chi^{\dagger 4}(y),\chi^4(x)]\rangle & \text{(C.112)}\\ &= \frac{4}{f}\frac{\alpha_1^2}{(4!)^2}\rho_{\chi^\dagger}(-k_1), & \text{(C.113)} \end{aligned}$$

where $k_1 = (4\dot{\phi}/f, \vec{0})$, and we have extended the integral to infinity by recalling from the main text that after adding the hermitian conjugate the integral over time gave an energy conservation delta function. If we add in all the other contributions to the friction term we find

$$-\gamma\dot{\phi} = -\frac{4}{f}\frac{\alpha_1^2}{(4!)^2}\left(\rho_\chi(k_1) - \rho_{\chi^\dagger}(-k_1)\right) - \frac{5}{f}\frac{\alpha_2^2}{(4!)^2}\left(\rho_\chi(k_2) - \rho_{\chi^\dagger}(-k_2)\right), \tag{C.114}$$

where $k_2 = (5\dot{\phi}/f, \vec{0})$.

We can read off from Eq. C.25 that the fluctuations

$$A = \frac{16\alpha_1^2}{(4!)^2 f^2}\left(\Pi^{<}_{\chi}(k_1) + \Pi^{<}_{\chi^\dagger}(-k_1)\right) + \frac{25\alpha_2^2}{(4!)^2 f^2}\left(\Pi^{<}_{\chi}(k_2) + \Pi^{<}_{\chi^\dagger}(-k_2)\right). \tag{C.115}$$

Now consider the fluctuation dissipation theorem (Eq. C.110) at leading order by noting that

$q = \pm 4$

$$\rho_{\chi}(k) \approx \frac{1}{2}\beta(k^0 - 4\mu_{\rm dist})\Pi^{<}_{\chi}(0), \tag{C.116}$$

$$\rho_{\chi^\dagger}(k) \approx \frac{1}{2}\beta(k^0 + 4\mu_{\rm dist})\Pi^{<}_{\chi^\dagger}(0). \tag{C.117}$$

Note that we have applied the fluctuation dissipation theorem separately to the ± 4 charged operators because these obey Eq. C.104 separately but not as a sum. This is expression is another way to see what we mentioned in the main text that at leading order the chemical potential leaves the fluctuation unaffected while reducing the friction. By setting the argument of Π to zero we see that the expression for A is completely independent of $\mu_{\rm dist}$.

Plugging Eq. C.116 and Eq. C.117 into Eq. C.114, and setting $k_1, k_2 = 0$ in Eq. C.115 gives exactly the relationship between the friction term and the fluctuations that we found in Eq. C.27. Also, notice that if $\mu_{\rm dist}$ is set to zero in Eqs. C.116 and C.117 then we would obtain $A = 2\gamma T$, which shows that $\mu_{\rm dist} = 0$ was an important implicit assumption in previous uses of the fluctuation-dissipation theorem in warm inflation.

C.6.5 From Thermal Fluctuations to Inflaton Fluctuations

In this Appendix, we derive the results used in Sec. C.4.2 of the main text. This work is necessarily partly numerical as the relevant Greens functions cannot be obtained completely analytically. Additionally, as mentioned in the main text, while this aspect of warm inflation has been addressed many times before, see Refs. [96, 97, 103, 104], many of the papers disagree with each other. The disagreement may in part be numerical, however many of the papers disagree

with each other on what is the form of the equation to solve. [4] In this Appendix, we discuss our approach to the problem, which follows Ref. [104].

In what follows, we use the mostly positive metric

$$ds^2 = -dt^2 + a(t)^2 dx^2, \tag{C.118}$$

and use dots to signify time derivatives and $'$ to signify derivatives with respect to the conformal time $\eta = -1/(aH)$. We will work in spatially flat gauge and in the limit $M_p \to \infty$. Additionally, we will take the temperature dependence of the frictional coefficient to be $\gamma \propto T^c$ and the equation of state of the thermal bath to be w.

C.6.5.1 The Differential Equations

We first discuss the differential equation that turns thermal fluctuations into inflaton fluctuations. In the limit where we are considering k modes corresponding to wavelengths much longer than the correlation length of the thermal bath, we can treat the fluid as an ideal fluid. As a result there are three scalar fluctuations that we need to consider, the inflaton fluctuations, the temperature fluctuations, and fluctuations in the local velocity of the bath. We express these three fluctuations as

$$\phi = \phi_0 + \dot{\phi}_0 \Phi \qquad T^0_{\text{bath},0} = -\rho_0(1+\mathcal{E}) \qquad \partial_i T^i_{\text{bath},0} = -\frac{4\rho_0}{3a^2}\nabla^2\Psi. \tag{C.119}$$

[4] All references agree that the equations to use are the conservation of the stress energy tensor, Eq. C.121, and the equation of motion for the inflaton, Eq. C.120. We disagree with Ref. [96, 97, 103, 105] because in Eq. C.122, we include the fluctuation term ξ coming from using the inflaton equation of motion, Eq. C.120, whereas they drop this term. Additionally in Eq.40 of Ref. [96], they copied Eq.A12 of Ref. [142] incorrectly dropping the additional term $2\Gamma\dot{\phi}\delta\dot{\phi}$, leading to additional differences with the rest of the references.

