

ON OBSERVATIONAL SIGNATURES  
OF STRING THEORY  
IN THE COSMIC MICROWAVE BACKGROUND

A Dissertation

Presented to the Faculty of the Graduate School  
of Cornell University

in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy

by

Gang Xu

January 2013

© 2013 Gang Xu  
ALL RIGHTS RESERVED

ON OBSERVATIONAL SIGNATURES  
OF STRING THEORY  
IN THE COSMIC MICROWAVE BACKGROUND

Gang Xu, Ph.D.

Cornell University 2013

In this dissertation, we discuss the possibility of detecting evidence for string theory in the sky. Inflation, the leading paradigm for describing the very early universe, is ultraviolet (UV) sensitive. In order to understand the microphysics of inflation, it is necessary to study inflation models in a UV-complete theory, such as string theory. Recent developments in flux compactification lead to several mechanisms for inflation in which quantum corrections are under control. Here we present the observational consequences of two notable examples: warped D-brane inflation and axion monodromy inflation. After the first explicit models of warped D-brane inflation were constructed, many analyses considered the single-field approximation of these models in detail. We take a statistical approach to study the full six-field dynamics. By evolving millions of realizations of the model numerically, we find that the probability of inflation follows a power law, which is independent of the initial conditions, of the distributions from which we draw the coefficients, and of the physical parameters. Without making slow-roll approximations, we look for multifield contributions to the power spectrum from bending of the trajectories; we find that the contributions of more than two fields are non-negligible for a sizable fraction of the samples we use. However, after imposing the observational constraints on the tilt, the imprints of these multifield effects on the power spectrum and bis-

pectrum are too small to be detected. The other example we study is axion monodromy inflation. It has an approximately linear potential and produces a detectable tensor signature. This linear potential is corrected sinusoidally by instantons. We study the effects on the scalar power spectrum and bispectrum from these sinusoidal corrections. We find that there are models that satisfy all the microphysical constraints from string theory and the constraints from the current cosmological data. Moreover, a non-negligible fraction of these models also produce experimentally interesting signatures: detectable undulations of the power spectrum and/or enhanced resonant non-Gaussianity.

## BIOGRAPHICAL SKETCH

Gang Xu was born in the small fishing village of Dalian in Liaoning Province of China (population: 6 million villagers). At the age of 2, she started to recite the multiplication table and at the age of 5, she began to earn her dinner by arithmetic. She was not allowed to have hair longer than 15cm before being admitted to Tsinghua University, and studied in the program of Fundamental Sciences.

She began her graduate study at Cornell in 2005. In her experience as a teaching assistant at Cornell, she learned many valuable lessons from her students, including that it would be very inappropriate to attend a Thanksgiving dinner without wearing a turkey costume. Later on, her husband suggested to her to put exclamation marks at the end of every sentence of her papers and she identified that as a turkey-costume trap.

She started her research career by calculating a two-loop scattering amplitude in QCD and felt it was too complicated. So she switched to the field of string cosmology, where she calculated the non-perturbative corrections to warped D-brane inflation to all orders and created an inflaton potential of 333 terms. This dissertation is the result of her work in string cosmology.

To Dan, with love.

## ACKNOWLEDGEMENTS

First and foremost, I am indebted to my advisor Liam McAllister for all of his guidance throughout my graduate life. He has been an incredible mentor. Not only is he very generous with the time he is willing to spend discussing physics with me, but also he is generous with food support given to his late night string theory study group. He is a great role model to follow. He displays such passion for physics that I have been motivated to work as hard as I could and even more importantly, to love what I do. Outside academics, he acts more like a friend: he participates in student-organized events like board game nights, ultimate frisbee, and paint ball (he enthusiastically chased his students and hunted them down). But most of all, I am grateful for his understanding of my jumps of logic and unusual usage of words, without which, much of the work in this dissertation would have been much more difficult.

I would also like to thank Henry Tye for sharing his experience in life and his knowledge with me. I am grateful for his role as my mentor when I was a visiting scholar at Hong Kong University of Science and Technology.

I would also like to thank Matthias Neubert, who led me through my first project and inspired me with his magical manipulations with Mathematica.

I have learned a lot from many faculty members at Cornell and would like to thank Jim Alexander, Csaba Csáki, Éanna Flanagan, Kurt Gottfried, Yuval Grossman, Georg Hoffstaetter, André LeClair, Erich Mueller, Maxim Perelstein, Anders Ryd, James P. Sethna, Saul Teukolsky, and Julia Thom for sharing their insights.

I would like to thank the students and postdocs at Cornell, including Thomas Bachlechner, Josh Berger, Monika Blanke, David Curtin, Sergei Dyda, Sohang Gandhi, Girma Hailu, Ben Heidenreich, Johannes Heinonen, Naresh

Kumar, David Marsh, Itay Nachshon, Andrew Noble, Bibhushan Shakya, Stefan Sjörs, Flip Tanedo, Yu-Hsin Tsai, and Timm Wrase.

I am especially grateful to my collaborators Nishant Agarwal, Rachel Bean, Enrico Pajer, Raphael Flauger, Sébastien Renaux-Petel and Alexander Westphal.

I would also like to acknowledge Niayesh Afshordi, Daniel Baumann, Cliff Burgess, Mafalda Dias, Adrienne Erickcek, Ben Freivogel, Ghazal Geshnizjani, Jinn-Ouk Gong, Matt Johnson, Shamit Kachru, Takeshi Kobayashi, Louis Leblond, Paul McFadden, Paul McGuirk, Patrick Peter, Erik Schnetter, Gary Shiu, Eva Silverstein, Mark Trodden, Bret Underwood, Wessel Valkenburg, Scott Watson, Yi Wang and Jiajun Xu for inspiring discussions about the work presented in this dissertation.

I am grateful for the financial support I received from the NSF under grant PHY-0757868 and from the Institute for Advanced Study, Hong Kong University of Science and Technology. I am grateful for the hospitality I received as a visiting graduate fellow at Perimeter Institute for Theoretical Physics.

My friends, including Ziyang Cheng, Bryan Daniels, Mohammad Hamidian, Yoav Kallus, Ben Kreis, Johannes Lischner, Stefan Natu, Tchefor Ndukum, Stephen Poprocki, Darren Puigh, Sumiran Pujari, Tristan Rocheleau, Yao Weng, and Peijun Zhou made my time in Ithaca unforgettable.

I would like to thank my parents in-law, Wendy and David, and brothers in-law, Mitri and Sam, for their support.

The love and support of my parents, Jinyi Xu and Feng Cai, made this dissertation possible. Most of all I thank my adorable husband, Dan, for his love.

## TABLE OF CONTENTS

Biographical Sketch . . . . .	iii
Dedication . . . . .	iv
Acknowledgements . . . . .	v
Table of Contents . . . . .	vii
List of Tables . . . . .	x
List of Figures . . . . .	xi
<b>1 Introduction</b>	<b>1</b>
1.1 Inflation . . . . .	1
1.2 Inflation and Cosmological Observables . . . . .	2
1.3 Single-field Slow-roll Inflation . . . . .	3
1.4 Inflation Models in String Theory . . . . .	7
1.4.1 Warped D-brane Inflation . . . . .	8
1.4.2 Axion Monodromy Inflation . . . . .	11
1.5 Organization of this Dissertation . . . . .	13
<b>2 Universality in D-brane Inflation</b>	<b>15</b>
2.1 Introduction . . . . .	15
2.2 Review of Warped D-brane Inflation . . . . .	17
2.3 Methodology . . . . .	22
2.3.1 Setup . . . . .	22
2.3.2 Constructing an ensemble of potentials . . . . .	23
2.3.3 Initial conditions . . . . .	26
2.3.4 Parameters summarized . . . . .	27
2.4 Results for the Homogeneous Background . . . . .	28
2.4.1 The probability of inflation . . . . .	29
2.4.2 An analytic explanation of the exponent $\alpha = 3$ . . . . .	33
2.4.3 Dependence on initial conditions . . . . .	35
2.4.4 The DBI effect . . . . .	38
2.5 Towards the Primordial Perturbations . . . . .	41
2.5.1 Angular kinetic energy and bending trajectories . . . . .	41
2.5.2 Single-field treatment of the perturbations . . . . .	45
2.6 Conclusions . . . . .	49
<b>3 A Statistical Approach to Multifield Inflation: Many-field Perturbations Beyond Slow Roll</b>	<b>53</b>
3.1 Introduction . . . . .	53
3.2 Method . . . . .	55
3.2.1 Background evolution in D-brane inflation . . . . .	55
3.2.2 Aspects of inflection point inflation . . . . .	57
3.2.3 Perturbations in multifield inflation . . . . .	58
3.3 The scalar mass spectrum and violations of slow roll . . . . .	65

3.3.1	Properties of the mass spectrum . . . . .	66
3.3.2	A random matrix model for the masses . . . . .	68
3.3.3	Violations of slow roll . . . . .	73
3.4	Multifield effects . . . . .	75
3.4.1	Decay of entropic perturbations . . . . .	76
3.4.2	Bending of the trajectory . . . . .	78
3.4.3	Multifield effects and many-field effects . . . . .	80
3.4.4	Constraints from scale invariance . . . . .	91
3.4.5	Non-Gaussianities . . . . .	95
3.5	Conclusions . . . . .	102
<b>4</b>	<b>Oscillations in the CMB from Axion Monodromy Inflation</b>	<b>107</b>
4.1	Introduction . . . . .	107
4.1.1	Review of axion monodromy inflation . . . . .	110
4.2	Background Evolution . . . . .	112
4.3	Spectrum of Scalar Perturbations . . . . .	115
4.3.1	Analytic solution of the Mukhanov-Sasaki equation . . . . .	117
4.3.2	Saddle-point approximation . . . . .	120
4.3.3	Particle production and deviations from the Bunch-Davies state . . . . .	124
4.3.4	Bispectrum of scalar perturbations . . . . .	130
4.4	Observational Constraints . . . . .	133
4.5	Microphysics of Axion Monodromy Inflation . . . . .	139
4.5.1	Axions in string theory . . . . .	139
4.5.2	Dimensional reduction and moduli stabilization . . . . .	141
4.5.3	Axion decay constants in string theory . . . . .	148
4.6	Microscopic Constraints . . . . .	151
4.6.1	Constraints from computability . . . . .	152
4.6.2	Constraints from backreaction on the geometry . . . . .	155
4.6.3	Constraints from higher-derivative terms . . . . .	158
4.6.4	Constraints on the axion decay constant . . . . .	160
4.6.5	Constraints on the amplitude of the modulations . . . . .	163
4.7	Combined Theoretical and Observational Constraints . . . . .	168
4.8	Conclusions . . . . .	171
<b>A</b>	<b>Chapter 1 of Appendix</b>	<b>181</b>
A.1	Structure of the Scalar Potential . . . . .	181
<b>B</b>	<b>Chapter 2 of Appendix</b>	<b>183</b>
B.1	Notation and Conventions . . . . .	183
B.2	Induced Shift of the Four-Cycle Volume . . . . .	184
B.2.1	A simple illustration of the suppression mechanism . . . . .	188
B.2.2	The conifold and its resolution . . . . .	190
B.2.3	The shift of the four-cycle volume . . . . .	193

B.3	The Kaluza-Klein Spectrum . . . . .	196
B.3.1	The effective theory . . . . .	196
B.3.2	Effects of the light Kaluza-Klein modes . . . . .	198
B.4	Numerical Examples . . . . .	200
B.4.1	Intersection numbers: set I . . . . .	200
B.4.2	Intersection numbers: set II . . . . .	201
<b>Bibliography</b>		<b>203</b>

## LIST OF TABLES

2.1	Summary of the power law fits to the success probabilities for various scenarios. . . . .	33
2.2	Summary of the power law fits to the success probabilities for various distributions. . . . .	33

## LIST OF FIGURES

2.1	Examples of downward-spiraling trajectories for a particular realization of the potential. . . . .	28
2.2	The rms value, $Q$ , of the coefficients $c_{LM}$ has a significant role in determining whether inflation can occur. . . . .	31
2.3	The likelihood of $N_e$ e-folds of inflation as a function of $N_e$ , for $\Delta_{\max} = 7.8$ and $N_f = 6$ . . . . .	32
2.4	Trajectories in the $\theta_1 - \theta_2$ plane, with $\log(x)$ vertical, for a fixed potential. . . . .	36
2.5	The ratio $KE_\psi/KE_r$ of angular to radial kinetic energies for trials with $\Delta_{\max} = 7.8$ and $N_f = 6$ . . . . .	43
2.6	Hubble slow roll parameters $\epsilon_{60}$ and $\eta_{60}$ as a functions of the maximum number of e-folds. . . . .	46
2.7	An inflection-point potential in one dimension, taken from [9]. . . . .	47
2.8	Scalar power $A_s$ compared to the measured central value, $A_s^* = 2.43 \times 10^{-9}$ , as a function of the total number of e-folds. . . . .	48
2.9	Scalar tilt $n_s$ and tensor-to-scalar ratio $r$ for cases for which the scalar power $A_s$ is consistent with WMAP7 at $2\sigma$ . . . . .	50
3.1	A single-field inflection point potential, taken from [9]. . . . .	58
3.2	The mass spectra of the six scalar fields, in units of $H^2$ . . . . .	66
3.3	The mass spectrum simulated in the matrix model given in equation (3.19), cf. [94]. . . . .	69
3.4	Histogram showing the relative probability of the number $n_f$ of scalar fields that are light enough to fluctuate during inflation, cf. equation (3.27). . . . .	73
3.5	The mass-squared of the adiabatic fluctuation in units of $H$ , evaluated 60 e-folds before the end of inflation, versus the total number of e-folds, $N_e$ . . . . .	74
3.6	The ratio of the one-field power to the naive power, versus the adiabatic mass-squared in units of $H$ . . . . .	74
3.7	The difference between the exact and naive spectral tilts, versus the adiabatic mass-squared in units of $H$ , for effectively single-field realizations (see §3.2.3). . . . .	75
3.8	The curves show the rapid decay of the total power of the five entropic modes over the course of inflation, for 19 realizations. . . . .	77
3.9	The curves show the rapid changes in $\eta_\perp$ during the first 15 e-folds for 9 realizations of inflation. . . . .	80
3.10	Left panel: the slow roll parameter $\eta_\perp$ evaluated at Hubble crossing, versus the total number of e-folds, $N_e$ . . . . .	81
3.11	Scalar power spectra as functions of the number of e-folds since Hubble crossing, for four different realizations. . . . .	83

3.12	The relative probability $P(\xi \geq \lambda)$ of different levels of corrections of the scalar power. . . . .	84
3.13	The ratio of the exact power to the one-field power, versus the total turn of the trajectory during inflation, for all realizations. . . . .	85
3.14	Many-field and two-field effects, for effectively multifield models (see §3.2.3). . . . .	87
3.15	The difference $\delta n_s$ between the exact and naive results for the tilt $n_s$ , versus the naive tilt, for effectively multifield realizations (see §3.2.3). . . . .	93
3.16	Left panel: the slow roll parameter $\eta_\perp$ evaluated at Hubble crossing, versus the total number of e-folds, $N_e$ , for models with tilt $n_s$ consistent with observations. . . . .	96
3.17	An example in which multifield effects — specifically, two-field effects — are large, and the power spectrum is consistent with observations. . . . .	97
4.1	The solid line is the analytical result for $\delta n_s$ as a function of $f$ , for $b = 0.08$ , while the dots are the numerical result obtained from an adaptation of the code used in [117]. . . . .	121
4.2	This plot shows the 68% and 95% likelihood contours in the $\delta n_s$ - $f$ , and $bf$ - $\log_{10} f$ plane, respectively, from the five-year WMAP data on the temperature angular power spectrum. . . . .	175
4.3	The left plot shows the angular power spectrum for the best fit point $f = 6.67 \times 10^{-4}$ , and $\delta n_s = 0.17$ . . . . .	175
4.4	These plots show the 68% and 95% likelihood contours for the five-year WMAP data on the temperature angular power spectrum in the $\delta n_s - \Omega_B h^2$ plane and $\delta n_s - \Delta\phi$ plane for an axion decay constant of $f = 3 \times 10^{-2}$ and $f = 1.5 \times 10^{-2}$ , respectively. . . . .	176
4.5	This figure shows a triangle plot for some of the parameters that were sampled in a Markov chain Monte Carlo for an axion decay constant of $f = 10^{-2}$ . . . . .	177
4.6	We show the (one- and two-sigma) likelihood contours for the temperature two-point function together with three contours that characterize the amplitude of the three point function, for $f_{res} = 200, 20, 2$ . . . . .	178
4.7	We superimpose the theoretical constraints, summarized in (4.155) and (4.156), on the constraints imposed by observations, which are shown in figure 4.6. . . . .	179
4.8	The blue dot represents the explicit numerical example presented in appendix B.4. . . . .	180
B.1	This diagram illustrates the positions of the A and B stacks of D3-branes in $\mathbb{R}^6$ , the choice of the angular coordinate $\theta$ , and, in blue, a (topologically trivial) four-cycle. . . . .	188

B.2 1st row: Plot of  $\mathcal{V}_{warped}^{(0)}$  and  $\delta\mathcal{V}_{warped}(\delta\theta_2)$  as functions of  $\mu$  at constant  $u = 0.01$ . . . . . 195

# CHAPTER 1

## INTRODUCTION

### 1.1 Inflation

Much of our current knowledge about the early universe results from the observations of the cosmic microwave background (CMB) radiation, the discovery of which established the Big Bang Theory as our standard model of cosmology.