To solve for these three variables, we have the three equations coming from the equation of motion of ϕ and the conservation of the stress energy tensor. The equation of motion for ϕ is

$$D^2\phi - V_{,\phi}(\phi) = \gamma u^\mu \partial_\mu \phi + \xi, \tag{C.120}$$

where u^μ specifies the rest frame of the thermal bath, ξ is the noise term, and D is the covariant derivative. Meanwhile conservation of the stress energy tensor gives

$$D_\mu T^\mu_\nu = D_\mu T^\mu_{\text{bath},\nu} + D_\mu T^\mu_{\phi,\nu} = 0 \tag{C.121}$$

$$D_\mu T^\mu_{\text{bath},\nu} = -D_\mu T^\mu_{\phi,\nu} = -\partial_\nu \phi \left(\gamma u^\mu \partial_\mu \phi + \xi\right) \tag{C.122}$$

where in the last line we have used the equation of motion, Eq. C.120. Removing the homogeneous pieces of these equations, we find the equations that govern the evolution of the perturbations to be

$$\begin{aligned}
&\Phi'' - \left(\frac{2}{\eta} + \frac{\gamma}{H\eta}\right)\Phi' + k^2\Phi = -\frac{c\gamma\mathcal{E}}{4H^2\eta^2} + \frac{\xi}{H^2\eta^2\dot{\phi}_0}, \\
&\mathcal{E}' - \frac{3(1+w)}{\eta}(1-\frac{c}{4})\mathcal{E} + \frac{4k^2H\eta\Psi}{3} - 6H(1+w)\Phi' = \frac{3(1+w)}{\gamma\eta\dot{\phi}_0}\xi, \\
&\Psi' - \frac{3}{\eta}\Psi - \frac{3w}{4H\eta}\mathcal{E} - \frac{9(1+w)}{4\eta}\Phi = 0.
\end{aligned} \tag{C.123}$$

C.6.5.2 Numerical Results

To proceed further, we need to solve Eqs. C.123 subject to the thermal fluctuations

$$\langle \xi_k(\eta)\xi_{k'}(\eta')\rangle = AH^4\eta^4\delta(\eta-\eta')(2\pi)^3\delta^3(k+k'). \tag{C.124}$$

The observed density fluctuations are written in terms of the gauge invariant perturbation ζ. After gauge fixing, $\zeta_k = -\frac{H}{\dot{\phi}_0}\delta\phi_k$, the observed density fluctuations are

$$\langle \zeta_k \zeta_{k'} \rangle = \frac{AH^2}{\dot{\phi}_0^2} F_2(Q) \frac{(2\pi)^3 \delta^3(k+k')}{k^3}, \tag{C.125}$$

where $Q \equiv \gamma/3H$. All that remains for us to do is determine numerically the function

$$F_2(Q) = k^3 \int_{-\infty}^{0} d\eta\, \eta^4 G_{\Phi_k}(0,\eta)^2, \tag{C.126}$$

where $G_{\Phi_k}(\eta,\eta')$ is the Greens function for the inflaton fluctuations Φ obtained by solving Eqs. C.123 with the replacement $\xi \to \dot{\phi}_0 \delta(\eta - \eta')$.

Since our approach can generate $\gamma \sim T^c$ with $c = 3, 1, -1$, we will be considering these three options when studying the two point function numerically. In all cases, we will restrict ourselves to a radiation bath obeying $w = 1/3$. We find a good approximation for our numerically obtained F_2 are

$$F_2(Q) \approx \begin{cases} 5 + 7.9 \tanh\left(\frac{1}{30Q}\right) + 1.5Q^2 + 7.8\times 10^{-3} Q^{4.2} + 1.0\times 10^{-6} Q^{6.6}, & \gamma \sim T^3 \\ \frac{1}{1.1+3.2Q} + 0.25\sqrt{Q} + 3.6\times 10^{-3} Q^2, & \gamma \sim T \\ \frac{1}{3.4+2.7Q+0.23Q^2}, & \gamma \sim T^{-1} \end{cases} \tag{C.127}$$

We compare these analytic approximations to the numerical solutions of F_2 in Fig. C.3.

We now discuss some of the physics that goes into these solutions. In particular, we will focus on the behavior of the modes present at large $Q = \gamma/3H$ as can be seen in Fig. C.3. The delta function source in Eq. C.124 produces fluctuations at all frequencies and the goal is