The CMB is incredibly smooth across the whole sky: it has a black body spectrum and its temperature in various directions differs by only about one part in  $10^5$ . At first glance this fact is surprising, as the CMB photons from different directions originated in regions that had never been in causal contact before the time when the CMB was formed according to the original Big Bang Theory.

Instead of believing that the universe began in a outrageously homogeneous status and hence evolved homogeneously to agree with the observed CMB, we would like to seek a theory with more predictive power. The theory of inflation introduced an early period of accelerated expansion to solve this "horizon problem" elegantly: before this brief but violent expansion the entire observable universe was in causal contact. Inflation also alleviates the other problems faced by the Big Bang Theory, such as why the universe is so homogeneous, flat, and empty.

On the other hand, although the CMB is smooth to a striking extent, the inhomogeneities observed in the CMB are large enough to seed the large-scale structure we observe in the universe.

Inflation[1, 2, 3] provides an elegant mechanism to generate these initial seeds via quantum mechanics. For different regions of space, small quantum fluctuations of the inflaton field(s) create spatial differences in the time at which inflation ends. After reheating, the differences in the ending time become energy density fluctuations, which are observed in the CMB.

Not only does the theory of inflation match with current observations incredibly well, different models of the inflation theory make different predictions about the cosmological observables, allowing them to be falsified by the upcoming CMB experiments.

## 1.2 Inflation and Cosmological Observables

The cosmological observables include the scalar power spectrum, inflationary gravitational waves and primordial non-Gaussianity.

For any given inflation model, we can calculate the amplitude of the primordial scalar power spectrum and the tilt, which measures how deviated the power spectrum is from scale-invariance. Experimentally, we can obtain the same information by measuring the two-point correlation function of the temperature anisotropies in the CMB. In specific inflation models, although the amplitude of the power spectrum can be easily tuned to agree with the experimental results, the tilt is constrained to lie within a relatively narrow range. Another very useful observable is the ratio of the amplitude of the primordial tensor and scalar power spectra. We have not observed any primordial gravitational waves, thus we can place an upper bound on this ratio. The combination of the constraints on the tilt and the tensor-to-scalar ratio already rule out certain

inflation models. If the Planck satellite observes primordial gravitational waves in the next few years, we will learn a lot about the inflation epoch. First, observation of a non-zero tensor-to-scalar ratio will tell us about the energy scale of inflation. Second, detectable gravitational waves imply a super-Planckian field excursion of the inflaton field.

The simplest inflation models predict the primordial fluctuations to be highly Gaussian, and Gaussian fluctuations are completely described by the power spectrum, or two-point correlation function. The non-Gaussianity generated in such models is too small to ever be detected. However, non-trivial kinetic terms, multiple fields and/or features in the inflaton potential can enhance the primordial non-Gaussianity by one or two orders of magnitude relative to the simplest models, creating a potentially observable signature. The CMB experiments attempt to detect non-Gaussianity by extracting information from the three-point function, the bispectrum.

### 1.3 Single-field Slow-roll Inflation

We will now examine in more details the successes and problems with the simplest phenomenological models of inflation.

In the simplest models of inflation, there is a single scalar field  $\phi$ , the inflaton. We assume the inflaton is minimally coupled to gravity and we can describe the time-evolution of the inflationary energy density with the action of a scalar field with canonical kinetic term and the inflaton potential,  $V(\phi)$ , which describes the

self-interactions of the inflaton.

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[ R - (\nabla\phi)^2 - 2V(\phi) \right], \quad (1.1)$$

in units where  $M_p^{-2} \equiv 8\pi G = 1$ . Assuming the background geometry is described by the Friedmann-Robertson-Walker (FRW) metric for the spacetime of the universe

$$ds^2 = -dt^2 + a^2(t)d\vec{r}^2. \quad (1.2)$$

where  $a(t)$  is the scale factor. Thus the dynamics of the scalar field is determined by

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi} = 0 \quad (1.3)$$

where  $H \equiv \frac{\dot{a}}{a}$  is the Hubble parameter and it is determined by

$$H^2 = \frac{1}{3} \left( \frac{1}{2} \dot{\phi}^2 + V(\phi) \right) \quad (1.4)$$

We define slow-roll parameters  $\epsilon$  and  $\eta$  as follows

$$\epsilon = -\frac{\dot{H}}{H^2} \quad (1.5)$$

and

$$\eta = -\frac{\ddot{\phi}}{H\dot{\phi}} \quad (1.6)$$

Accelerated expansion requires  $\epsilon < 1$ . In order for accelerated expansion to persist, we also require  $\eta < 1$ .

To agree with the observed curvature of the universe, the universe needs to expand at least by a factor of  $e^{60}$ . In order for the inflation to persist for such a long time, both slow-roll parameters need to remain small. We can relate the slow-roll parameters to the shape of the inflaton potential.

$$\epsilon \approx \epsilon_V = \frac{1}{2} \left( \frac{V_{,\phi}}{V(\phi)} \right)^2 \quad (1.7)$$

$$\eta \approx \eta_V = \frac{V_{,\phi\phi}}{V(\phi)} \quad (1.8)$$

Requiring slow-roll parameters to be small is equivalent to requiring inflaton potential to be flat. In other words, both the first derivative and the second derivative of the inflaton potential need to remain small in Planck units during the inflation epoch. In particular, the inflaton needs to be light.

We can calculate the number of e-folds in single-field slow-roll inflation,

$$N(\phi) \equiv \ln \frac{a_{end}}{a} = \int_t^{t_{end}} H dt = \int_{\phi}^{\phi_{end}} \frac{H}{\dot{\phi}} d\phi \approx \int_{\phi_{end}}^{\phi} \frac{V}{V_{,\phi}} d\phi,$$

where we used the slow-roll results. In terms of slow roll parameters, this can be written as

$$N(\phi) \approx \int_{\phi_{end}}^{\phi} \frac{d\phi}{\sqrt{2\epsilon_V}}. \quad (1.9)$$

We can calculate the power spectra of the fluctuations by expanding the action to the second order in fluctuations and derive the action of the curvature perturbation. The curvature perturbation is conserved outside the horizon, which enables us to evaluate everything at horizon crossing,  $k = aH$ . We then solve the equation of motion for the curvature perturbation and calculate its power spectra.

The results for the power spectra of the scalar and tensor fluctuations created by inflation are

$$\Delta_s^2(k) = \frac{1}{8\pi^2} \frac{H^2}{M_p^2} \frac{1}{\epsilon_V} \Big|_{k=aH}, \quad (1.10)$$

$$\Delta_t^2(k) = \frac{2}{\pi^2} \frac{H^2}{M_p^2} \Big|_{k=aH}, \quad (1.11)$$

The tensor-to-scalar ratio is

$$r \equiv \frac{\Delta_t^2}{\Delta_s^2} = 16 \epsilon_{V,\star} . \quad (1.12)$$

where  $\star$  means calculated at horizon crossing.

The scale dependence of the spectra is quantified by the spectral indices

$$n_s - 1 \equiv \frac{d \ln \Delta_s^2}{d \ln k} , \quad n_t \equiv \frac{d \ln \Delta_t^2}{d \ln k} . \quad (1.13)$$

Note that the tensor-to-scalar ratio relates directly to the evolution of the inflaton as a function of e-folds  $N$

$$r = \frac{8}{M_p^2} \left( \frac{V_{,\phi}}{V(\phi)} \right)^2 = \frac{8}{M_p^2} \left( \frac{\dot{\phi}}{H} \right)^2 = \frac{8}{M_p^2} \left( \frac{d\phi}{dN} \right)^2 . \quad (1.14)$$

The total field excursion can therefore be written as the following integral

$$\frac{\Delta\phi}{M_p} = \int_{N_{end}}^{N_{cmb}} dN \sqrt{\frac{r}{8}} . \quad (1.15)$$

During slow-roll evolution, one may obtain the following approximate relation [4]

$$\frac{\Delta\phi}{M_p} = \mathcal{O}(1) \times \left( \frac{r}{0.01} \right)^{1/2} , \quad (1.16)$$

where  $r(N_{cmb})$  is the tensor-to-scalar ratio on CMB scales. Large values of the tensor-to-scalar ratio,  $r > 0.01$ , therefore correlate with  $\Delta\phi > M_p$  or large-field inflation.

Single-field slow-roll inflation is extremely successful phenomenologically: with a suitable inflaton potential  $V(\phi)$ , prolonged inflation of 60 or more e-folds can be generated to solve the puzzles for the Big Bang theory, such as the horizon problem. The power spectrum generated by this type of model is almost scale-invariant, which agrees well with the experimental data. Single-field slow-roll inflation predicts an unmeasurable small amount of non-Gaussianity, and

non-Gaussianity has not been detected. Depending on the total excursion of the inflaton field in the field space, it can predict either observable gravitational waves or not.

However, from a theoretical point of view, single-field slow-roll inflation cannot be the complete story.

For one thing, we do not know the physical nature of this field, which we currently only treat as an order parameter; we cannot explain why only one scalar field is chosen to be the inflaton field, except that this is the simplest choice. For another, we cannot explain why the chosen potential  $V(\phi)$  can be so flat considering the corrections the potential may receive from higher-order operators.

String theory can help.

## 1.4 Inflation Models in String Theory

In most theories, inflation occurred at extremely high energies, far beyond the energies at which the known laws of physics have been tested. It is uncertain what physics governed this epoch. Thus, a phenomenological approach is taken: we first postulate an effective inflaton potential, then calculate the cosmological observables, and then we compare the predicted cosmological observables with the experimental data and obtain constraints on the parameters of this potential.

This approach is not entirely satisfactory. Without knowing the physics at the Planck scale, the phenomenological effective inflaton potential suffers from

incalculable corrections from higher order Planck-suppressed operators. And more often than not, a slight change in the inflaton potential will destroy the prolonged inflation and render the predictions useless.

The simplest example of these dangerous higher order operators is  $V(\phi)\frac{\phi^2}{M_p^2}$ . From the effective theory point of view, if  $V(\phi)$  is a four dimensional operator in the system, this will be a natural correction with order one coefficient. This correction will destroy the designed-to-be-flat potential  $V(\phi)$ , and inflation will last much less than the needed 60 e-folds.

Thus the inflaton potential should be derived from a fundamental theory of quantum gravity, such as string theory, where either we are able to calculate the corrections to the inflaton potential to all orders or we have a symmetry unbroken up to the Planck scale to forbid the unwanted higher-order operators. We will talk about these two different approaches in the following subsections.

### 1.4.1 Warped D-brane Inflation

Four-dimensional effective actions that arise from string theory and are used for inflation typically contain many scalar fields, the moduli. In particular, the compactification manifolds often have parameters that control the size and shape of the extra dimensions, which become scalar fields in the low energy effective theory in four dimensions. This is great for inflation, since there are problems in the model when the Higgs boson, the only fundamental scalar discovered so far, is considered to be the inflaton. But before we can use some of these scalar fields for inflation, we must stabilize the moduli.

The existence of unstabilized moduli is problematic for cosmology: these moduli typically have gravitational strength couplings, and their evolution can affect low energy observables such as the couplings of the Standard Model and Newton's gravitational constant, which are strongly bounded by experiments. In addition, the moduli, displaced from their zero-temperature minima, will store energy during inflation. Unless the moduli acquire a mass more than 30 TeV, they will either overclose the universe or dilute the products of Big Bang nucleosynthesis, depending on their masses. In particular, the overall compactification volume is the most dangerous one: with an environment of positive energy density, which is indicated by the present-day acceleration, it becomes energetically favorable for the compactification volume to increase, and we end up with runaway decompactification.

The moduli stabilization problem has recently been solved in the setting of flux compactifications [5] of type IIB string theory. The inclusion of two real integrally quantized three-form fluxes generally fixes all the complex structure moduli, which control the shape of the compact manifold, and the dilaton. To fix the Kähler moduli, which control the size of each even-dimensional cycle and the overall size of the internal space, non-perturbative effects from gaugino condensation on D7-branes or Euclidean D3-instantons are used.

Such a flux compactification may include a warped throat region, and the interaction between an anti-D3 brane sitting at the tip of the throat and a mobile D3-brane near the top of the throat could be weak enough to support inflation. In this case, the separation between the pairs of the branes is the inflaton. The inflaton potential has several terms [6]: the Coulomb interaction between the brane-antibrane pair, the coupling with the four dimensional scalar curva-

ture, and the non-perturbative effects from the specific compactification. The Coulomb potential is almost entirely flat; it provides a strong attractive force when the inflaton is very close to the tip of the throat and thus terminates inflation. The curvature coupling gives a fixed term in the inflaton potential that is too steep to support prolonged inflation. The hope is that the numerous non-perturbative effects will accidentally cancel out the curvature contribution and yield an approximate flat potential for the inflaton.

One particular compactification has been studied [7, 8, 9], and the authors demonstrated that it was possible to achieve an inflection point potential with significant fine-tuning. They assumed that the angular directions were already minimized and only studied the radial direction.

Although the previous treatment studied warped D-brane inflation consistently with a specific setup and in the single-field approximation, there are lots of questions that remain to be asked.

- Not all compactifications can be explicitly calculated, can we somehow learn information about how inflation happens in other compactifications?
- If we do not know the specific compactification scheme, but we know the general form of the contributions they make to the inflaton potentials, what can we learn about inflation in a general compactification?
- In some scenarios, the inflaton might not be able to find the preferred angular location. Instead of rolling down radially, it could follow a direction that is neither purely radial nor purely angular. The inflaton could change direction from time to time. Can we approximate the inflaton potential with a single-field potential without losing useful information?

- Is there anything we can learn about this setup that is general enough that we can extend our conclusions to scenarios that are completely different from warped D-brane inflation?
- What are the consequences of the five angular directions? How many fields are actually light enough to fluctuate, and thus affect the power spectrum? As inflaton will inevitably turn in this multi-field setup, can we observe this turning in the experimental data? Will this turning leave some traceable imprint in the spectrum and bi-spectrum? Would there be detectable isocurvature perturbations?

In chapter 2 and 3, we address most of these questions and give our answers.

## 1.4.2 Axion Monodromy Inflation

In the case of large-field inflation, where the inflaton traverses a distance of more than a Planck mass in its field space, the UV sensitivity of inflation is dramatically enhanced: the potential becomes sensitive to an infinite series of operators with arbitrary dimensions. Since the prolonged inflation in this case requires the inflaton potential to be flat over a super-Planckian range, if inflation arises by accident, it requires a cancellation among an infinite number of terms with order one coefficients. Instead, an approximate symmetry can be invoked to forbid these dangerous operators. The shift symmetry, in which the inflaton potential is left unchanged if the inflaton is moved in the field space by a constant, will protect the inflaton potential in a natural way.

String theory provides these ingredients naturally: there are many axions and they enjoy shift symmetries, which are only broken non-perturbatively.

Axion monodromy inflation[10, 11] is a successful model that uses the shift symmetry of the axions to protect the flat potential. The linear part of the inflaton potential comes from the Dirac-Born-Infeld (DBI) action, and this is the flat potential that can support 60 e-folds of inflation if the axion was somehow displaced by at least 11 Planck units initially. Since the inflaton transverses a super-Planckian distance in the field space, this model has the distinctive signal of observable primordial gravitational waves. The satellite Planck will be able to observe these primordial gravitational waves in the next few years.

But since the shift symmetry is broken by instantons non-perturbatively, this linear potential is sinusoidally corrected. We ask ourselves the following questions about the effect of this correction and attempt to answer them in chapter 4.

- Will this sinusoidal correction leave a fingerprint in the power spectrum that can be detected before the gravitational waves?
- Can the current data already rule out part of the parameter space?
- Will these "bumps" on the inflaton potential be large enough to generate detectable resonant non-Gaussianity?
- What are the constraints that string theory microphysics places on the parameters?
- Are there any models of this type which completely satisfy the constraints of the microphysics, and are also experimentally interesting - i.e. not yet ruled out, but falsifiable in the next few years?

## 1.5 Organization of this Dissertation

In Chapter 2 we study the six-field dynamics of D3-brane inflation statistically for a general scalar potential on the conifold. In Chapter 3, using the ensemble of six-field inflationary models studied in Chapter 2, we study the multifield contribution to the scalar power spectrum. In Chapter 4, we study the observational signatures of axion monodromy inflation.