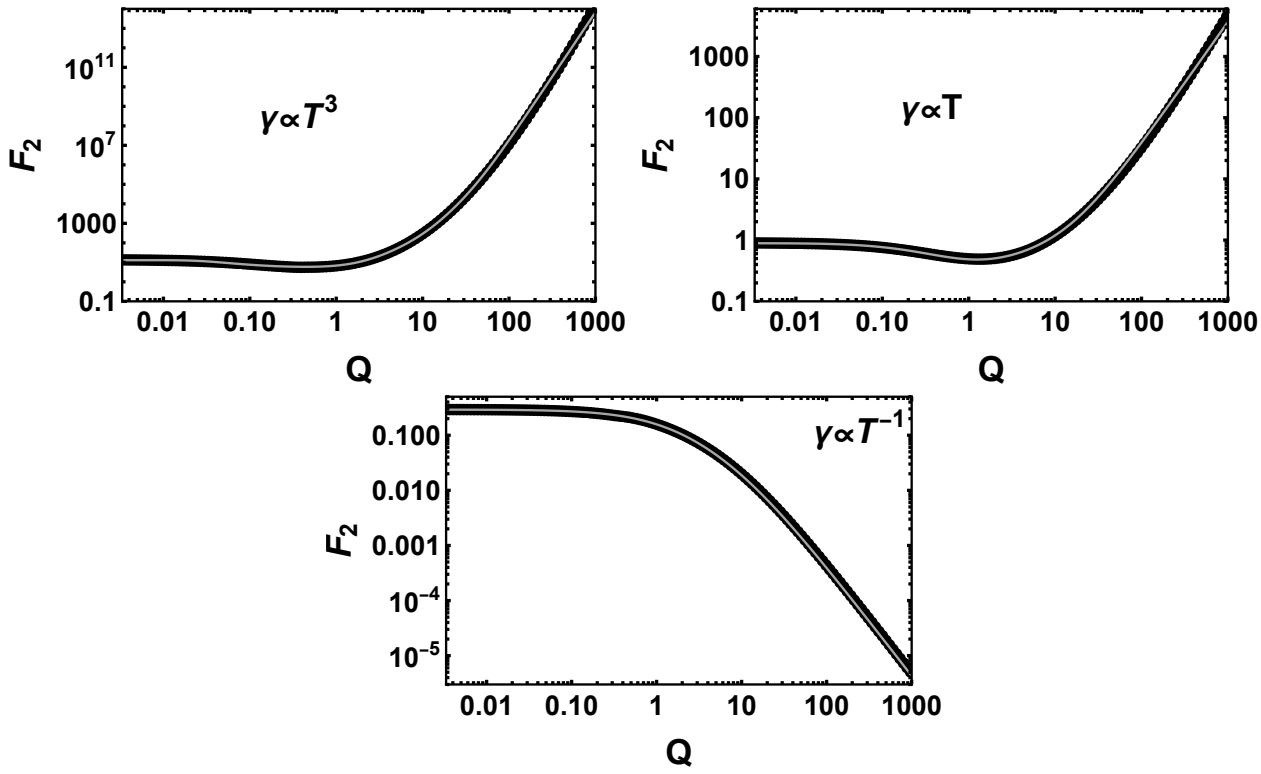


Figure C.3: A comparison of the numerical value of F_2 versus the numerical fits shown in Eq. C.127. The fits are shown in the solid lines while the dots represent the numerical solution.

to identify and understand the dynamics of the modes that eventually dominate the inflaton fluctuations we observe today. We will discuss things chronologically as the physical momentum goes from large to small.

$c > 0$: When the physical momentum is larger than $\sqrt{\gamma H}$, inflaton perturbations are underdamped oscillators. In this region of parameter space, the WKB approximation shows that these modes dilute away. As such inflaton modes that begin their life as thermal fluctuations with large momenta like this are necessarily subdominant.

Next we consider when the physical momentum is smaller than $\sqrt{\gamma H}$ but still larger than Hubble. In this region, there is an exponential instability. The critical consideration behind this

instability is how does a $\delta\phi$ fluctuation deposit energy and momentum into the fluid. In the momentum regime we are considering, momentum deposition is more important than energy deposition, and most of energy goes into pressure gradients. Because pressure points outwards, a large pressure gradient causes the temperature to drop. As a result, an upward fluctuation in $\delta\phi$ deposits a lot of momentum and causes temperature to decrease, which causes the friction to decrease and thus speeding up the growth of $\delta\phi$ leading to a runaway.

When the physical momentum is smaller than Hubble, Hubble friction is more important than anything else in the story. As is familiar, Hubble friction brings perturbations to a grinding halt. As a result, the only superhorizon mode that survives until late times is the one that becomes a constant.

We can now trace the origin of the perturbations that dominate the end result. These modes were originally produced by the delta function thermal fluctuation when their momentum was of order $\sqrt{\gamma H}$. These modes grow in size until they are redshifted out of the instability band and become superhorizon. At that point they freeze out and become constant.

$c < 0$: The $c < 0$ story is similar to the $c > 0$ except for in the instability region, $H < k_{\text{physical}} < \sqrt{\gamma H}$. As before, momentum deposition is the dominant effect and so the logic is similar to before. An upward fluctuation in $\delta\phi$ deposits a lot of momentum and causes temperature to decrease, which causes the friction to increase (as opposed to $c > 0$ which caused a decrease) and slowing down the growth of $\delta\phi$ leading to a feedback mechanism that stabilizes any perturbation. Thus there is no exponential growth and only friction.

Because when $c < 0$, there is no enhancement of modes, the relevant modes are dominantly produced when $k_{\text{physical}} \sim H$ so as to minimize frictional effects diminishing the amplitude before

freeze-out. In this almost slow roll limit, the γ dependence of the Greens function is clear since the only place that γ appears is as ξ/γ. So that the Greens function $\sim 1/\gamma$ and $F_2 \sim 1/\gamma^2$ as seen in Fig. C.3.

C.6.5.3 Derivation of the Spectral Tilt

Fitting to experimental data requires calculating the spectral tilt

$$n_s - 1 \equiv \frac{d\log \Delta_R}{dN}, \tag{C.128}$$

where $dN = Hdt$. From Eq. C.125 we see that the spectral tilt is given by

$$n_s - 1 = 2\frac{d\log H}{dN} - 2\frac{d\log \dot{\phi}}{dN} + \frac{d\log A}{dN} + \frac{d\log F_2}{d\log Q}\frac{d\log Q}{dN}. \tag{C.129}$$

We must write the derivative here in terms of the slow-roll parameters. By directly taking the derivatives of the LHS in Eqs. C.46, and C.47 one finds

$$\frac{d\log H}{dN} = -\epsilon \tag{C.130}$$

$$\frac{d\log \dot{\phi}}{dN} = \epsilon - \eta - \frac{Q}{1+Q}\frac{d\log Q}{dN}. \tag{C.131}$$

To write the final term in terms of ϵ and η we use $Q \sim T^c/H$ to write

$$\frac{d\log Q}{dN} = \epsilon + c\frac{d\log T}{dN}. \tag{C.132}$$

Finally one can recall that $\rho_r \sim T^4$, and use Eq. C.48 to find

$$\left(4 + c\frac{Q-1}{Q+1}\right)\frac{d\log T}{dN} = \frac{3+Q}{1+Q}\epsilon - 2\eta. \tag{C.133}$$

This is easily rearranged to give $d\log T/dN$ in terms of ϵ, and η. If we specialise to $c = 3$, and plug into Eq. C.129, then we arrive at Eq. C.63.