In Chapter 2 we examine an ensemble of over  $7 \times 10^7$  realizations of warped D-brane inflation with full six-field dynamics by numerically evolving the equations of motion. We find that the detailed properties of the statistical distribution from which we draw the coefficients have small effects on our results. We find that prolonged inflation occurs when there is an accidental cancellation among numerous terms in the potential: during the successful trials, the D3-brane moved rapidly in all directions for the first few e-folds, spiraling down to an inflection point, where most of the evolution became effectively single-field. We find that the probability of  $N_e$  e-folds of inflation is a power law and we explain the origin of this exponent analytically.

In Chapter 3 we investigate an ensemble of six-field inflationary models in the same setup as in Chapter 2, the warped D-brane inflation. We focus on studying the multifield contributions to the scalar power spectrum. Without making a slow-roll approximation, we calculate the primordial perturbations numerically. As we find in Chapter 2, most of inflation occurs around an inflection point. Before the inflaton reaches the inflection point, violations of slow roll and bending trajectories are common, and "many-field" effects, in which three or more fields influence the perturbations, are often important. However, the

scalar power spectrum is typically blue above the inflection point, becoming red only below the inflection point. Thus in a large fraction of models consistent with constraints on the tilt, the imprint in the perturbations left by the multifield evolution cannot be observed. In the cases where multifield effects are important, the quasi-single-field inflation is realized microphysically, but the cubic and quartic couplings are not big enough to produce detectable non-Gaussianity.

In Chapter 4 we study the observable signatures of axion monodromy inflation. In this scenario, the approximately linear inflaton potential is periodically modulated. We first study the phenomenological potential that capture the essential effects and find that these modulations yield two striking observational signatures: undulations in the scalar power spectrum and resonant enhancement of the bispectrum. We then compare our results with experimental data and place constraints on the amplitude and frequency of the periodic modulations. Furthermore, we investigate the realization in string theory and use the compactification data to compute the decay constant and the magnitude of the modulation. We find that it is possible to produce models with detectable modulations of the scalar power spectrum and/or bispectrum, which are consistent with the observational data and the known microphysical constraints.

## CHAPTER 2

### UNIVERSALITY IN D-BRANE INFLATION

#### 2.1 Introduction

Inflation [1, 2, 3] provides a compelling explanation for the large-scale homogeneity of the universe and for the observed spectrum of cosmic microwave background (CMB) anisotropies. However, in a large fraction of the multitude of inflationary models — e.g., in most small-field models — the success and the predictions of inflation are sensitive to small changes in the inflaton Lagrangian and initial conditions. Without a priori measures on the space of scalar field Lagrangians and on the corresponding phase space, it is difficult to test a given model of inflation.

In this paper we find robust predictions in a surprising place: warped D-brane inflation [5], a well-studied scenario for inflation in string theory in which six (or more) dynamical fields are governed by a scalar potential with hundreds of terms. We show that the collective effect of many terms in the potential is accurately described by a simple and predictive phenomenological model. The essential idea behind our approach is that in an inflationary model whose potential involves the sum of many terms depending on multiple fields, one can expect a degree of emergent simplicity, which may be thought of as central limit behavior.

Our primary method is a comprehensive Monte Carlo analysis. Recent results [6] provide the structure of the scalar potential in warped D-brane inflation, i.e. a list of all possible terms in the potential, with undetermined, model-

dependent coefficients. A realization of warped D-brane inflation then consists of a choice of coefficients together with a choice of initial conditions. We construct an ensemble of realizations, drawing the coefficients from a range of statistical distributions and truncating the potential to contain 27, 237, and 334 independent terms, corresponding to contributions from Planck-suppressed operators with maximum dimensions of 6, 7, and  $\sqrt{28} - 3/2 \approx 7.79$ , respectively. We then numerically evolve the equations of motion for the homogeneous background and identify robust observables that have demonstrably weak dependence on the statistical distribution, on the degree of truncation, and on the initial data. In particular, we find that the probability of  $N_e$  e-folds of inflation is a power law,  $P(N_e) \propto N_e^{-3}$ , and we present a very simple analytical model of inflection point inflation that reproduces this exponent.

To study the primordial perturbations, we focus on the subset of realizations in which the dynamics during the final 60 e-folds is that of single-field slow roll inflation. (In the remaining realizations, multifield effects can be significant, and a dedicated analysis is required.) For these cases, we find that primordial perturbations consistent with WMAP7 [13] constraints on the scalar spectral index,  $n_s$ , are possible only in realizations yielding  $N_e \gtrsim 120$  e-folds. In favorable regions of the parameter space, a universe consistent with observations arises approximately once in  $10^5$  trials.

The plan of this paper is as follows. In §2.2 we recall the setup of warped D-brane inflation, and in §2.3 we explain how we construct and study an ensemble of realizations. Results for the homogeneous background evolution appear in §2.4, while the perturbations are studied in §2.5. We conclude in §3.5. Appendix A.1 summarizes the structure of the inflaton potential, following [6].

## 2.2 Review of Warped D-brane Inflation

In the simplest models of warped D-brane inflation,<sup>1</sup> the inflaton field  $\phi$  is identified as the separation between a D3-brane and an anti-D3-brane along the radial direction of a warped throat region of a flux compactification [5]. (We will have much to say about more complicated inflationary trajectories that involve angular motion.) The D3-brane potential receives a rich array of contributions, from the Coulomb interaction of the brane-antibrane pair, from the coupling to four-dimensional scalar curvature, and from nonperturbative effects that stabilize the Kähler moduli of the compactification. The curvature coupling yields a significant inflaton mass, and in the absence of any comparable contributions, the slow roll parameter  $\eta$  obeys  $\eta \approx 2/3$  [5], which is inconsistent with prolonged inflation.

The moduli-stabilizing potential does generically make significant contributions to the inflaton potential, and many authors have taken the attitude that within the vast space of string vacua, in some fraction the moduli potential will by chance provide an approximate cancellation of the inflaton mass, so that  $\eta \ll 1$ . To do better, one needs to know the form of the moduli potential. The nonperturbative superpotential was computed in [17] for a special class of configurations in which a stack of D7-branes falls inside the throat region. For this case, Refs. [7, 8, 9, 18] studied the possibility of inflation, and found that fine-tuned inflation, at an approximate inflection point, is indeed possible [7, 8, 9].

This situation is unsatisfactory in several ways. First, the restriction to compactifications in which D7-branes enter the throat is artificial, and serves to enhance the role of known terms in the inflaton potential (those arising from in-

---

<sup>1</sup>See [14], [15], [16] for foundational work on brane inflation.

interactions with the nearby D7-branes) over more general contributions from the bulk of the compactification. Second, the analyses of [7, 8, 9, 18, 19, 20, 21, 22] treated special cases in which the D3-brane tracked a minimum along some or all of the angular directions of the conifold, sharply reducing the dimensionality of the system, but there is no reason to believe that this situation is generic. Therefore, although these works do provide consistent treatments of inflation in special configurations, an analysis that studies the full six-dimensional dynamics in a general potential is strongly motivated.<sup>2</sup>

The results of [6, 26] provide the necessary information about the D3-brane potential. As explained in detail in [6], the most general potential for a D3-brane on the conifold corresponds to a general supergravity solution in a particular perturbation expansion around the Klebanov-Strassler solution. The most significant terms in this potential arise from supergravity modes corresponding to the most relevant operators in the dual CFT, and by consulting the known spectrum of Kaluza-Klein modes, one can write down the leading terms in the inflaton potential, up to undetermined Wilson coefficients. The physical picture is that effects in the bulk of the compactification, e.g. gaugino condensation on D7-branes, or distant supersymmetry breaking, distort the upper reaches of the throat, leading to perturbations of the solution near the location of the D3-brane.

To describe the D3-brane action, we begin with the background geometry, which is a finite region of the warped deformed conifold. Working far above the tip and ignoring logarithmic corrections to the warp factor, the line element is

$$ds^2 = \left(\frac{R}{r}\right)^2 g_{ij} dy^i dy^j = \left(\frac{R}{r}\right)^2 (dr^2 + r^2 ds_{T^{1,1}}^2), \quad (2.1)$$

---

<sup>2</sup>See [23, 24] for detailed studies of multifield effects at the end of D-brane inflation in the framework of [9], and [25] for a systematic exploration of the likelihood of inflation in this context.

where  $r$  is the radial direction of the cone, and the base space  $T^{1,1}$  is parameterized by five angles,  $0 \leq \theta_1, \theta_2 \leq \pi$ ,  $0 \leq \phi_1, \phi_2 < 2\pi$ ,  $0 \leq \psi < 4\pi$ , which we shall collectively denote by  $\Psi$ . The radius  $R$  is given by  $R^4 = \frac{27}{4}\pi g_s N \alpha'^2$ , with  $N \gg 1$  the D3-brane charge of the throat. At radial coordinate  $r_{UV} \approx R$ , the throat smoothly attaches to the remainder of the compact space, which we refer to as the bulk. On the other hand, the deformation is significant in the vicinity of the tip, at  $r_0 \approx a_0 R$ , with  $a_0$  denoting the warp factor at the tip. We will perform our analysis in the region  $a_0 R \ll r < r_{UV}$ , where the singular conifold approximation is applicable, and will work with a rescaled radial coordinate  $x \equiv \frac{r}{r_{UV}} < 1$ . Finally, because the D3-brane kinetic term is insensitive to warping at the two-derivative level (cf. §2.4.4), the metric on the inflaton field space is the *unwarped* metric  $g_{ij}$ .

The D3-brane potential is usefully divided into four parts,

$$V(x, \Psi) = V_0 + V_C(x) + V_{\mathcal{R}}(x) + V_{\text{bulk}}(x, \Psi), \quad (2.2)$$

which we will discuss in turn. First, the constant  $V_0$  represents possible contributions from distant sources of supersymmetry breaking, e.g. in other throats. Next, the Coulomb potential  $V_C$  between an anti-D3-brane at the bottom of the throat and the mobile D3-brane has the leading terms

$$V_C = D_0 \left( 1 - \frac{27 D_0}{64 \pi^2 T_3^2 r_{UV}^4} \frac{1}{x^4} \right), \quad (2.3)$$

where  $T_3$  is the D3-brane tension and  $D_0 = 2a_0^4 T_3$ . Higher-multipole terms in the Coulomb potential depend on the angles  $\Psi$ , but are suppressed by additional powers of  $a_0$  and may be neglected in our analysis.

Upon defining the scale  $\mu^4 = (V_0 + D_0) \left( \frac{T_3 r_{UV}^2}{M_{\text{pl}}^2} \right)$ , the leading contribution from

curvature, corresponding to a conformal coupling, may be written

$$V_{\mathcal{R}} = \frac{1}{3}\mu^4 x^2. \quad (2.4)$$

Finally, the structure of the remaining terms has been obtained in [6]:

$$V_{\text{bulk}}(x, \Psi) = \mu^4 \sum_{LM} c_{LM} x^{\delta(L)} f_{LM}(\Psi). \quad (2.5)$$

Here  $LM$  are multi-indices encoding the quantum numbers under the  $SU(2) \times SU(2) \times U(1)$  isometries of  $T^{1,1}$ , the functions  $f_{LM}(\Psi)$  are angular harmonics on  $T^{1,1}$ , and  $c_{LM}$  are constant coefficients. The exponents  $\delta(L)$  have been computed in detail in [6], building on the computation of Kaluza-Klein masses in [27]:

$$\delta = 1, 3/2, 2, \sqrt{28} - 3, 5/2, \sqrt{28} - 5/2, 3, \sqrt{28} - 2, 7/2, \sqrt{28} - 3/2, \dots \quad (2.6)$$

From the viewpoint of the low-energy effective field theory, a term in the potential proportional to  $x^{\delta(L)}$  arises from a Planck-suppressed operator with dimension  $\Delta = \delta(L) + 4$ . In particular, the conformal coupling to curvature corresponds to an operator of dimension six,  $\mathcal{O}_6 = (V_0 + D_0)\phi^2$ .

Two technical remarks are in order. First, for simplicity of presentation we have included higher-order curvature contributions in the list of bulk terms, rather than in a separate category. Second, perturbations of the unwarped metric  $g_{ij}$  lead to terms in the D3-brane potential that were not analyzed in [6], but can be important in some circumstances.<sup>3</sup> We do not implement the angular structure of these terms in full detail, but we have verified that these contributions lead to negligible corrections to our results.

The coefficients  $c_{LM}$  could be computed in principle in a specific realization in which all details of the compactification are available, but in practice must

---

<sup>3</sup>We thank Sohag Gandhi for very helpful discussions of this point.

be treated as unknown parameters. Our approach is to assume that all possible terms are present, with coefficients  $c_{LM}$  of comparable magnitude. Specifically, we will draw the  $c_{LM}$  from a range of statistical distributions and then verify that the (unknown) detailed statistical properties of the  $c_{LM}$  are not important for the inflationary phenomenology, while the overall scale of the  $c_{LM}$  does matter significantly.

We set the overall scale of the bulk contributions by noting that the moduli potential and the remainder of the potential are tied by the requirement that the cosmological constant should be small after brane-antibrane annihilation. With our definition of  $\mu$ , the scaling arguments presented in Appendix A of [26] suggest that in typical KKLT compactifications,  $c_{LM} \sim O(1)$ .

Let us remark that the potential (2.2) is not the most general function on the conifold: it is the most general D3-brane scalar potential on the conifold (within the fairly broad assumptions of [6].) Many terms in (2.2) enjoy correlations that would be absent in a totally general function, and which arise here because certain physical sources, such as fluxes, contribute in correlated ways to different terms in the potential. We defer a full description of the construction of the potential to Appendix A.1.

Finally, we note that in compactifications preserving discrete symmetries that act nontrivially on the throat region, the structure of the D3-brane potential is altered by the exclusion of terms that are odd under the discrete symmetries [26]. Exploring the phenomenology of the corresponding models is an interesting question that is beyond the scope of this work.<sup>4</sup>

---

<sup>4</sup>We thank Daniel Baumann for helpful discussions of this point.

## 2.3 Methodology

To characterize the dynamics of D3-brane inflation in a general potential, we perform a Monte Carlo analysis, numerically evolving more than  $7 \cdot 10^7$  distinct realizations of the model. In this section we explain our recipe for constructing an ensemble of realizations. In §2.3.1, we obtain the equations of motion and introduce the parameters required to specify the potential. In §2.3.2 we describe how we draw the coefficients in the potential from statistical distributions, and in §2.3.3 we indicate how we choose initial conditions.

### 2.3.1 Setup

The inflaton field is characterized by one radial coordinate and five angular coordinates. At the two-derivative level (see §2.4.4 for a discussion of DBI effects), the equations of motion for the homogeneous background are the Klein-Gordon equations obtained from the Lagrangian

$$\mathcal{L} = a^3 \left( \frac{1}{2} T_3 g_{ij} \dot{y}^i \dot{y}^j - V(y) \right), \quad (2.7)$$

where  $a$  is the scale factor, along with the Friedmann and acceleration equations,

$$3H^2 = \frac{1}{2} T_3 g_{ij} \dot{y}^i \dot{y}^j + V(y), \quad (2.8)$$

$$\dot{H} = -\frac{1}{2} T_3 g_{ij} \dot{y}^i \dot{y}^j. \quad (2.9)$$

Here  $y^i$  denotes the six coordinates, dots indicate derivatives with respect to time,  $H \equiv \dot{a}/a$ , and  $g_{ij}$  is the metric on the inflaton field space, which is the conifold.

Three important microphysical parameters are the D3-brane tension,  $T_3$ ; the

length of the throat,  $\phi_{\text{UV}} \equiv r_{\text{UV}} \sqrt{T_3}$ ; and the warp factor at the tip,  $a_0$ . The combination  $2T_3 a_0^4 \equiv D_0$  determines the overall scale of inflation, while  $\phi_{\text{UV}}$  dictates the size of the field space.

The result of [28] gives an upper bound on the inflaton field range,  $\phi_{\text{UV}} < \frac{2M_{\text{pl}}}{\sqrt{N}}$ , with  $N \gg 1$  the D3-brane charge of the warped throat. Consistent with this, we take  $\phi_{\text{UV}} = 0.1$ . Working in units where  $M_{\text{pl}}^{-2} = 8\pi G = 1$  for the remainder, we set the D3-brane tension to be  $T_3 = 10^{-2}$ , and for our analysis of the background evolution, we take  $a_0 = 10^{-3}$ . We have verified that changing these parameters does not substantially alter our results for the homogeneous background. However, changing  $T_3 a_0^4$  — which we accomplish by changing  $a_0$  — does affect the scale of inflation, and hence the normalization of the scalar perturbations. Therefore, in our study of the perturbations in §2.5, we scan over a range of values for  $a_0$ , focusing on values most likely to lead to a WMAP-normalized spectrum [13]. This is a fine-tuning that we will not attempt to quantify, as there is no agreed-upon measure for  $a_0$ .

### 2.3.2 Constructing an ensemble of potentials

In principle the D3-brane potential (2.2) has an infinite number of terms, but for  $x \equiv \frac{r}{r_{\text{UV}}} < 1$  one can truncate (2.2) at some maximum exponent  $\delta = \delta_{\text{max}} \equiv \Delta_{\text{max}} - 4$ . Because of the critical role of the inflaton mass term, truncating to  $\Delta_{\text{max}} < 6$  would fail to capture essential physical properties, so we must have  $\Delta_{\text{max}} \geq 6$ . As  $\Delta$  increases, the number of independent terms grows very rapidly, because there are many angular harmonics  $f_{LM}$  for each  $\Delta$ . Limited by computational power, we truncate the potential at  $\Delta_{\text{max}} = \sqrt{28} + 5/2 \approx 7.8$ . We perform identical

analyses for  $\Delta_{\max} = 6, 7,$  and  $7.8,$  corresponding respectively to 27, 237 and 334 independent terms in the potential, in order to assess whether our results are sensitive to the cutoff.

Many studies of D-brane inflation treat the evolution of the radial position, the volume of a particular four-cycle, and sometimes one angular coordinate, cf. e.g. [7, 8, 9, 18, 19, 20, 21, 22], rather than the full multifield dynamics. Although angular evolution in the framework of [9] has been studied in detail in [23, 24], the focus of these works was the onset of angular instabilities at the end of inflation. We will find that angular evolution before the onset of inflation also plays a critical role.

To understand how our results differ from treatments with fewer dynamical fields, we study the impact of stepwise increases in the number  $N_f$  of evolving fields. We artificially, but self-consistently, freeze  $6 - N_f$  of the angular fields by not imposing the corresponding equations of motion, creating realizations that depend on  $N_f$  variables. While these realizations have less physical meaning than the full potential, they provide some insight into the role of the angular fields. For simplicity we study  $N_f = 1, N_f = 2,$  and the full case  $N_f = 6.$

As we are not assuming that D7-branes wrapping a four-cycle descend into the throat region, we will not model the evolution of the Kähler moduli. Although it would be very interesting (and challenging) to study the cosmological dynamics of Kähler moduli in the bulk, the universality found in the present analysis makes it plausible that additional fields would have little effect on the inflationary phenomenology.

Turning now to the Wilson coefficients  $c_{LM},$  we do not assume a specific com-

pactification, but instead draw the  $c_{LM}$  from a range of statistical distributions. We define the root mean square (rms) size,  $\langle c_{LM}^2 \rangle^{1/2} \equiv Q$ , where the brackets denote the ensemble average, and by assumption<sup>5</sup> the rms size is independent of  $L$  and  $M$ . It is then convenient to write

$$c_{LM} = Q \hat{c}_{LM}, \quad (2.10)$$

and draw the  $\hat{c}_{LM}$  from some distribution  $\mathcal{M}$  that has unit variance but is otherwise arbitrary.

The physical picture is that  $Q$  depends on the distance to the nearest stack of D7-branes effecting Kähler moduli stabilization. The estimates performed in [26] indicate that  $Q \sim \mathcal{O}(1)$  for D7-branes in the upper region of the throat. We anticipate that as the nearest D7-branes are moved farther into the bulk,  $Q$  will diminish to some extent, though we are not aware of a regime in which the bulk contributions are strictly negligible.

If the inflationary phenomenology depended in detail on the nature of  $\mathcal{M}$ , e.g. if the success of inflation depended sensitively on the higher moments of  $\mathcal{M}$ , then no general predictions would be possible. Let us clarify that dependence on the rms size  $Q$  of the  $c_{LM}$ , corresponding to the typical size of the bulk contribution to the inflaton potential, is to be expected and is not problematic. Difficulty would arise if, for example, two distributions with unit variance but with distinct skewness or kurtosis led to disparate predictions.

There are strong motivations for expecting that some statistical properties of the potential will be independent of  $\mathcal{M}$ . For example, if a symmetric  $N \times N$  ma-

---

<sup>5</sup>A strong trend in  $Q$  as a function of  $\Delta$  could change the relative importance of terms with large  $\Delta$ , and hence affect our conclusions about the robustness of the truncation to  $\Delta \leq \Delta_{\max}$ . We are not aware of a well-motivated proposal for such a trend, but it could be worthwhile to investigate this further.

trix has its entries drawn from some distribution with appropriately bounded moments, then in the large  $N$  limit the statistical properties of the eigenvalues are indistinguishable from those obtained from entries drawn from a Gaussian distribution with mean zero [29]. By experimenting with different distributions, we will identify observables which, like the eigenvalue distribution in random matrix theory, are robust against changes in the statistics of the inputs. In practice, for much of our analysis we choose  $\mathcal{M}$  to be a Gaussian distribution with mean zero, and then carefully verify for a range of other distributions that our results receive negligible corrections.

### 2.3.3 Initial conditions

The phase space of initial conditions for a D3-brane in the conifold is 12-dimensional: six dimensions for the initial positions  $x_0, \Psi_0$  and another six dimensions for the initial velocities  $\dot{x}_0, \dot{\Psi}_0$ . A grid-based scan across the full 12-dimensional space would be very computationally intensive even with only a few points along each dimension. Fortunately, five of the six dimensions are angular coordinates on the coset space  $T^{1,1}$ , which has a large isometry group,  $SU(2) \times SU(2) \times U(1)$ . These isometries can be used to reduce the dimensionality of the initial phase space, in the following way. A generic configuration of sources in the compact space will break the isometry group completely, but in a large ensemble of realizations, we expect that there are no preferred regions on  $T^{1,1}$ : the ensemble averages should respect the isometries even though any individual realization breaks the isometries. Thus, without loss of generality we may pick a fixed point  $\Psi_0$  on  $T^{1,1}$  for the initial position. For numerical purposes it is convenient to begin away from the coordinate singularities, so we choose

$\Psi_0$  to be  $\theta_1 = \theta_2 = \phi_1 = \phi_2 = \psi = 1.0$ .

The initial angular velocities  $\dot{\Psi}_0$  are slightly more complicated.<sup>6</sup> To describe a general angular velocity, it suffices to specify the magnitude of the velocity in each  $S^2$  and in the fiber  $S^1$ . For simplicity we focus on velocity in the fiber,  $\dot{\psi}$ , and take the remaining components of the initial velocity to vanish. We expect, and find, similar results for initial velocities in either  $S^2$ , but we postpone a complete scan of the phase space to future work.

We are left with a three-dimensional space of initial configurations spanned by the radial position  $x_0$ , the radial velocity  $\dot{x}_0$ , and the angular velocity  $\dot{\Psi}_0 = \dot{\psi}_0$ . Of course, our evolution occurs in the full 12-dimensional phase space: the simplification applies only to the initial conditions. For a portion of our Monte Carlo analysis, we set  $\dot{x}_0 = \dot{\psi}_0 = 0$ , so that the D3-brane begins at rest. In §2.4.3 we describe the effect of nonvanishing initial velocities.

### 2.3.4 Parameters summarized

To summarize, we fix  $T_3 = 10^{-2}$ ,  $a_0 = 10^{-3}$ , and  $\phi_{UV} = 0.1$  for our analysis of the background evolution. We truncate the D3-brane potential to include contributions from operators with maximum dimension  $\Delta_{\max} = 6, 7$ , and  $7.8$ , and we take  $N_f = 1, 2, 6$  of the D3-brane coordinates to be dynamical fields. The coefficients  $c_{LM}$  have rms size  $Q$ , and the rescaled quantities  $\hat{c}_{LM} = c_{LM}/Q$  are drawn from a distribution  $\mathcal{M}$  that has unit variance. We begin at  $x = x_0 \equiv 0.9$ ,  $\Psi = \Psi_0 \equiv \{\theta_1 = \theta_2 = \phi_1 = \phi_2 = \psi = 1.0\}$ , with arbitrary radial velocity  $\dot{x}_0$ , arbitrary angular velocity  $\dot{\psi}_0$  in the  $\psi$  direction, and all other angular velocities vanishing.

---

<sup>6</sup>We are grateful to Raphael Flauger for helpful discussions of this point.

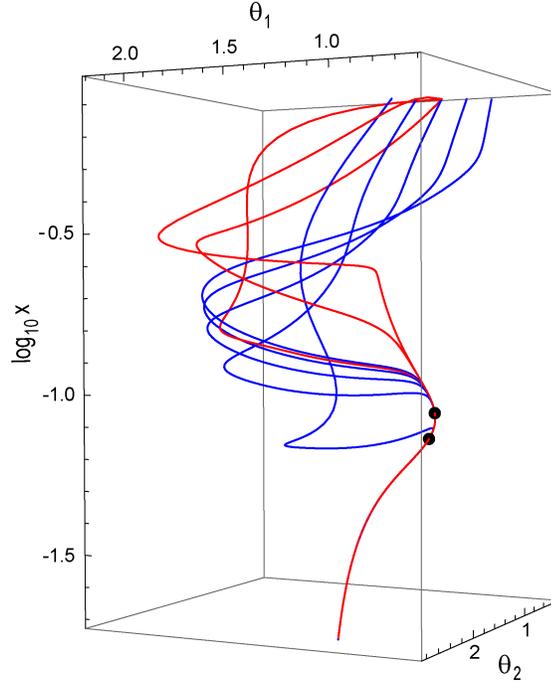


Figure 2.1: Examples of downward-spiraling trajectories for a particular realization of the potential. The black dots mark 60 and 120 e-folds before the end of inflation (7 of the 8 curves shown achieve  $N_e > 120$ ); inflation occurs along an inflection point that is not necessarily parallel to the radial direction. Red curves have nonvanishing initial angular velocities  $\dot{\Psi}_0$ , while blue curves have  $\dot{\Psi}_0 = 0$ .

We would now like to understand how the observables depend on the input parameters  $Q$ ,  $\Delta_{\max}$ ,  $N_f$ , and  $\mathcal{M}$ , and on the initial data  $x_0$ ,  $\dot{x}_0$ ,  $\dot{\psi}_0$ .

## 2.4 Results for the Homogeneous Background

As a first step, we study the evolution of the homogeneous background. In §2.4.1, we show that for fixed initial conditions, the probability of  $N_e$  e-folds of inflation is a power law, and we show that the exponent is robust against

changes in the input parameters  $\Delta_{\max}$ ,  $N_f$ , and  $\mathcal{M}$ . In §2.4.2 we present a simple analytic model that reproduces this power law. We study the effect of varying the initial conditions in §2.4.3, and we discuss DBI inflation in §2.4.4.

### 2.4.1 The probability of inflation

We find it useful to divide possible trajectories into three classes. The D3-brane can be ejected from the throat, reaching  $x > 1$  and leaving the domain of validity of our analysis; it can become trapped in a local or global minimum of the potential; and it can reach the bottom<sup>7</sup> of the throat, triggering the hybrid exit and reheating, after a certain number of e-folds of inflation.

A central question is what fraction of realizations solve the horizon problem by producing  $N_e \geq 60$  e-folds of inflation, and then plausibly transition to the hot Big Bang. We will not model reheating in detail, but we will insist that only e-folds of inflation *that precede a hybrid exit* are counted towards  $N_e$ . That is, false vacuum inflation in a metastable minimum, or slow roll inflation preceded by ejection and unknown dynamics in the bulk, do not contribute to  $N_e$ .

Specifically, for  $k$  distinct trials we define

$$P(N_e > 60) = \frac{\#(N_e > 60)}{\#(N_e > 60) + \#(N_e \leq 60) + \#(\text{ejected}) + \#(\text{trapped})}, \quad (2.11)$$

i.e. trials leading to ejection or trapping are included in the denominator, so that  $P(N_e > 60)$  reflects the probability of  $N_e > 60$  e-folds of inflation preceding a hybrid exit in a general realization.

---

<sup>7</sup>In practice, we define the bottom of the throat to be at  $x = 20a_0$  in order to remain well above the region where the throat rounds off and the singular conifold approximation fails.

We now examine how the ‘success probability’  $P(N_e > 60)$  depends on the input parameters  $Q$ ,  $\Delta_{\max}$ ,  $N_f$ , and  $\mathcal{M}$ . First, Figure 2.2 shows that  $P(N_e > 60)$  depends strongly on  $Q$ , and the optimal value of  $Q$  depends on  $\Delta_{\max}$  and on  $N_f$ . When all six fields are dynamical ( $N_f = 6$ ), the probability of inflation is optimized for  $Q \sim 0.04$ , while for  $N_f = 1$ ,  $Q \sim 1$  can yield sufficient inflation.

To understand this result, we recall that in the presence of a single harmonic contribution to the inflaton potential, after minimization of the angular potential, the radial potential is expulsive [26]. More generally, the bulk contributions to the potential provide the only possibility of counterbalancing the Coulomb and curvature contributions, which both draw the D3-brane towards the tip. For  $Q = 0$ , the Coulomb and curvature contributions are not counterbalanced, and the D3-brane falls quickly towards the tip without driving inflation. For  $Q \sim 1$ , a single harmonic contribution term could marginally balance the inward force; the net effect of 334 such terms then plausibly leads to rapid expulsion from the throat. This result is consistent with our finding that the optimal value of  $Q$  diminishes as the number of terms in the potential increases, as shown in Figure 2.2.

In Figure 2.3 we display a histogram of Monte Carlo trials that give more than 40 e-folds of inflation for  $Q \in [0, 2.0]$  with  $\Delta_{\max} = 7.8$  and all six fields evolving ( $N_f = 6$ ). We find that we can characterize the probability of inflation, for scenarios yielding  $N_e \gg 10$  e-folds, by a function  $P(N_e) = \mathcal{A}(N_e/60)^{-\alpha}$ . On fitting the data in Figure 2.3, we find that  $\alpha = 3.22 \pm 0.07$  and  $\mathcal{A} = 1.7 \cdot 10^{-6}$ .

In Table 2.1 we summarize the power law fits to the probability of inflation as one considers different numbers of fields,  $N_f$ , and different truncations of the potential,  $\Delta_{\max}$ , assuming that  $\mathcal{M}$  is Gaussian, and taking zero initial angular

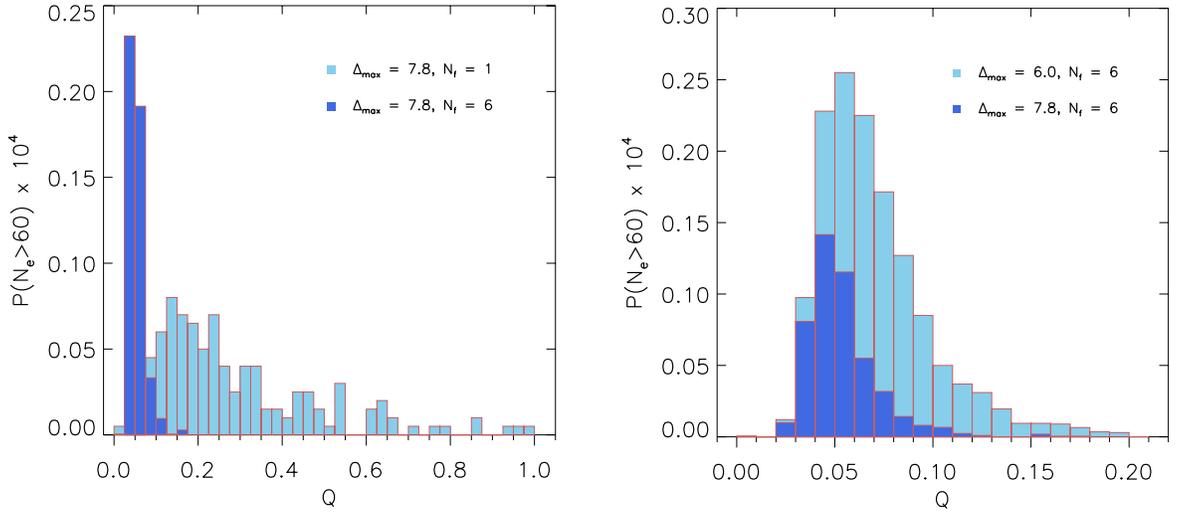


Figure 2.2: The rms value,  $Q$ , of the coefficients  $c_{LM}$  has a significant role in determining whether inflation can occur. [Left panel] The success probability  $P(N_e > 60)$  for two different numbers of fields,  $N_f = 1$  and  $N_f = 6$ , with  $\Delta_{\max} = 7.8$ . [Right panel] The success probability for two different degrees of truncation,  $\Delta_{\max} = 6$  and  $\Delta_{\max} = 7.8$ , with  $N_f = 6$ .

and radial velocities at  $x = 0.9$  and fixed angular position  $\Psi_0$ .