C.6.6 Linear Response Theory and the Adiabatic Approximation

Most other approaches to the perturbations and fluctuations of warm inflation utilize linear response theory [136, 137]. In this section we briefly review linear response theory and the adiabatic approximation emphasizing how they are simply related to how we went about calculating expectation values.

Let us take an interaction between the bath and the inflaton to be $V_{\rm int}(\phi,\chi) = \epsilon f(\phi)g(\chi)$. In the approach we took in the main text, we would write the equation of motion as

$$\Box\phi + 3H\dot{\phi} + V'(\phi) = -\epsilon_{\rm in}\langle f'(\phi)g(\chi)\rangle_{{\rm in},\beta}, \tag{C.134}$$

and subsequently utilize perturbation theory to calculate the expectation value of the in-in correlator. Linear response theory is when one calculates the expectation value in Eq. C.134 to first order in perturbation theory. To first order in perturbation theory with $V_{\rm int}$, the above expectation value is

$$-_{\rm in}\langle f'(\phi)g(\chi)\rangle_{{\rm in},\beta} = i\epsilon\int d^4y\theta(t_x - t_y)\langle[f'(\phi_x)g(\chi_x), f(\phi_y)g(\chi_y)]\rangle_\beta + \mathcal{O}(\epsilon^2), \tag{C.135}$$

where the subscripts indicate what fields are functions of x versus y and we are using the subscript $_{\text{in}}$ to indicate what expectation values are calculated including the $e^{-i\epsilon V_{\text{int}}t}$ and which are not. We now massage the expression a bit by plugging in the background values of ϕ and adding and subtracting $\delta V'(\phi_x)$

$$\begin{aligned} -{}_{\text{in}}\langle f'(\phi)g(\chi)\rangle_{\text{in},\beta} &= i\epsilon f'(\phi_x)\int d^4y\theta(t_x-t_y)(f(\phi_y)-f(\phi_x))\langle[g(\chi_x),g(\chi_y)]\rangle_\beta - \delta V'(\phi_x) \\ \delta V'(\phi_x) &= -i\epsilon f'(\phi_x)\int d^4y\theta(t_x-t_y)f(\phi_x)\langle[g(\chi_x),g(\chi_y)]\rangle_\beta. \end{aligned} \tag{C.136}$$

The first term in this expression is what is typically used in linear response theory with $\delta V(x)$ being absorbed into the definition of $V(x)$. We can now summarize the difference between the two approaches, which we note will give exactly the same expression. Linear response theory is the ϵ^1 term in perturbation theory with "local" effects (terms that only depend on ϕ_x) pulled into the potential. These last terms were extremely briefly noted in Sec. C.2.2. The remaining terms are "non-local", namely they can contribute to derivatives terms of the inflaton, e.g. $\dot{\phi}_x$.

Finally, when evaluating Eq. C.136, the adiabatic approximation is typically used. The adiabatic approximation is

$$\begin{aligned} f(\phi_y)-f(\phi_x) &\approx f'(\phi_x)\dot{\phi}_x(t_y-t_x) \\ -\epsilon_{\text{in}}\langle f'(\phi)g(\chi)\rangle_{\text{in},\beta} &= -\gamma\dot{\phi}_x \\ \gamma &= i\epsilon^2 f'(\phi_x)^2\int d^4y\theta(t_x-t_y)(t_x-t_y)\langle[g(\chi_x),g(\chi_y)]\rangle_\beta. \end{aligned} \tag{C.137}$$

This Taylor series in time is equivalent to our Taylor series $\dot{\phi}/f \ll T$ that was done in e.g. Eq. C.13. Because of the special manner in which our inflaton coupled to the χ field, this

approximation is not needed as the time integrals give nice delta functions without the need of a Taylor series. However, the expansion makes the integrals easy to do, which is why we did the expansion.

In summary, as long as back reaction is taken into account, what we have done is identical to the typically considered linear response theory with the adiabatic approximation. We simply framed the calculation differently.

Bibliography

[1] Nima Arkani-Hamed and Juan Maldacena. Cosmological Collider Physics. 3 2015.

[2] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. Phys. Rev. D, 23:347–356, 1981.

[3] Daniel Baumann. Inflation. In Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small, pages 523–686, 2011.

[4] Scott Dodelson and Fabian Schmidt. Modern Cosmology. Academic Press, 2020.

[5] A. Friedman. On the Curvature of space. Z. Phys., 10:377–386, 1922.

[6] Georges Lemaitre. A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae. Annales Soc. Sci. Bruxelles A, 47:49–59, 1927.

[7] H. P. Robertson. Kinematics and World-Structure. Astrophys. J., 82:284–301, 1935.

[8] A. G. Walker. On Milne's Theory of World-Structure. Proc. Lond. Math. Soc. s, 2-42(1):90–127, 1937.

[9] Arno A. Penzias and Robert Woodrow Wilson. A Measurement of excess antenna temperature at 4080-Mc/s. Astrophys. J., 142:419–421, 1965.

[10] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. Astron. J., 116:1009–1038, 1998.

[11] S. Perlmutter et al. Measurements of Ω and Λ from 42 High Redshift Supernovae. Astrophys. J., 517:565–586, 1999.

[12] D. M. Scolnic et al. The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. Astrophys. J., 859(2):101, 2018.

[13] Ted Jacobson. Introduction to quantum fields in curved space-time and the Hawking effect. In School on Quantum Gravity, pages 39–89, 8 2003.

[14] T. S. Bunch and P. C. W. Davies. Quantum Field Theory in de Sitter Space: Renormalization by Point Splitting. Proc. Roy. Soc. Lond. A, 360:117–134, 1978.

[15] G. W. Gibbons and S. W. Hawking. Cosmological Event Horizons, Thermodynamics, and Particle Creation. Phys. Rev. D, 15:2738–2751, 1977.