The probability of obtaining 60 e-folds of inflation does not change dramatically if one truncates the potential at  $\Delta_{\max} = 6.0$  or  $\Delta_{\max} = 7.8$ , so our results appear insensitive to the precise placement of the truncation. As the number of fields,  $N_f$ , increases, the range of  $Q$  yielding inflation becomes restricted, but we also find that the probability of achieving inflation within this  $Q$  range increases, cf. Figure 2.2. In fact, the power law fit of the success probability remains fairly consistent as  $N_f$  is varied, provided that one marginalizes over  $Q$ .

Although we have seen that  $P(N_e > 60)$  is highly sensitive to the value of  $Q$ , we find that  $P(N_e > 60)$  has negligible dependence on the shape of the distribution  $\mathcal{M}$  from which the  $\hat{c}_{LM}$  are drawn. Specifically, we have obtained power law

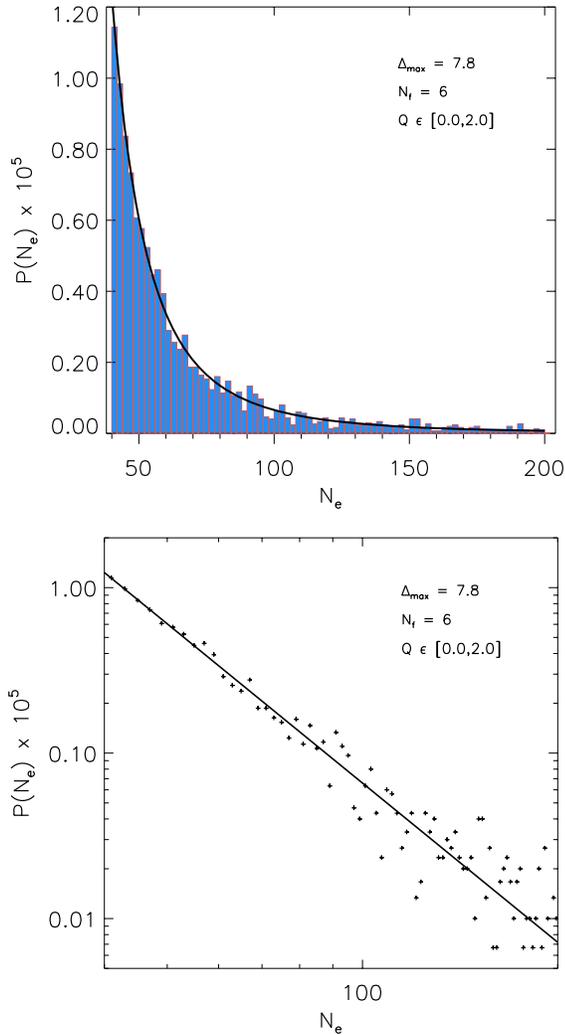


Figure 2.3: The likelihood of  $N_e$  e-folds of inflation as a function of  $N_e$ , for  $\Delta_{\max} = 7.8$  and  $N_f = 6$ . We find  $P(N_e) \propto N_e^{-\alpha}$ , with  $\alpha = 3.22 \pm 0.07$  at the 68% confidence level. The left panel shows the power law fit to the histogram and the right panel shows the same fit on a log-log plot.

fits of the success probability for ensembles in which  $\mathcal{M}$  is a Gaussian, shifted Gaussian, triangular, or uniform distribution. As shown in Table 2.2, we find negligible changes in  $\mathcal{A}$  and  $\alpha$ . We expect that there exist pathological distributions, e.g. with rapidly growing higher moments, that could change our findings, but we are not aware of a microphysical argument for such a distribution.

$N_f$	$\Delta_{\max}$	$\mathcal{A}$	$\alpha$	$P(N_e > 60)$	$P(N_e > 120)$	$P(\text{ej})$	$P(\text{min})$
6	7.8	$1.7 \cdot 10^{-6}$	$3.22 \pm 0.07$	$4.6 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$	0.97	$3.9 \cdot 10^{-3}$
6	7	$3.2 \cdot 10^{-6}$	$3.19 \pm 0.10$	$9.0 \cdot 10^{-5}$	$2.0 \cdot 10^{-5}$	0.96	$8.8 \cdot 10^{-3}$
6	6	$4.3 \cdot 10^{-6}$	$2.85 \pm 0.14$	$1.3 \cdot 10^{-4}$	$3.3 \cdot 10^{-5}$	0.92	$5.4 \cdot 10^{-2}$
2	7.8	$3.0 \cdot 10^{-5}$	$3.30 \pm 0.25$	$9.6 \cdot 10^{-5}$	$2.4 \cdot 10^{-5}$	0.77	$6.7 \cdot 10^{-2}$
2	7	$5.6 \cdot 10^{-5}$	$2.97 \pm 0.17$	$1.8 \cdot 10^{-4}$	$4.9 \cdot 10^{-5}$	0.66	$1.2 \cdot 10^{-1}$
2	6	$1.5 \cdot 10^{-5}$	$2.99 \pm 0.12$	$4.7 \cdot 10^{-4}$	$1.3 \cdot 10^{-4}$	0.51	$1.9 \cdot 10^{-1}$
1	7.8	$2.7 \cdot 10^{-6}$	$2.83 \pm 0.22$	$8.8 \cdot 10^{-5}$	$2.6 \cdot 10^{-5}$	0.36	$5.1 \cdot 10^{-2}$
1	7	$4.3 \cdot 10^{-6}$	$3.03 \pm 0.21$	$1.2 \cdot 10^{-4}$	$2.9 \cdot 10^{-5}$	0.29	$6.1 \cdot 10^{-2}$
1	6	$1.3 \cdot 10^{-5}$	$2.86 \pm 0.15$	$4.2 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$	0.17	$7.4 \cdot 10^{-2}$

Table 2.1: Summary of the power law fits to the success probabilities for various scenarios, with  $Q$  varied over the range  $[0.0, 2.0]$ . We fit the Monte Carlo data for  $N_e > 40$  to a power law of the form  $P(N_e) = \mathcal{A}(N_e/60)^{-\alpha}$ . The two final columns indicate the probabilities that the D3-brane will be ejected from the throat or trapped in a minimum, respectively.

$\mathcal{M}$	$\Delta_{\max}$	$\mathcal{A}$	$\alpha$	$P(N_e > 60)$	$P(N_e > 120)$
Gaussian	7.8	$1.6 \cdot 10^{-6}$	$3.21 \pm 0.14$	$4.6 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$
shifted	7.8	$1.7 \cdot 10^{-6}$	$3.42 \pm 0.10$	$4.7 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$
triangular	7.8	$1.8 \cdot 10^{-6}$	$3.21 \pm 0.09$	$4.9 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$
uniform	7.8	$1.7 \cdot 10^{-6}$	$3.18 \pm 0.08$	$4.8 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$
Gaussian	6	$4.3 \cdot 10^{-6}$	$2.96 \pm 0.09$	$1.3 \cdot 10^{-4}$	$3.4 \cdot 10^{-5}$
shifted	6	$3.7 \cdot 10^{-6}$	$3.03 \pm 0.08$	$1.1 \cdot 10^{-4}$	$2.9 \cdot 10^{-5}$
triangular	6	$4.3 \cdot 10^{-6}$	$2.94 \pm 0.08$	$1.3 \cdot 10^{-4}$	$3.5 \cdot 10^{-5}$
uniform	6	$4.3 \cdot 10^{-6}$	$3.06 \pm 0.07$	$1.3 \cdot 10^{-4}$	$3.6 \cdot 10^{-5}$

Table 2.2: Summary of the power law fits to the success probabilities for various distributions  $\mathcal{M}$ , with  $Q$  varied over the range  $[0.0, 2.0]$ , and for initial positions chosen randomly on  $T^{1,1}$ . We fit the Monte Carlo data for  $N_e > 40$  to a power law of the form  $P(N_e) = \mathcal{A}(N_e/60)^{-\alpha}$ .

## 2.4.2 An analytic explanation of the exponent $\alpha = 3$

In our ensemble of potentials, inflation typically occurs near an approximate inflection point of the potential. We now show that a very simple model of

single-field inflection point inflation, along the lines proposed in [30], predicts  $\alpha = 3$ , in excellent agreement with our numerical results.<sup>8</sup>

An approximate inflection point of a function  $V(\phi)$  of a single field  $\phi$  is a location where  $V'' = 0$  and  $V'$  is small in appropriate units. We choose the origin of  $\phi$  to correspond to the zero of  $V''$ , so that

$$V(\phi) = c_0 + c_1\phi + c_3\phi^3 + \dots, \quad (2.12)$$

with the  $c_i$  being constants. Assuming that the constant term dominates, the number of e-folds of inflation is

$$N_e \approx \frac{c_0}{\sqrt{c_1 c_3}} \quad (2.13)$$

in the regime of interest where the  $c_i$  are small.

The approach suggested in [30] is to obtain the probability of  $N_e$  e-folds of inflation by computing

$$P(N_e) = \int \prod_{i=1}^k d\xi_i F(\xi_1, \dots, \xi_k) \delta(N_e - f(\xi_1, \dots, \xi_k)), \quad (2.14)$$

where the  $\xi_i$  are the parameters of the model,  $f$  is the number of e-folds as a function of these parameters, and  $F$  is a measure on the parameter space. Determining  $F$  from first principles is very subtle, and is beyond the scope of this work. However, to compare to our numerical results involving *relative* probabilities of different numbers of e-folds, we need only use a measure  $F$  that properly represents the measure  $\mathcal{M}$  that we have imposed on the coefficients in our ensemble. At very small values of the  $c_i$ , we can approximate  $\mathcal{M}$  as a constant, and so we take  $F(c_0, c_1, c_3) = 1$ . Thus, we need to evaluate

$$P(N_e) = \int dc_1 dc_3 \delta\left(N_e - \frac{c_0}{\sqrt{c_1 c_3}}\right), \quad (2.15)$$

---

<sup>8</sup>We thank G. Shiu and H. Tye for very helpful discussions of this point.

Performing the integral and again using the smallness of the  $c_i$ , we find

$$P(N_e) \approx -\frac{4c_0^2}{N_e^3} \log(c_0) \quad (2.16)$$

so that  $\alpha = 3$ , which compares very well to our numerical results displayed in Table 2.1.

In the homogeneous background analysis described in §2.4.1, the power in scalar perturbations is unconstrained. However, in §2.5 we will assemble realizations whose scalar perturbations are consistent with the WMAP7 [13] normalization. To compare to the ensemble of §2.5 with fixed scalar power, we must compute

$$P(N_e) = \int \prod_{i=1}^k d\xi_i F(\xi_1, \dots, \xi_k) \delta(N_e - f(\xi_1, \dots, \xi_k)) \delta(A_s^* - A_s(\xi_1, \dots, \xi_k)), \quad (2.17)$$

where  $A_s(\xi_1, \dots, \xi_k)$  is the amplitude of the scalar perturbations as a function of the parameters  $\xi_i$ , and  $A_s^*$  is the central value measured by WMAP7. For the inflection point model (2.12), when the scalar power is fixed as in (2.17), one again finds  $\alpha = 3$ , just as in the case (2.14) with unconstrained scalar power. Moreover, the ensemble of §2.5 with fixed scalar power is consistent with  $\alpha = 3$ , providing a second check of our analytical model.

### 2.4.3 Dependence on initial conditions

By construction, there is no preferred angular position selected by the ensemble of potentials: upon averaging over all possible source locations in the bulk, we recover ensemble average rotational invariance. However, in any particular realization, the potential will be quite different at different angular locations,

so it is meaningful to ask about the effect of varying the initial angular position in a given realization. Moreover, changes in the initial radial position can significantly alter the dynamics. In §2.4.3 we determine the effects of altering the initial radial and angular positions, while the effects of varying the initial velocities are presented in §2.4.3.

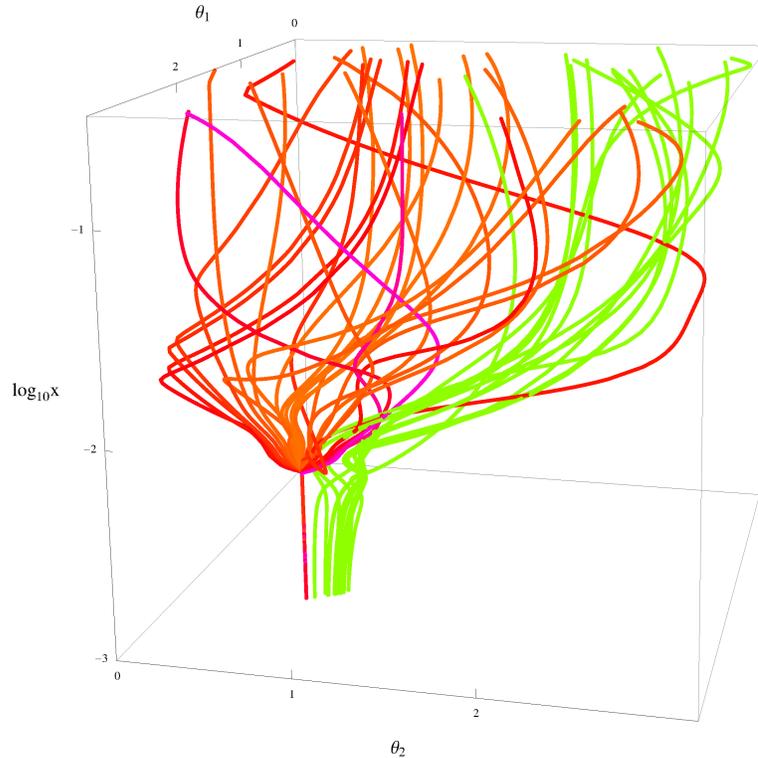


Figure 2.4: Trajectories in the  $\theta_1 - \theta_2$  plane, with  $\log(x)$  vertical, for a fixed potential. Notice the attractor behavior in the angular directions. Green trajectories correspond to  $\approx 5$  e-folds of expansion, while the remaining colors correspond to trajectories with  $\approx 150$  e-folds.

### Dependence on the initial position

Prior works on initial conditions for D-brane inflation have found that in many examples, the inflaton needs to begin with small velocity just above the inflec-

tion point in order to yield substantial inflation. In our ensemble, overshooting is not a problem: in most realizations yielding at least 60 e-folds, it suffices to begin the evolution with small velocity high up in the throat, e.g. at  $x = 0.9$ , while the inflection point is generally in the vicinity of  $x = 0.1$  or even smaller. In fact, increasing the initial radial position typically *increases* the amount of inflation. We suggest that this increase could be due to an increased opportunity to find the inflection point during a prolonged period of radial infall.

The amelioration of the overshoot problem in our ensemble is a reflection of the difference between potentials that are fine-tuned by hand and potentials that are chosen randomly. In the former case, there is a natural tendency to fine-tune the potential to be just flat enough for 60 e-folds of inflation given perfect initial conditions, but no flatter. In contrast, when scanning through the space of possible potentials, one can actually find more robust examples. As we have seen, successful realizations are reasonably common.

The success of inflation has very mild dependence on the initial angular positions: we find that in realizations of the potential that yield more than 60 e-folds of inflation for one set of initial positions, an order-unity fraction of the space of initial angular positions leads to the same outcome. Indeed, we find attractor behavior in the space of initial angles, as illustrated in Figure 2.4.

### **Dependence on the initial velocity**

When the D3-brane begins with a radial velocity of order  $10^{-6}$  of the local limiting speed (cf. §2.4.4), corresponding to an initial kinetic energy that is  $\approx 10\%$  of the potential energy, it strikes the bottom of the throat within a fraction of an

e-fold. However, initial *angular* velocity of the same magnitude has a different effect: the D3-brane is quickly ejected from the throat.

We have found that two distinct causes contribute to this ejection effect: first, D3-branes with large angular velocities can overcome potential barriers in the angular directions and thereby explore a larger fraction of  $T^{1,1}$ , including regions where the potential is strongly expulsive. Second, angular momentum produces a barrier to radial infall, as in standard central force problems. Inward-directed radial velocity and comparably large angular velocity have counterbalancing effects in many cases, suggesting that slow roll inflation could arise in special regions of phase space where the initial velocities are not small, but have compensating effects. We leave this as an interesting question for future work.

#### 2.4.4 The DBI effect

In certain parameter regimes, higher-derivative contributions to the D3-brane kinetic energy can support a phase of DBI inflation [31, 32]. The DBI Lagrangian is

$$\mathcal{L} = a^3 \left( -T(y) \sqrt{1 - \frac{T_3 g_{ij} \dot{y}^i \dot{y}^j}{T(y)}} - V(y) + T(y) \right), \quad (2.18)$$

where in the  $AdS_5 \times T^{1,1}$  approximation with warp factor  $e^A = x$ ,  $T(y) = T_3 x^4$ . DBI inflation can occur if the D3-brane velocity approaches the local limiting speed, i.e. if

$$1 - \frac{T_3 g_{ij} \dot{y}^i \dot{y}^j}{T(y)} \equiv \frac{1}{\gamma^2} \rightarrow 0. \quad (2.19)$$

In our Monte Carlo trials, we did not observe a single example with  $\gamma - 1 > 10^{-8}$ , so the DBI effect was never relevant in our system.

To understand this result, we recall that steep potentials are generically required to accelerate D3-branes to approach the local speed of light, and not every potential that is too steep to support slow roll inflation is actually steep enough to drive DBI inflation in a given warped background.<sup>9</sup> Specifically, DBI inflation requires [32]

$$\left(\frac{V'}{V}\right)^2 \gg \frac{T(y)}{V}, \quad (2.20)$$

so that for a fixed potential, DBI inflation could be achieved by appropriately reducing the background warp factor  $T(y)$ . However, microphysical constraints prevent the warp factor from becoming arbitrarily small: when the infrared scale of a throat becomes small compared to the scale of supersymmetry breaking (due to e.g. fluxes or antibranes in a different region of the compactification), then relevant supersymmetry-breaking perturbations of the throat sourced in the ultraviolet lead to large corrections to the infrared geometry (cf. the discussion in [6]). This constraint enforces

$$V \lesssim 2T_3 a_0^4 \equiv 2T(y)|_{\text{tip}} \leq 2T(y). \quad (2.21)$$

For comparison, the general arguments of [26] concerning the scale of compactification corrections give  $V_{\text{bulk}} \sim T_3 a_0^4$ , and our choice to expand around  $Q = 1$  is consistent with these results. We have taken  $T_3 = 10^{-2}$  and  $a_0 = 10^{-3}$  throughout, so the condition (2.20) could only be satisfied if a rather steep potential arose by chance.

In fact, an additional effect reduces the likelihood of DBI inflation in our analysis. For consistency we have restricted our numerical evolution to the region where the singular conifold approximation is applicable, which excludes the region of greatest warping, the tip of the deformed conifold. In practice, we

---

<sup>9</sup>We thank Enrico Pajer for instructive discussions of this issue.

impose  $x > 20 a_0$ , so that the minimum value of  $T(y)$  explored in our simulations exceeds the global minimum value by a factor of  $20^4$ . It would be very interesting to extend our analysis to the tip region, and as the unperturbed Klebanov-Strassler solution is well understood, it would be straightforward to incorporate purely kinematic corrections involving deviations of the field space metric from that of the singular conifold. However, characterizing the structure of the potential in this region would be much more challenging, and would require understanding the most general non-supersymmetric, perturbed solution for the tip region, along the lines of [6] but departing from the approximately-conformal region.

One could also ask whether, even when the potential is not steep enough to accelerate the D3-brane to near the local speed of light, large initial velocities might still trigger a phase of DBI inflation. We have found that cases with initial radial kinetic energy larger than 10% of the initial potential energy typically strike the bottom without entering the DBI regime. This result is compatible with prior investigations such as [33], which found that exceptionally steep Coulomb potentials, which could arise in cones whose base spaces have extremely small angular volume, are required to produce DBI phases.

In summary, we find that the combination of the mild Coulomb potential in the Klebanov-Strassler throat, and general contributions to the D3-brane potential from moduli stabilization in the bulk, do not suffice to support DBI inflation in the  $AdS_5 \times T^{1,1}$  region. It would be interesting to understand whether DBI inflation arises in the tip region [34, 35].

## 2.5 Towards the Primordial Perturbations

D3-brane inflation generically involves at least six<sup>10</sup> dynamical fields, and the study of the primordial perturbations is rather intricate. Most notably, entropy perturbations can be converted outside the horizon into curvature perturbations [37], provided that the inflaton trajectory bends in a suitable way.

In §2.5.1 we review the prospects for isocurvature-curvature conversion in warped D-brane inflation, following [23]. We will find that although our setup provides an efficient framework for computing multifield effects, there is a wide range of parameter space in which these effects can be neglected. Therefore, in §2.5.2 we restrict our attention to the subset of cases that admit a single-field description, and straightforwardly obtain the CMB observables. A complete analysis of perturbations in the general case is postponed to a future publication.

### 2.5.1 Angular kinetic energy and bending trajectories

Typical trajectories that lead to prolonged inflation begin with relatively rapid angular and radial motion, then gradually spiral down to slow roll inflation along an inflection point, which is not necessarily parallel to the radial direction. Eventually, the D3-brane leaves the inflection point and accelerates, ending slow roll; it then plummets towards the anti-D3-brane, triggering tachyon condensation and annihilation of the brane-antibrane pair.

We will begin by assessing the prevalence of multifield effects in our ensem-

---

<sup>10</sup>In addition to the six D3-brane coordinates, the compactification moduli can also evolve.

ble of realizations. As a first, rough measure of the importance of multiple fields, one can examine the angular kinetic energy during inflation. In Figure 2.5, we show the ratio of angular kinetic energy to radial kinetic energy as a function of the number of e-folds before the end of inflation for selected examples. In realizations yielding  $N_e \approx 60$  e-folds, the ratio of angular kinetic energy to radial kinetic energy is often of order unity when observable scales exit the horizon, but diminishes thereafter. In realizations yielding  $N_e \gtrsim 120$  e-folds, the transients are much diminished, but we still find cases (cf. Figure 2.5) in which the angular kinetic energy is of order the radial kinetic energy, and is approximately constant, throughout inflation. These cases involve slow roll inflation along an inflection point that is not parallel to the radial direction.

Although some degree of bending is commonplace, we do not find that the inflaton trajectory is substantially lengthened as a result of meandering [38] in six dimensions. We find that the total distance  $\ell$  in field space traversed in a realization with six active fields is negligibly larger than that for a realization with only one active field,  $\ell(N_f = 6) \lesssim 1.01\ell(N_f = 1)$ .

Next, we turn to a more precise characterization of multifield contributions to the primordial perturbations. A comprehensive study of multifield effects in D-brane inflation [23] has been performed in the framework of [9], i.e. in terms of explicit embeddings of D7-branes in the Klebanov-Strassler solution. One important lesson of [23] concerns the necessary conditions for isocurvature-curvature conversion at the end of inflation to make a significant contribution to CMB temperature anisotropies. Under fairly general assumptions, this contribution is negligible unless slow roll persists into the deformed conifold region, and the Coulomb potential is subdominant to the moduli potential at the time

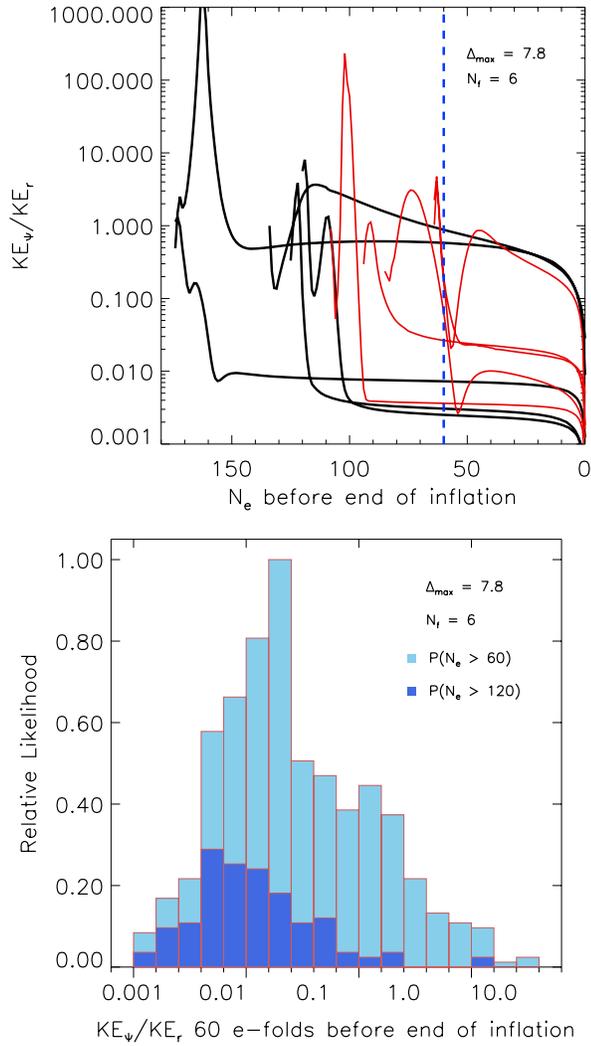


Figure 2.5: The ratio  $KE_\psi/KE_r$  of angular to radial kinetic energies for trials with  $\Delta_{\max} = 7.8$  and  $N_f = 6$ . [Left panel] Evolution of  $KE_\psi/KE_r$  for trials yielding  $60 \leq N_e \lesssim 120$  e-folds of inflation (red lines), and  $120 \leq N_e \lesssim 180$  e-folds (black lines). Notice that in some cases the angular kinetic energy is non-negligible, and nearly constant, in the final 60 e-folds, corresponding to an inflection point trajectory that is not purely radial. [Right panel] Histogram of  $KE_\psi/KE_r$  60 e-folds before the end of inflation for potentials that yield more than 60 (light blue) or more than 120 (dark blue) e-folds of inflation.

of tachyon condensation [23]. Our analysis applies only in the region above the tip of the deformed conifold, so we cannot consistently capture large multifield effects from the *end* of inflation.<sup>11</sup>

A further possibility is that a sharp bend in the trajectory partway through inflation will produce substantial isocurvature-curvature conversion and render invalid a single-field treatment of the perturbations.

To quantify the contributions from additional fields, we calculate  $\eta$  in two components,  $\eta^{\parallel}$  and  $\eta^{\perp}$ , as defined in [39] and [40]. The acceleration of the inflaton parallel to its instantaneous trajectory is captured by  $\eta^{\parallel}$ , while  $\eta^{\perp}$  encodes the rate at which the inflaton trajectory bends perpendicular to itself. Therefore,  $\eta^{\perp}$  is an efficient measure of the role of multiple fields in producing the primordial perturbations [39]. We define as “effectively single-field” a realization in which the stringent cut  $\eta^{\perp}/\eta^{\parallel} < 0.05$  is obeyed for the entirety of the last 60 e-folds.

Interestingly, we do find that in a small fraction of cases, abrupt angular motion occurs after a period of inflation driven by radial motion: the inflaton shifts rapidly from one angular minimum to another, then resumes radially-directed inflation. These examples with  $\eta^{\perp}/\eta^{\parallel} \gg 0.05$  require a full multifield treatment of the perturbations, and there is the intriguing possibility of substantial non-Gaussianity from superhorizon evolution of isocurvature perturbations. We defer consideration of these interesting cases to a dedicated analysis [41].

---

<sup>11</sup>As explained in §2.4.4, incorporating these effects would require an extension of the results of [6] to the tip region, which is beyond the scope of this work.

## 2.5.2 Single-field treatment of the perturbations

We now consider observational constraints on the substantial fraction of examples in which  $\eta^\perp/\eta^\parallel < 0.05$ , so that the primordial perturbations are well-approximated by the single-field result.

We begin by computing the Hubble slow roll parameters,

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = -\frac{d \log H}{dN_e}, \quad (2.22)$$

$$\eta \equiv \epsilon - \frac{1}{2} \frac{d \log \epsilon}{dN_e}, \quad (2.23)$$

We describe the power spectrum of curvature fluctuations using a normalization  $A_s$  and scalar spectral index, or ‘tilt’,  $n_s$ ,  $\Delta_{\mathcal{R}}^2(k) = A_s(k/k_0)^{n_s-1}$ . In terms of  $\eta_{60}$  and  $\epsilon_{60}$ , one has  $A_s = \frac{V_{60}}{24\pi^2\epsilon_{60}}$  and  $n_s - 1 = 2\eta_{60} - 4\epsilon_{60}$ , where the subscripts denote evaluation 60 e-folds before the end of inflation.

The scalar power has significant dependence on the parameter  $D_0 = 2a_0^4 T_3$ , which measures the height of the Coulomb potential. We therefore scan over a range of values of  $D_0$  (in practice, we fix  $T_3$  and scan over  $a_0$ ), and for each successful trial that yields at least 60 e-folds, we compute the slow roll parameters,  $A_s$ , and  $n_s$ . In Figure 2.6, we show scatter plots of  $\epsilon_{60}$  and  $\eta_{60}$  as a function of the maximum number of e-folds.<sup>12</sup>

Notice the paucity of examples with  $N_e \lesssim 120$ . As  $P(N_e) \propto N_e^{-3}$ , a large fraction of trials yield  $N_e$  in this range, but most such examples are excluded by the cut  $\eta^\perp/\eta^\parallel < 0.05$ . This can be understood from Figure 2.5: realizations yielding

<sup>12</sup>Figures 2.6, 2.8 and 2.9 share a set of  $4.9 \cdot 10^6$  Monte Carlo trials at  $\Delta_{\max} = 6$ , out of which 8301 trials yield more than 60 e-folds and 140 also satisfy the WMAP7 constraints on  $A_s$  at  $2\sigma$ . Figure 2.9 additionally includes  $5 \cdot 10^5$  trials at  $\Delta_{\max} = 7.8$ , out of which 750 examples yield more than 60 e-folds and 9 examples also satisfy the constraints on  $A_s$  at  $2\sigma$ . All the data points given in Figures 2.6 and 2.8 obey  $\eta^\perp/\eta^\parallel < 0.05$ , while in Figure 2.9, data points with  $\eta^\perp/\eta^\parallel \geq 0.05$  are included, and indicated by red or purple dots.

$N_e \ll 120$  typically have substantial angular evolution in the final 60 e-folds, so that the single-field approximation is inapplicable.

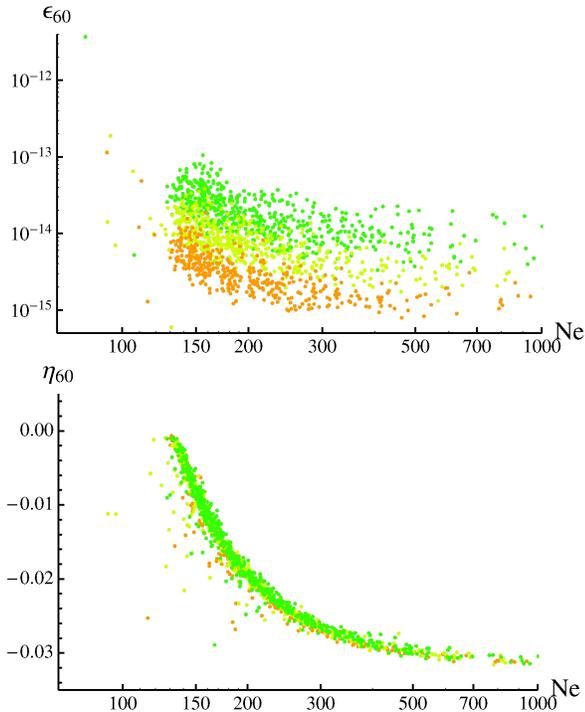


Figure 2.6: Hubble slow roll parameters  $\epsilon_{60}$  and  $\eta_{60}$  as a functions of the maximum number of e-folds. The scatter in  $\epsilon_{60}$  for  $N_e \gtrsim 120$  results from our scan over different values of the inflationary scale, as encoded in  $a_0$ . Notice the absence of corresponding scatter in  $\eta_{60}$ . Color coding:  $-5 < \log_{10} a_0 < -4.75$ , green;  $-4.75 < \log_{10} a_0 < -4.5$ , chartreuse;  $-4.5 < \log_{10} a_0 < -4.25$ , orange. All points shown have trajectories with negligible bending,  $\eta_{\perp} < .05\eta_{\parallel}$ .

We observe that  $\eta_{60}$  is strongly correlated with the number of e-folds. Importantly, when the single-field slow roll approximation is valid, only cases with  $N_e \gtrsim 120$  e-folds are observationally consistent, since it is only for these cases that we have  $V'' < 0$ , ensuring  $n_s < 1$ . This is not surprising (cf. [9], Figure 3.1): for single-field inflation in an approximate inflection point that is flat enough to yield exactly 60 e-folds of inflation, the CMB anisotropies are generated when

the inflaton is above the inflection point, so that the potential is concave up, and hence  $n_s > 1$ . In a corresponding potential that yields 120 e-folds of inflation, observable modes exit the horizon when the inflaton is near to the inflection point (because 60 e-folds have elapsed and 60 e-folds remain), so that for  $\epsilon_{60} \ll 1$ , one has  $n_s \approx 1$ .

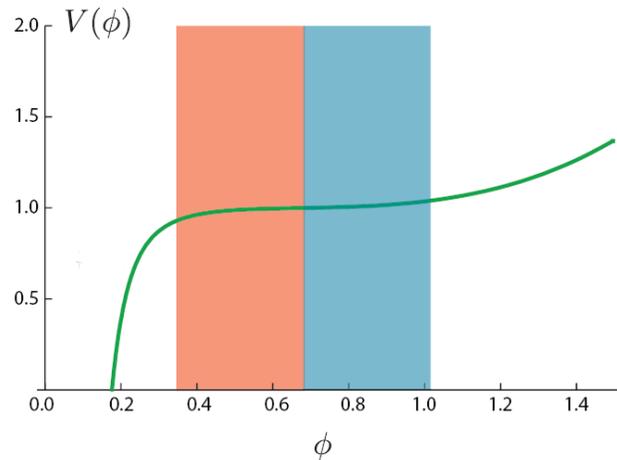


Figure 2.7: An inflection-point potential in one dimension, taken from [9]. The color coding indicates that if the inflaton is above the inflection point 60 e-folds before the end of inflation,  $V'' > 0$  and the scalar power spectrum is blue. A red spectrum is possible if the inflaton has passed the inflection point 60 e-folds before the end of inflation.