[16] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. Astron. Astrophys., 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[17] Y. Akrami et al. Planck 2018 results. X. Constraints on inflation. Astron. Astrophys., 641:A10, 2020.

[18] Arushi Bodas, Edward Broadberry, and Raman Sundrum. Grand unification at the cosmological collider with chemical potential. JHEP, 01:115, 2025.

[19] Petar Suman, Dong-Gang Wang, Wuhyun Sohn, James R. Fergusson, and E. P. S. Shellard. How Significant are Cosmological Collider Signals in the Planck Data? 11 2025.

[20] Giovanni Cabass, Oliver H. E. Philcox, Mikhail M. Ivanov, Kazuyuki Akitsu, Shi-Fan Chen, Marko Simonović, and Matias Zaldarriaga. BOSS Constraints on Massive Particles during Inflation: The Cosmological Collider in Action. 4 2024.

[21] Wuhyun Sohn, Dong-Gang Wang, James R. Fergusson, and E. P. S. Shellard. Searching for Cosmological Collider in the Planck CMB Data. 4 2024.

[22] Gabriel Jung, Michele Citran, Bartjan van Tent, Léa Dumilly, and Nabila Aghanim. Constraints on primordial non-Gaussianity from Planck PR4 data. Astron. Astrophys., 702:A204, 2025.

[23] Soubhik Kumar, Qianshu Lu, Zhong-Zhi Xianyu, and Yisong Zhang. Cosmological Collider Searches beyond the Hubble Scale with Planck Data. 3 2026.

[24] Juan Martin Maldacena. Non-Gaussian features of primordial fluctuations in single field inflationary models. JHEP, 05:013, 2003.

[25] Giovanni Cabass, Enrico Pajer, and Fabian Schmidt. How Gaussian can our Universe be? JCAP, 01:003, 2017.

[26] Marcelo Alvarez et al. Testing Inflation with Large Scale Structure: Connecting Hopes with Reality. 12 2014.

[27] Azadeh Moradinezhad Dizgah and Cora Dvorkin. Scale-Dependent Galaxy Bias from Massive Particles with Spin during Inflation. JCAP, 01:010, 2018.

[28] Azadeh Moradinezhad Dizgah, Hayden Lee, Julian B. Muñoz, and Cora Dvorkin. Galaxy Bispectrum from Massive Spinning Particles. JCAP, 05:013, 2018.

[29] Kazuhiro Kogai, Kazuyuki Akitsu, Fabian Schmidt, and Yuko Urakawa. Galaxy imaging surveys as spin-sensitive detector for cosmological colliders. JCAP, 03:060, 2021.

[30] Xingang Chen, Yi Wang, and Zhong-Zhi Xianyu. Schwinger-Keldysh Diagrammatics for Primordial Perturbations. JCAP, 12:006, 2017.

[31] Xingang Chen and Yi Wang. Quasi-Single Field Inflation and Non-Gaussianities. JCAP, 04:027, 2010.

[32] Daniel Baumann and Daniel Green. Signatures of Supersymmetry from the Early Universe. Phys. Rev. D, 85:103520, 2012.

[33] Hayden Lee, Daniel Baumann, and Guilherme L. Pimentel. Non-Gaussianity as a Particle Detector. JHEP, 12:040, 2016.

[34] Lian-Tao Wang and Zhong-Zhi Xianyu. In Search of Large Signals at the Cosmological Collider. JHEP, 02:044, 2020.

[35] Arushi Bodas, Soubhik Kumar, and Raman Sundrum. The Scalar Chemical Potential in Cosmological Collider Physics. JHEP, 02:079, 2021.

[36] Arushi Bodas, Edward Broadberry, Raman Sundrum, and Zhaohui Xu. Charged loops at the cosmological collider with chemical potential. JHEP, 01:083, 2026.

[37] Xingang Chen, Yi Wang, and Zhong-Zhi Xianyu. Neutrino Signatures in Primordial Non-Gaussianities. JHEP, 09:022, 2018.

[38] Lian-Tao Wang and Zhong-Zhi Xianyu. Gauge Boson Signals at the Cosmological Collider. 4 2020.

[39] Nima Arkani-Hamed, Howard Georgi, and Matthew D. Schwartz. Effective field theory for massive gravitons and gravity in theory space. Annals Phys., 305:96–118, 2003.

[40] Jogesh C. Pati and Abdus Salam. Unified Lepton-Hadron Symmetry and a Gauge Theory of the Basic Interactions. Phys. Rev. D, 8:1240–1251, 1973.

[41] H. Georgi and S. L. Glashow. Unity of All Elementary Particle Forces. Phys. Rev. Lett., 32:438–441, 1974.

[42] R. L. Workman et al. Review of Particle Physics. PTEP, 2022:083C01, 2022.

[43] Michael E. Peskin and Daniel V. Schroeder. An Introduction to quantum field theory. Addison-Wesley, Reading, USA, 1995.

[44] Paul Langacker. Grand Unified Theories and Proton Decay. Phys. Rept., 72:185, 1981.

[45] Tommy Ohlsson. Proton decay. Nucl. Phys. B, 993:116268, 2023.

[46] Yoshiharu Kawamura. Gauge symmetry breaking from extra space S**1 / Z(2). Prog. Theor. Phys., 103:613–619, 2000.

[47] Yoshiharu Kawamura. Triplet doublet splitting, proton stability and extra dimension. Prog. Theor. Phys., 105:999–1006, 2001.

[48] Lawrence J. Hall and Yasunori Nomura. Grand unification in higher dimensions. Annals Phys., 306:132–156, 2003.

[49] Soubhik Kumar and Raman Sundrum. Seeing Higher-Dimensional Grand Unification In Primordial Non-Gaussianities. JHEP, 04:120, 2019.