The seven-year WMAP (WMAP7) constraints on  $A_s$  and  $n_s$  are  $A_s = (2.43 \pm 0.11) \times 10^{-9}$  and  $n_s = 0.963 \pm 0.014$  at  $k = 0.002 \text{ Mpc}^{-1}$ , at the 68% confidence level [13]. In Figure 2.8 we show a scatter plot of  $A_s$  compared to the measured central value,  $A_s^* = 2.43 \times 10^{-9}$ , indicating cases that are allowed at  $2\sigma$  by the WMAP7 constraint. Notice that there is no fine-tuned choice of  $D_0$ , or of other input parameters, that guarantees a WMAP-normalized spectrum of perturbations: there is significant dispersion in  $\epsilon_{60}$  in our ensemble of inflationary models, with

corresponding dispersion in the scalar power.

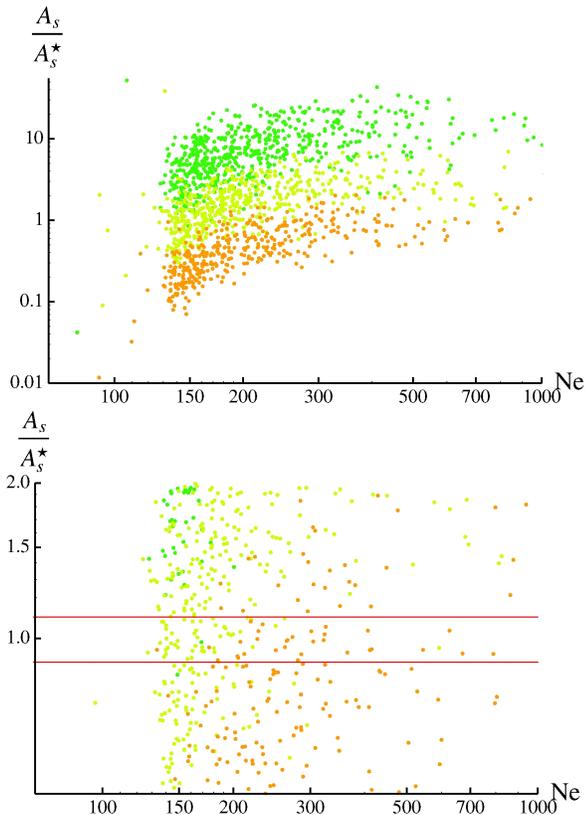


Figure 2.8: Scalar power  $A_s$  compared to the measured central value,  $A_s^* = 2.43 \times 10^{-9}$ , as a function of the total number of e-folds. On the right, we show a zoomed in version. Within the red lines, the scalar power is consistent with WMAP7 at  $2\sigma$ . All points shown have  $\eta_{\perp} < .05\eta_{\parallel}$ , and the color coding is as in Figure 2.6.

Next, for the subset of cases in which  $A_s$  is consistent with experiment, we calculate the corresponding tilt  $n_s$  and tensor-to-scalar ratio  $r$ , displaying the results in Figure 2.9. Evidently, observational constraints on the tilt are readily satisfied for essentially all cases with  $N_e \gtrsim 120$  e-folds, while cases with  $60 \lesssim N_e \lesssim 120$  e-folds solve the horizon and flatness problems but are not consistent with experiment. We stress that this is a straightforward, albeit interesting, consequence of the inflection point form of the potential that is characteristic of

successful realizations in our ensemble.

Within our model there is a microphysical upper bound  $r \leq r_{\max}$  on the tensor-to-scalar ratio resulting from the geometric bound on  $\phi_{\text{UV}}$  [28]. However, in single-field slow roll cases whose scalar perturbations are consistent with WMAP7, we find that  $r \ll r_{\max}$ . It is sometimes argued that small values of  $r$  require substantial fine-tuning. Evaluating the absolute likelihood of inflation in this scenario requires information about a priori measures, and is beyond the scope of this work. Even so, our analysis indicates that the requisite degree of fine-tuning is not extreme: in optimal regions of the parameter space, but without direct fine-tuning of the potential, we find examples consistent with all observations approximately once in  $10^5$  trials, and it is clear from Figure 2.9 that these scenarios have  $r \lesssim 10^{-12}$ .

## 2.6 Conclusions

We have performed a comprehensive Monte Carlo analysis of D3-brane inflation for a general scalar potential on the conifold, obtaining robust predictions in spite of — indeed, arguably because of — the complexity of the inflaton action. Our work builds on recent results [6] that provide the structure of the potential, i.e. a list of possible terms in the potential with undetermined coefficients. Most previous works (cf. e.g. [7, 8, 9, 18, 22, 33]) have treated special configurations in which suitably-aligned D7-branes stabilize the angular positions of the inflationary D3-brane, leading to one-field or two-field dynamics. We have characterized the six-dimensional dynamics of the homogeneous background, without restricting the scalar potential. We have not fine-tuned the potential by

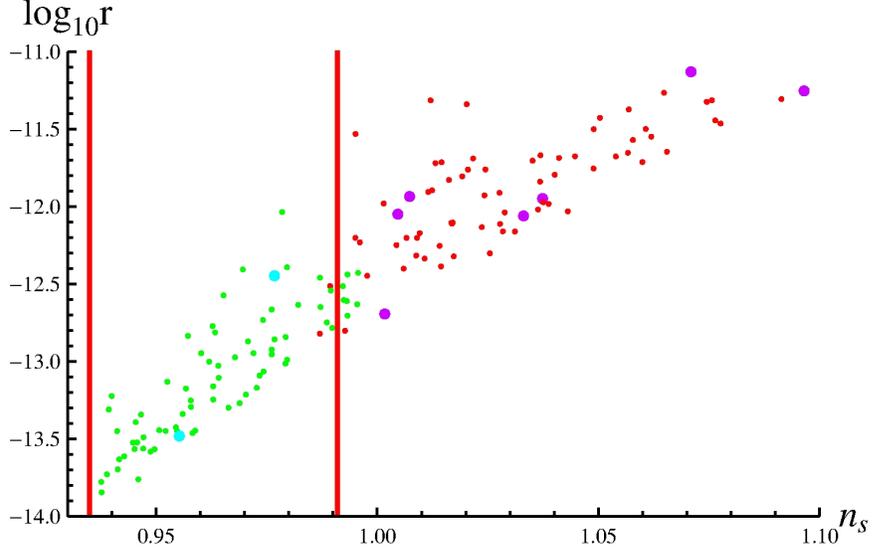


Figure 2.9: Scalar tilt  $n_s$  and tensor-to-scalar ratio  $r$  for cases for which the scalar power  $A_s$  is consistent with WMAP7 at  $2\sigma$ . The red lines indicate the region allowed at  $2\sigma$  by the WMAP7 constraints on  $n_s$ . Green and red dots arise from realizations with  $\Delta_{\max} = 6$ , while cyan and purple dots have  $\Delta_{\max} = 7.8$ . Green and cyan dots correspond to trajectories with negligible bending ( $\eta_{\perp} < .05\eta_{\parallel}$ ), while red and purple dots have  $\eta_{\perp} \geq .05\eta_{\parallel}$  and plausibly require a multifield analysis.

hand, but have instead drawn the coefficients in the scalar potential from suitable distributions, creating an ensemble of potentials, a subset of which led to inflation by chance.

We found that the probability  $P(N_e)$  of  $N_e$  e-folds of inflation is a power law,  $P(N_e) \propto N_e^{-\alpha}$ , with  $\alpha \approx 3$ . The exponent is robust against changes in our truncation of the potential, in the statistical distribution from which the coefficients in the potential are drawn, in the initial conditions, and in the number of dynamical fields. Moreover, we derived  $\alpha = 3$  from a simple analytical model of inflection point inflation. This power-law behavior has significant implications for the prospect of detecting transients arising from the onset of inflation

(cf. [30]): among all histories with at least 60 e-folds, histories with at least 65 e-folds — in which most transients are stretched to unobservable scales — are considerably more likely.

When the inflaton starts at a radial location that is far above the inflection point, angular motion combined with gradual radial infall frequently allow the inflaton to reach the inflection point with a velocity small enough to permit inflation. Moreover, we found attractor behavior in the angular directions: in an order-unity fraction of the space of initial angular positions, the inflaton spirals down to the inflection point. However, large amounts of radial or angular kinetic energy, of order the initial potential energy, are compatible with inflation only in exceptional cases.

DBI inflation did not arise by chance in our ensemble: the potential was never steep enough. It would be interesting to understand whether this finding can be generalized or is an artifact of the limitations of our treatment.

We have obtained the scalar perturbations for the subset of realizations in which a single-field description is applicable throughout the final 60 e-folds, deferring a comprehensive study of the multifield evolution of perturbations to future work. In optimal regions of the parameter space, we found that 60 or more e-folds of inflation arose approximately once in  $10^3$  trials, but because constraints on  $A_s$  and  $n_s$  enforce  $N_e \gtrsim 120$ , observational constraints were satisfied approximately once in  $10^5$  trials. Outside the optimal regions, the chance of inflation diminished rapidly. As we lack a meaningful *a priori* measure on the space of parameters and initial conditions, we have not attempted to quantify the total degree of fine tuning, but our results provide considerable information about relative likelihoods.

In the range of parameters where realizations consistent with observations of the scalar power spectrum are most likely, we found that the tensor-to-scalar ratio obeys  $r \lesssim 10^{-12}$  in all examples allowed by WMAP7, which is much smaller than the maximum allowed by the Lyth bound. Our statements about the perturbations apply only to realizations that are consistently described by slow roll, effectively single-field inflation, which we checked by computing the rate of bending of the trajectory. We anticipate that including more general multifield cases could populate additional regions of the  $n_s - r$  plane.

Our findings have interesting implications beyond the setting of D-brane inflation. The fact that our conclusions are unaffected by the statistical distribution used to generate the coefficients in the potential suggests that in inflationary models in which there are many competing terms in the scalar potential, the details of the individual terms can be less important than the collective structure. A simplification of this form has been previously noted [42] in inflation driven by  $D \gg 1$  fields with quadratic potentials [43], where random matrix theory could be applied directly.<sup>13</sup> Our present results suggest that this emergent simplicity may be more general, and it would be valuable to understand whether a general inflationary model in a field space of dimension  $D \gg 1$  has characteristic properties at large  $D$ .

---

<sup>13</sup>See [45] for recent related work.

## CHAPTER 3

### A STATISTICAL APPROACH TO MULTIFIELD INFLATION: MANY-FIELD PERTURBATIONS BEYOND SLOW ROLL

#### 3.1 Introduction

Inflation [1, 2, 3] provides a superb explanation for the observed spectrum of cosmic microwave background (CMB) anisotropies. The simplest and best-understood models of inflation involve a single field slowly rolling down a relatively flat potential, but more complicated models involving multiple light fields and/or violations of slow roll are arguably more natural, and provide the prospect of distinctive signatures.

Strong theoretical arguments motivate the consideration of inflationary models with many light fields. In theories with spontaneously broken supersymmetry, naturalness suggests that scalars receive masses that are at least of order the gravitino mass. If the inflationary energy is the dominant source of supersymmetry breaking in the early universe, scalars that couple with gravitational strength to this energy will acquire masses of order the inflationary Hubble parameter,  $H$ . Thus, moduli in the theory do not decouple from the inflationary dynamics, and can be light enough to fluctuate.<sup>1</sup> This picture is well-attested in flux compactifications of string theory, which typically include tens or hundreds of moduli with masses clustered around  $H$ . In cases where the moduli potential is computable, one generally finds a complicated, high-dimensional potential energy landscape with structure dictated by the spectrum

---

<sup>1</sup>See [46] for a discussion of some of the cosmological signatures of models with spontaneously broken supersymmetry.

of Planck-suppressed operators in the theory. The nature of inflation in such a potential is an important and urgent problem.

However, despite intensive efforts to understand multifield inflation over the past decade, most analyses of explicit models consider only two light fields. Moreover, even though significant analytical tools have been developed to trace the evolution of primordial perturbations outside the Hubble radius in models violating the slow roll approximation (see §3.2.3 for a brief review of prior results), a large majority of works on the subject do make a slow roll expansion during Hubble exit. An understanding of truly general models involving several fields with masses of order  $H$  and arbitrary dynamics remains necessary.

In this work we study the primordial perturbations produced by inflation in a class of six-field potentials obtained in string theory [6], corresponding to a D3-brane moving in a conifold region of a stabilized compactification.<sup>2</sup> Our approach is statistical: instead of directly fine-tuning the potential to achieve inflation, we draw potentials at random from a well-specified ensemble and study realizations that inflate by chance. We then compute the exact dynamics of the linear perturbations numerically, making no slow roll approximation. To reveal the physical processes underlying the resulting perturbations, we recompute the perturbations using a range of approximate, truncated descriptions that retain different subsets of the entropic modes, and we then cross-correlate the exact and approximate answers. For example, we identify multiple-field contributions to the perturbations by comparing the exact spectrum to the spectrum calculated in a single-field truncation (in which no slow roll approximation is

---

<sup>2</sup>This system was recently studied by Dias, Frazer, and Liddle in [47]. In §3.4.1 we explain how our conclusions are qualitatively different from those of [47]: in contrast to [47], we find that an adiabatic limit is reached during inflation, so that the curvature perturbations can be predicted without modeling the details of reheating.

made).

The organization of this paper is as follows. In §3.2 we describe our method for generating inflationary trajectories and computing the corresponding primordial perturbations. In §3.3 we show that the scalar mass spectrum follows from a simple matrix model, and we assess the incidence and consequences of slow roll violations. In §3.4 we study two-field and many-field contributions to the scalar power spectrum. We conclude in §3.5.

## 3.2 Method

In this section we briefly review warped D-brane inflation, describe how we identify an ensemble of inflating solutions, and then explain how we compute the primordial perturbations.

### 3.2.1 Background evolution in D-brane inflation

In warped D-brane inflation [5], inflation is driven by a D3-brane moving toward an anti-D3-brane in a warped throat region of a flux compactification. The structure of the potential for this configuration has been derived in [6, ?]. The operators in the effective Lagrangian are dictated by the throat geometry, and have been computed explicitly, but the Wilson coefficients are determined by the detailed configuration of sources (fluxes, Euclidean D-branes, etc.) in the bulk of the compactification. Thus, with present knowledge the Wilson coefficients can at best be modeled statistically.

In [48], two of us, in collaboration with N. Agarwal and R. Bean, studied a large number of realizations of the potential, drawing the Wilson coefficients from statistical distributions, and found that although the typical scale of all six scalar masses is  $H$ , accidental cancellations among many terms could nevertheless lead to prolonged inflation. The inflationary trajectory took a characteristic form: the D3-brane initially moved rapidly in the angular directions of the conifold, spiraled down to an inflection point in the potential, and then settled into an inflating phase. It was established in [48] that the inflationary phenomenology has negligible dependence on the detailed form of the statistical distribution of the Wilson coefficients: a sort of universality emerges in this complicated ensemble.<sup>3</sup>

Our method for studying the evolution of the homogeneous background is identical to that of [48], to which we refer for further details. We began with an ensemble of more than 13 million potentials describing a D3-brane in a conifold geometry. For each potential we started with zero kinetic energy at a fixed location in the conifold<sup>4</sup> and evolved the background equations of motion numerically, as detailed in [48]. After discarding trials in which the inflaton became stuck in a local minimum or was ejected from the throat region, we obtained 18731 realizations yielding at least 66 e-folds<sup>5</sup> of expansion followed by a hybrid exit.<sup>6</sup>

---

<sup>3</sup>See [19, 23, 24, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58] for related work on multifield evolution in D-brane inflation.

<sup>4</sup>Trajectories with nonvanishing initial kinetic energy were studied in [48], where it was found that as long as the initial kinetic energy is no larger than the initial potential energy, the effect on the phenomenology is minimal. Similarly, it was shown in [48] that the dependence on the initial radial location is minimal, while the rotational invariance of our ensemble of potentials implies that we do not lose any generality by fixing the initial angular positions.

<sup>5</sup>Although we are interested in requiring at least 60 e-folds of expansion to solve the horizon problem, we insist on a total of 66 e-folds in order to be able to impose Bunch-Davies initial conditions well before the CMB exits the Hubble radius.

<sup>6</sup>In practice we take the end of inflation to occur when the D3-brane reaches a fixed radial location slightly above the tip.