[50] A De Rujula, H Georgi, and S.L Glashow. Trinification of all elementary particle forces. Fifth Workshop on Grand Unification, 1984.

[51] Christopher D. Carone and Justin M. Conroy. Higgsless GUT breaking and trinification. Phys. Rev. D, 70:075013, 2004.

[52] Christopher D. Carone and Justin M. Conroy. Five-dimensional trinification improved. Phys. Lett. B, 626:195–201, 2005.

[53] Nobuhito Maru and Akira Okawa. Non-Gaussianity from X, Y gauge bosons in Cosmological Collider Physics. 1 2021.

[54] Nobuhito Maru and Akira Okawa. Cosmological collider signals of non-Gaussianity from higgs boson in GUT. Int. J. Mod. Phys. A, 38(14):2350075, 2023.

[55] Xingang Chen, Yi Wang, and Zhong-Zhi Xianyu. Standard Model Mass Spectrum in Inflationary Universe. JHEP, 04:058, 2017.

[56] Lian-Tao Wang, Zhong-Zhi Xianyu, and Yi-Ming Zhong. Precision calculation of inflation correlators at one loop. JHEP, 02:085, 2022.

[57] Donald Marolf and Ian A. Morrison. The IR stability of de Sitter: Loop corrections to scalar propagators. Phys. Rev. D, 82:105032, 2010.

[58] Zhehan Qin and Zhong-Zhi Xianyu. Phase information in cosmological collider signals. JHEP, 10:192, 2022.

[59] Zhong-Zhi Xianyu and Hongyu Zhang. Bootstrapping one-loop inflation correlators with the spectral decomposition. JHEP, 04:103, 2023.

[60] Timothy Cohen, Daniel Green, and Yiwen Huang. Operator origin of anomalous dimensions in de Sitter space. Phys. Rev. D, 111(10):103513, 2025.

[61] Zhehan Qin. Cosmological correlators at the loop level. JHEP, 03:051, 2025.

[62] Sergio L. Cacciatori, Henri Epstein, and Ugo Moschella. Loops in de Sitter space. JHEP, 07:182, 2024.

[63] Tom Westerdijk and Chen Yang. Bananas are unparticles: differential equations and cosmological bootstrap. JHEP, 09:089, 2025.

[64] Paolo Benincasa, Giacomo Brunello, Manoj K. Mandal, Pierpaolo Mastrolia, and Francisco Vazão. One-loop corrections to the Bunch-Davies wave function of the universe. Phys. Rev. D, 111(8):085016, 2025.

[65] Leonardo Senatore and Matias Zaldarriaga. On Loops in Inflation. JHEP, 12:008, 2010.

[66] Leonardo Senatore and Matias Zaldarriaga. On Loops in Inflation II: IR Effects in Single Clock Inflation. JHEP, 01:109, 2013.

[67] Guilherme L. Pimentel, Leonardo Senatore, and Matias Zaldarriaga. On Loops in Inflation III: Time Independence of zeta in Single Clock Inflation. JHEP, 07:166, 2012.

[68] Jacques Bros, Henri Epstein, Michel Gaudin, Ugo Moschella, and Vincent Pasquier. Triangular invariants, three-point functions and particle stability on the de Sitter universe. Commun. Math. Phys., 295:261–288, 2010.

[69] Manuel Loparco, Joao Penedones, Kamran Salehi Vaziri, and Zimo Sun. The Källén-Lehmann representation in de Sitter spacetime. JHEP, 12:159, 2023.

[70] Matthijs Hogervorst, João Penedones, and Kamran Salehi Vaziri. Towards the non-perturbative cosmological bootstrap. JHEP, 02:162, 2023.

[71] Frank Olver, Daniel Lozier, Ronald Boisvert, and Charles Clark. The NIST Handbook of Mathematical Functions. Cambridge University Press, New York, NY, 2010-05-12 00:05:00 2010.

[72] R. N. Mohapatra. Supersymmetric grand unification: An Update. In ICTP Summer School in Particle Physics, pages 336–394, 6 1999.

[73] M. Kawasaki, Masahide Yamaguchi, and T. Yanagida. Natural chaotic inflation in supergravity. Phys. Rev. Lett., 85:3572–3575, 2000.

[74] Renata Kallosh, Andrei Linde, and Tomas Rube. General inflaton potentials in supergravity. Phys. Rev. D, 83:043507, 2011.

[75] Lawrence J. Hall and Yasunori Nomura. Gauge unification in higher dimensions. Phys. Rev. D, 64:055003, 2001.

[76] Lawrence J. Hall and Yasunori Nomura. Gauge coupling unification from unified theories in higher dimensions. Phys. Rev. D, 65:125012, 2002.

[77] Lawrence J. Hall and Yasunori Nomura. A Complete theory of grand unification in five-dimensions. Phys. Rev. D, 66:075004, 2002.

[78] Arjun Berera. Warm inflation. Phys. Rev. Lett., 75:3218–3221, 1995.

[79] Arjun Berera and Li-Zhi Fang. Thermally induced density perturbations in the inflation era. Phys. Rev. Lett., 74:1912–1915, 1995.

[80] Arjun Berera. Thermal properties of an inflationary universe. Phys. Rev. D, 54:2519–2534, 1996.

[81] Arjun Berera. Warm inflation at arbitrary adiabaticity: A Model, an existence proof for inflationary dynamics in quantum field theory. Nucl. Phys. B, 585:666–714, 2000.

[82] Arjun Berera, Ian G. Moss, and Rudnei O. Ramos. Warm Inflation and its Microphysical Basis. Rept. Prog. Phys., 72:026901, 2009.

[83] Kim V. Berghaus, Peter W. Graham, and David E. Kaplan. Minimal Warm Inflation. JCAP, 03:034, 2020. [Erratum: JCAP 10, E02 (2023)].

[84] Junichi Yokoyama and Andrei D. Linde. Is warm inflation possible? Phys. Rev. D, 60:083509, 1999.

[85] Mar Bastero-Gil, Arjun Berera, Rudnei O. Ramos, and Joao G. Rosa. Warm Little Inflaton. Phys. Rev. Lett., 117(15):151301, 2016.

[86] Mar Bastero-Gil, Arjun Berera, Rudnei O. Ramos, and João G. Rosa. Towards a reliable effective field theory of inflation. Phys. Lett. B, 813:136055, 2021.

[87] Luca Visinelli. Natural Warm Inflation. JCAP, 09:013, 2011.

[88] Hiranmaya Mishra, Subhendra Mohanty, and Akhilesh Nautiyal. Warm natural inflation. Phys. Lett. B, 710:245–250, 2012.

[89] Ricardo Z. Ferreira and Alessio Notari. Thermalized Axion Inflation. JCAP, 09:007, 2017.

[90] Ricardo Z. Ferreira and Alessio Notari. Thermalized axion inflation: natural and monomial inflation with small r. Phys. Rev. D, 97(6):063528, 2018.

[91] Vahid. Kamali. Warm pseudoscalar inflation. Phys. Rev. D, 100(4):043520, 2019.

[92] Peter Brockway Arnold and Cheng-Xing Zhai. The Three loop free energy for pure gauge QCD. Phys. Rev. D, 50:7603–7623, 1994.

[93] Eric Braaten and Agustin Nieto. Free energy of QCD at high temperature. Phys. Rev. D, 53:3421–3437, 1996.

[94] Anson Hook and Gustavo Marques-Tavares. Relaxation from particle production. JHEP, 12:101, 2016.

[95] Nayara Fonseca, Enrico Morgante, Ryosuke Sato, and Géraldine Servant. Axion fragmentation. JHEP, 04:010, 2020.

[96] Chris Graham and Ian G. Moss. Density fluctuations from warm inflation. JCAP, 07:013, 2009.

[97] Mar Bastero-Gil, Arjun Berera, Ian G. Moss, and Rudnei O. Ramos. Theory of non-Gaussianity in warm inflation. JCAP, 12:008, 2014.

[98] Mikko Laine and Aleksi Vuorinen. Basics of Thermal Field Theory, volume 925. Springer, 2016.

[99] Kim V. Berghaus, Peter W. Graham, David E. Kaplan, Guy D. Moore, and Surjeet Rajendran. Dark energy radiation. Phys. Rev. D, 104(8):083520, 2021.

[100] Marco Drewes and Sebastian Zell. On sphaleron heating in the presence of fermions. JCAP, 06:038, 2024.

[101] Kim V. Berghaus, Matthew Forslund, and Mark Vincent Guevarra. Warm inflation with a heavy QCD axion. JCAP, 10:103, 2024.

[102] Kim V. Berghaus, Marco Drewes, and Sebastian Zell. Warm Inflation with the Standard Model. 3 2025.

[103] Mar Bastero-Gil, Arjun Berera, and Rudnei O. Ramos. Shear viscous effects on the primordial power spectrum from warm inflation. JCAP, 07:030, 2011.

[104] Mehrdad Mirbabayi and Andrei Gruzinov. Shapes of non-Gaussianity in warm inflation. JCAP, 02:012, 2023.

[105] Gabriele Montefalcone, Vikas Aragam, Luca Visinelli, and Katherine Freese. WarmSPy: a numerical study of cosmological perturbations in warm inflation. JCAP, 01:032, 2024.

[106] Edward Broadberry, Anson Hook, and Sagnik Mondal. Warm Inflation with Pseudo-scalar couplings. 5 2025.

[107] Valentin Assassi, Daniel Baumann, and Daniel Green. On Soft Limits of Inflationary Correlation Functions. JCAP, 11:047, 2012.

[108] Xingang Chen and Yi Wang. Quasi-Single Field Inflation with Large Mass. JCAP, 09:021, 2012.

[109] Toshifumi Noumi, Masahide Yamaguchi, and Daisuke Yokoyama. Effective field theory approach to quasi-single field inflation and effects of heavy fields. JHEP, 06:051, 2013.

[110] Abraham Loeb and Matias Zaldarriaga. Measuring the small - scale power spectrum of cosmic density fluctuations through 21 cm tomography prior to the epoch of structure formation. Phys. Rev. Lett., 92:211301, 2004.

[111] P. Daniel Meerburg, Moritz Münchmeyer, Julian B. Muñoz, and Xingang Chen. Prospects for Cosmological Collider Physics. JCAP, 03:050, 2017.

[112] G. W. Gibbons and S. W. Hawking. Cosmological Event Horizons, Thermodynamics, and Particle Creation. Phys. Rev. D, 15:2738–2751, 1977.

[113] Chon Man Sou, Xi Tong, and Yi Wang. Chemical-potential-assisted particle production in FRW spacetimes. JHEP, 06:129, 2021.

[114] Xi Tong and Zhong-Zhi Xianyu. Large spin-2 signals at the cosmological collider. JHEP, 10:194, 2022.

[115] Viatcheslav F. Mukhanov, H. A. Feldman, and Robert H. Brandenberger. Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions. Phys. Rept., 215:203–333, 1992.

[116] Karim A. Malik and David Wands. Cosmological perturbations. Phys. Rept., 475:1–51, 2009.

[117] Y. Akrami et al. Planck 2018 results. IX. Constraints on primordial non-Gaussianity. Astron. Astrophys., 641:A9, 2020.

[118] Steven Weinberg. Quantum contributions to cosmological correlations. Phys. Rev. D, 72:043514, 2005.

[119] E. C. G. Stueckelberg. Interaction energy in electrodynamics and in the field theory of nuclear forces. Helv. Phys. Acta, 11:225–244, 1938.

[120] Soubhik Kumar and Raman Sundrum. Heavy-Lifting of Gauge Theories By Cosmic Inflation. JHEP, 05:011, 2018.

[121] Olivier Doré et al. Cosmology with the SPHEREX All-Sky Spectral Survey. 12 2014.

[122] Stefano Camera, Mario G. Santos, and Roy Maartens. Probing primordial non-Gaussianity with SKA galaxy redshift surveys: a fully relativistic analysis. Mon. Not. Roy. Astron. Soc., 448(2):1035–1043, 2015. [Erratum: Mon.Not.Roy.Astron.Soc. 467, 1505–1506 (2017)].

[123] Abraham Loeb and Matias Zaldarriaga. Measuring the small-scale power spectrum of cosmic density fluctuations through 21 cm tomography prior to the epoch of structure formation. Phys. Rev. Lett., 92:211301, 2004.

[124] Julian B. Muñoz, Yacine Ali-Haïmoud, and Marc Kamionkowski. Primordial non-gaussianity from the bispectrum of 21-cm fluctuations in the dark ages. Phys. Rev. D, 92(8):083508, 2015.

[125] Neil Barnaby, Ryo Namba, and Marco Peloso. Phenomenology of a Pseudo-Scalar Inflaton: Naturally Large Nongaussianity. JCAP, 04:009, 2011.

[126] Peter Adshead, Lauren Pearce, Marco Peloso, Michael A. Roberts, and Lorenzo Sorbo. Phenomenology of fermion production during axion inflation. JCAP, 06:020, 2018.

[127] Stephon Alexander, S. James Gates, Leah Jenks, K. Koutrolikos, and Evan McDonough. Higher Spin Supersymmetry at the Cosmological Collider: Sculpting SUSY Rilles in the CMB. JHEP, 10:156, 2019.

[128] James M. Bardeen, Paul J. Steinhardt, and Michael S. Turner. Spontaneous Creation of Almost Scale - Free Density Perturbations in an Inflationary Universe. Phys. Rev. D, 28:679, 1983.

[129] Paolo Creminelli and Matias Zaldarriaga. Single field consistency relation for the 3-point function. JCAP, 10:006, 2004.

[130] Clifford Cheung, Paolo Creminelli, A. Liam Fitzpatrick, Jared Kaplan, and Leonardo Senatore. The Effective Field Theory of Inflation. JHEP, 03:014, 2008.

[131] G. F. Giudice and A. Masiero. A Natural Solution to the mu Problem in Supergravity Theories. Phys. Lett. B, 206:480–484, 1988.

[132] Charlotte Sleight. A Mellin Space Approach to Cosmological Correlators. JHEP, 01:090, 2020.

[133] H. Bacry and J. Levy-Leblond. Possible kinematics. J. Math. Phys., 9:1605–1614, 1968.

[134] Andrei D. Linde. A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems. Phys. Lett. B, 108:389–393, 1982.

[135] Katherine Freese, Joshua A. Frieman, and Angela V. Olinto. Natural inflation with pseudo - Nambu-Goldstone bosons. Phys. Rev. Lett., 65:3233–3236, 1990.

[136] Arjun Berera and Rudnei O. Ramos. Construction of a robust warm inflation mechanism. Phys. Lett. B, 567:294–304, 2003.

[137] Mar Bastero-Gil, Arjun Berera, and Rudnei O. Ramos. Dissipation coefficients from scalar and fermion quantum field interactions. JCAP, 09:033, 2011.

[138] Gabriele Montefalcone, Vikas Aragam, Luca Visinelli, and Katherine Freese. Observational constraints on warm natural inflation. JCAP, 03:002, 2023.

[139] Mohamed M. Anber and Lorenzo Sorbo. Naturally inflating on steep potentials through electromagnetic dissipation. Phys. Rev. D, 81:043534, 2010.

[140] Mohamed M. Anber and Lorenzo Sorbo. Non-Gaussianities and chiral gravitational waves in natural steep inflation. Phys. Rev. D, 85:123537, 2012.

[141] Larry D. McLerran, Emil Mottola, and Mikhail E. Shaposhnikov. Sphalerons and Axion Dynamics in High Temperature QCD. Phys. Rev. D, 43:2027–2035, 1991.

[142] Ian G Moss and Chun Xiong. Non-Gaussianity in fluctuations from warm inflation. JCAP, 04:007, 2007.

[143] Ian G. Moss and Timothy Yeomans. Non-gaussianity in the strong regime of warm inflation. JCAP, 08:009, 2011.

[144] Yue Qiu and Lorenzo Sorbo. Spectrum of tensor perturbations in warm inflation. Phys. Rev. D, 104(8):083542, 2021.

[145] P. Klose, M. Laine, and S. Procacci. Gravitational wave background from non-Abelian reheating after axion-like inflation. JCAP, 05:021, 2022.

[146] Faqir C. Khanna, Adolfo P. C. Malbouisson, Jorge M. C. Malbouisson, and Ademir R. Santana. Thermal quantum field theory - Algebraic aspects and applications. World Scientific, 2009.

[147] Ashok K. Das. Finite Temperature Field Theory. World Scientific, New York, 1997.

[148] Y. Takahashi and H. Umezawa. Thermo field dynamics. Int. J. Mod. Phys. B, 10:1755–1805, 1996.

[149] Gerard 't Hooft. Why Do We Need Local Gauge Invariance in Theories With Vector Particles? An Introduction. NATO Sci. Ser. B, 59:101–115, 1980.

[150] Matthew Baumgart and Raman Sundrum. De Sitter Diagrammar and the Resummation of Time. JHEP, 07:119, 2020.