A word of caution about choices of measure is necessary. As shown in [48], the inflationary phenomenology is substantially independent of the statistical distribution from which the Wilson coefficients are drawn, and of the measure taken on the space of homogeneous initial conditions. However, we do not include a measure factor that weights histories according to the amount of expansion, and this must be borne in mind when interpreting our results.

### 3.2.2 Aspects of inflection point inflation

When inflation arises in the ensemble of potentials considered in this work, it does so at an approximate inflection point [7, 8, 9, 48]. Before proceeding we will recall a few elementary properties of single-field inflection point inflation that have significant ramifications for our analysis.

Inflection point inflation begins in a region of field space where the potential is positively curved and evolves to a region where the curvature is negative (see Fig. 3.1). As the size of the inflection point region in Planck units is typically extremely small (see e.g. [48]), the potential slow roll parameters  $\epsilon_V \equiv \frac{1}{2}M_p^2 \left(\frac{V'}{V}\right)^2$  and  $\eta_V \equiv M_p^2 \frac{V''}{V}$  obey  $\epsilon_V \ll |\eta_V|$ , so that  $n_s \approx 1 + 2\eta_V$ . Hence, the scalar power spectrum calculated in the single-field slow roll approximation is initially blue and becomes red. Correspondingly, the tilt measured in the CMB is dictated by the point in field space at which observable modes exit the Hubble radius. Inflection points producing a total of  $N_e \approx 120$  e-folds of inflation have approximately 60 e-folds below the inflection point and 60 e-folds above, so that the primordial spectrum on large angular scales is scale-invariant. Inflection points producing  $N_e > 120$  e-folds have red spectra, while inflection points producing

$N_e < 120$  e-folds have blue spectra and are in tension with observations.

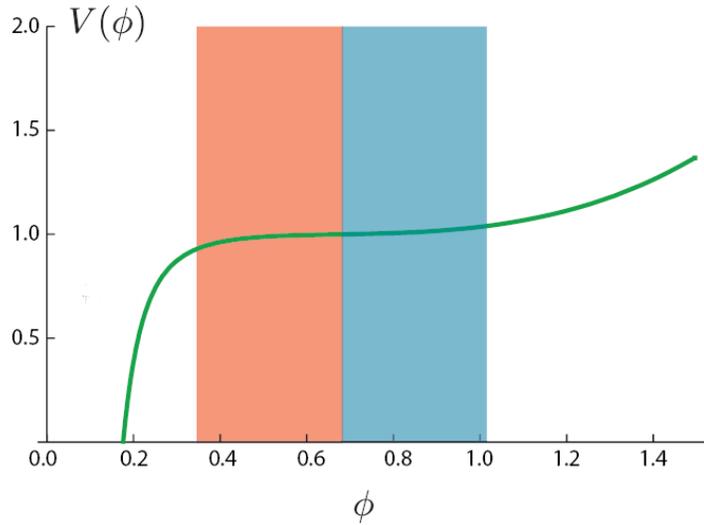


Figure 3.1: A single-field inflection point potential, taken from [9]. If the inflaton is above the inflection point 60 e-folds before the end of inflation,  $V'' > 0$  and the scalar power spectrum is blue, as indicated by the shading. A red spectrum requires that the inflaton has passed the inflection point 60 or more e-folds before the end of inflation.

### 3.2.3 Perturbations in multifield inflation

In this section, we review an efficient approach to computing the primordial perturbations in multifield inflation, following [39] and [59].

## Equations of motion for exact multifield treatment

Our starting point is the action for a collection of scalar fields  $\phi^I$ , endowed with a metric<sup>7</sup>  $G_{IJ}(\phi^K)$  on field space, interacting through a potential  $V(\phi^I)$ , and minimally coupled to gravity:

$$S = \int d^4x \sqrt{-g} \left( \frac{R}{2} - \frac{1}{2} G_{IJ} \nabla_\mu \phi^I \nabla^\mu \phi^J - V(\phi^I) \right) \quad (3.1)$$

(see [93] for the explicit expression for the field space metric in the case of the conifold).

The background metric is assumed to be of the spatially flat Friedmann-Lemaître-Robertson-Walker form

$$ds^2 = -dt^2 + a^2(t) d\vec{x}^2, \quad (3.2)$$

where  $t$  is cosmic time and  $a(t)$  denotes the scale factor. If the fields depend only on  $t$ , the equations of motion take the simple form

$$3H^2 = \frac{1}{2} \dot{\sigma}^2 + V, \quad (3.3)$$

$$\dot{H} = -\frac{1}{2} \dot{\sigma}^2, \quad (3.4)$$

$$\mathcal{D}_t \dot{\phi}^I + 3H \dot{\phi}^I + G^{IJ} V_{,J} = 0, \quad (3.5)$$

where dots denote derivatives with respect to  $t$ ,  $H \equiv \dot{a}/a$  is the Hubble parameter,  $\frac{1}{2} \dot{\sigma}^2 \equiv \frac{1}{2} G_{IJ} \dot{\phi}^I \dot{\phi}^J$  is the kinetic energy of the fields, and, here and in the following,  $\mathcal{D}_t A^I \equiv \dot{A}^I + \Gamma^I_{JK} \dot{\phi}^J A^K$  for a field space vector  $A^I$ .

The dynamics of linear perturbations about this background is governed by the second-order action [39, 59, 60]

$$S_{(2)} = \int dt d^3x a^3 \left( G_{IJ} \mathcal{D}_t Q^I \mathcal{D}_t Q^J - \frac{1}{a^2} G_{IJ} \partial_i Q^I \partial^i Q^J - M_{IJ} Q^I Q^J \right), \quad (3.6)$$

<sup>7</sup>Relativistic motion, corresponding to Dirac-Born-Infeld inflation [31, 32], did not arise in our ensemble [33, 48], so it suffices to consider the two-derivative kinetic term.

where the  $Q^I$  are the field fluctuations in the spatially flat gauge and the mass (squared) matrix is given by

$$M_{IJ} = V_{;IJ} - \mathcal{R}_{IKLJ} \dot{\phi}^K \dot{\phi}^L - \frac{1}{a^3} \mathcal{D}_t \left[ \frac{a^3}{H} \dot{\phi}_I \dot{\phi}_J \right]. \quad (3.7)$$

Here  $V_{;IJ} \equiv V_{,IJ} - \Gamma_{IJ}^K V_{,K}$  is the covariant Hessian,  $\mathcal{R}_{IKLJ}$  is the Riemann tensor associated to the field space metric, and field space indices are raised and lowered using  $G_{IJ}$ . From equation (3.6) one easily deduces the equations of motion for the linear fluctuations (in Fourier space):

$$\mathcal{D}_t \mathcal{D}_t Q^I + 3H \mathcal{D}_t Q^I + \frac{k^2}{a^2} Q^I + M_J^I Q^J = 0. \quad (3.8)$$

### The adiabatic/entropic decomposition

Following [37, 39], it is useful to decompose the field fluctuations into the so-called (instantaneous) adiabatic and entropic perturbations. The adiabatic perturbation is defined as

$$Q_\sigma \equiv e_{\sigma I} Q^I, \quad (3.9)$$

where  $e_\sigma^I \equiv \dot{\phi}^I / \dot{\sigma}$  is the unit vector pointing along the background trajectory in field space, whereas entropic fluctuations represent fluctuations off the background trajectory. The adiabatic fluctuation is directly proportional to the comoving curvature perturbation  $\mathcal{R}$ ,

$$\mathcal{R} = \frac{H}{\dot{\sigma}} Q_\sigma, \quad (3.10)$$

while the genuinely multifield effects are embodied by the entropic fluctuations.

One of the entropic modes plays a distinguished role: this is the ‘first’ entropic fluctuation  $Q_{s_1} \equiv e_{s_1 I} Q^I$ , which is the fluctuation along the direction of

acceleration perpendicular to the background trajectory, where

$$e_{s_1}^I \equiv -\frac{\perp^{IJ} V_{,J}}{\sqrt{\perp^{IJ} V_{,I} V_{,J}}}, \quad (3.11)$$

and  $\perp^{IJ} \equiv G^{IJ} - e_{\sigma}^I e_{\sigma}^J$  is the projection operator on the entropic subspace. The first entropic mode instantaneously couples to the adiabatic perturbation, but the remaining entropic modes do not.

The adiabatic equation of motion can be written in the compact form

$$\ddot{Q}_{\sigma} + 3H\dot{Q}_{\sigma} + \left(\frac{k^2}{a^2} + \mu_{\sigma}^2\right) Q_{\sigma} = (2H\eta_{\perp} Q_{s_1})' - \left(\frac{\dot{H}}{H} + \frac{V_{,\sigma}}{\dot{\sigma}}\right) 2H\eta_{\perp} Q_{s_1}, \quad (3.12)$$

where

$$\eta_{\perp} \equiv -\frac{V_{,s_1}}{H\dot{\sigma}} \quad (3.13)$$

is a very important dimensionless parameter measuring the size of the coupling between the adiabatic mode and the first entropic mode. Here  $V_{,s_1} \equiv e_{s_1}^I V_{,I}$ , and similarly for analogous quantities, and the adiabatic mass (squared)  $\mu_{\sigma}^2$  is given by

$$\frac{\mu_{\sigma}^2}{H^2} \equiv 3\eta - \eta^2 + \epsilon\eta + \dot{\eta}/H, \quad (3.14)$$

with  $\epsilon \equiv -\frac{\dot{H}}{H^2}$  and  $\eta \equiv -\frac{1}{2} \frac{\dot{\epsilon}}{H\epsilon}$ .

A brief overview of existing methods to study cosmological fluctuations in multifield inflation is appropriate at this stage. Several analytical methods have been developed to follow the evolution of perturbations outside the Hubble radius without making any slow roll approximation, including the  $\delta N$  formalism [60, 61, 62, 63], the transfer functions method [64], the gradient expansion method [65, 66, 67, 68], the covariant method [69, 70, 71], and the transport equations method [72, 73, 74, 75]. A number of works incorporate perturbative corrections to the slow roll approximation during the epoch of Hubble exit,

including [39, 76, 77, 78, 79, 80], but exact results are scarce for systems with more than two fluctuating fields. Related investigations of the effects of heavy fields in multifield inflation include [81, 82, 83, 84, 85, 86, 87, 88], while numerical studies for two-field systems include [78, 89, 90].

We stress that an exact numerical treatment is needed, at present, for models involving significant bending around Hubble crossing, which as we will see are common in our ensemble.

### **Methods for computing the perturbations**

For each of the 18731 realizations of inflation yielding at least 66 e-folds of expansion followed by a hybrid exit, we numerically integrated the exact equations of motion (3.8) for the linearized perturbations corresponding to the scale exiting the Hubble radius 60 e-folds before the end of inflation. To relate this result to conceptually simpler models, we also computed the perturbations using an array of approximate descriptions, which we now specify. To distinguish references to these specific models from more general uses of words such as “exact”, “naive”, etc., we will put the model name in a distinctive font.

The naive model simply computes  $\frac{H^2}{8\pi^2\epsilon}$  at Hubble exit,<sup>8</sup> and incorporates neither slow roll violations nor multifield effects. The adiabatic or one-field model makes no slow roll approximation, but discards the effects of all entropic modes: only fluctuations tangent to the trajectory at any given time are retained.

The two-field model keeps the instantaneous adiabatic and first entropic fluc-

---

<sup>8</sup>All quantities described as ‘evaluated at Hubble exit’ in this work are actually averaged from one e-fold before Hubble exit until one e-fold after Hubble exit, in order to smooth quantities that may be rapidly changing.

tuations. For this case the first entropic equation of motion takes the form

$$\ddot{Q}_{s_1} + 3H\dot{Q}_{s_1} + \left(\frac{k^2}{a^2} + \mu_{s_1}^2\right) Q_{s_1} = -2\dot{\sigma}\eta_{\perp}\dot{\mathcal{R}}, \quad (3.15)$$

where

$$\mu_{s_1}^2 \equiv V_{;s_1s_1} - \dot{\sigma}^2 R_{s_1\sigma\sigma s_1} - H^2\eta_{\perp}^2. \quad (3.16)$$

The three-field, four-field, and five-field models (which we will not use in this work) retain additional entropic perturbations along corresponding basis vectors in the decomposition of [39]. We will refer to the six-field model, which includes all five entropic modes and therefore incorporates all the physics of the linear perturbations, as the `exact` model. The scalar power spectra resulting in each model are named in a similar manner, e.g.  $\mathcal{P}_{\text{one}}$  is the scalar power computed in the `one-field` model.

By comparing the results of these approximate descriptions, one can unambiguously identify certain physical effects in the perturbations. Specifically,  $\mathcal{P}_{\text{one}} \neq \mathcal{P}_{\text{naive}}$  signifies violations of slow roll, while  $\mathcal{P}_{\text{exact}} \neq \mathcal{P}_{\text{k}}$  demonstrates that at least  $k + 1$  fields contributed to the perturbations.

Finally, a discussion about numerical implementation of the quantization and evolution of the perturbations is in order. As is well known (see for instance [89, 91, 92]), to determine the late-time power spectrum by numerically evolving the coupled equations (3.8), one cannot simply impose a single choice of initial conditions and solve the system (3.8) once to obtain the six perturbations  $Q^I$ . This would introduce spurious interference terms in the power spectrum between what are supposed to be independent variables. On the contrary, one should identify six variables that are independent deep inside the Hubble radius, each corresponding to an independent set of creation and annihilation operators, and solve the system of equations (3.8) six times, each time imposing

the Bunch-Davies initial conditions for only one of the independent variables, while setting the other variables to zero initially.<sup>9</sup> One then extracts power spectra by summing the relevant quantities over all six runs. This is the numerical analogue of the fact that the various creation and annihilation operators are independent, so that their effects add incoherently (see §3.4.3 for a precise illustration of this method).

Deep inside the Hubble radius, one can neglect the mass matrix in equation (3.6), so that identifying a set of independent variables is equivalent to identifying a set of vielbeins for the field space metric  $G_{IJ}$ , which is easily accomplished numerically. We follow this strategy for the quantization, imposing initial conditions six e-folds before Hubble crossing, while still solving the system of equations (3.8) in the natural coordinate basis on the conifold (i.e.  $r, \theta_1, \phi_1, \theta_2, \phi_2, \psi$  — see [93]), which we find numerically efficient.

The method above applies to the exact model, but some modifications are required for the  $k$ -field models with  $k < 6$ , where by definition only  $k$  independent variables are included. While there is no possible subtlety for the one-field model, for the two-field model for instance, the easiest set of vielbeins consists of the adiabatic and first entropic basis vectors. From a numerical perspective, the kinematical basis of [39] corresponds to a particular set of vielbeins with transparent physical meaning.

---

<sup>9</sup>In the  $k$ -field models with  $k < 6$  described above, we set  $6-k$  entropic modes to zero throughout the evolution, not just initially.

## Terminology for restricted ensembles

For convenience, we now define important subsets of our ensemble. The full ensemble consists of all realizations yielding at least 66 e-folds of inflation. The effectively single-field subset consists of all realizations in which multifield contributions to the scalar power are at most 1% corrections: specifically, we require that

$$\xi_{\text{exact/one}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{one}} - 1| < 0.01. \quad (3.17)$$

The effectively multifield ensemble is the complementary subset, consisting of all cases with  $\xi_{\text{exact/one}} \geq 0.01$ . Similarly, we define

$$\xi_{\text{exact/two}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{two}} - 1|, \quad (3.18)$$

which measures the extent to which a two-field description is insufficient. The set of models with  $\xi_{\text{exact/two}} \geq 0.01$  may be termed ‘effectively many-field’. Finally, the observationally allowed ensemble consists of all realizations satisfying the WMAP7 constraints on the tilt<sup>10</sup> of the scalar power spectrum at  $2\sigma$ .

### 3.3 The scalar mass spectrum and violations of slow roll

Before describing the primordial perturbations, we will first characterize the spectrum of scalar masses. We begin with empirical observations about the mass spectrum (§3.3.1), and then turn to obtaining several of these properties from a random matrix model (§3.3.2). In §3.3.3 we present key consequences of the scalar mass spectrum, focusing on violations of slow roll and the resulting imprint in the power spectrum.

---

<sup>10</sup>See §3.4.4 for a discussion of constraints on the amplitude and running of the scalar power spectrum.

### 3.3.1 Properties of the mass spectrum

For each inflationary trajectory we computed the Hessian matrix  $V^{ij}$  of the scalar potential, in terms of the canonical coordinates constructed at the relevant location on the conifold, at the moment that the CMB exited the Hubble radius. We denote the ordered eigenvalues of the Hessian as  $m_1^2 \leq \dots \leq m_6^2$ , with corresponding eigenvectors<sup>11</sup>  $\psi_1, \dots, \psi_6$ . A histogram of  $m_1^2 \leq \dots \leq m_6^2$  appears in Fig. 3.2.

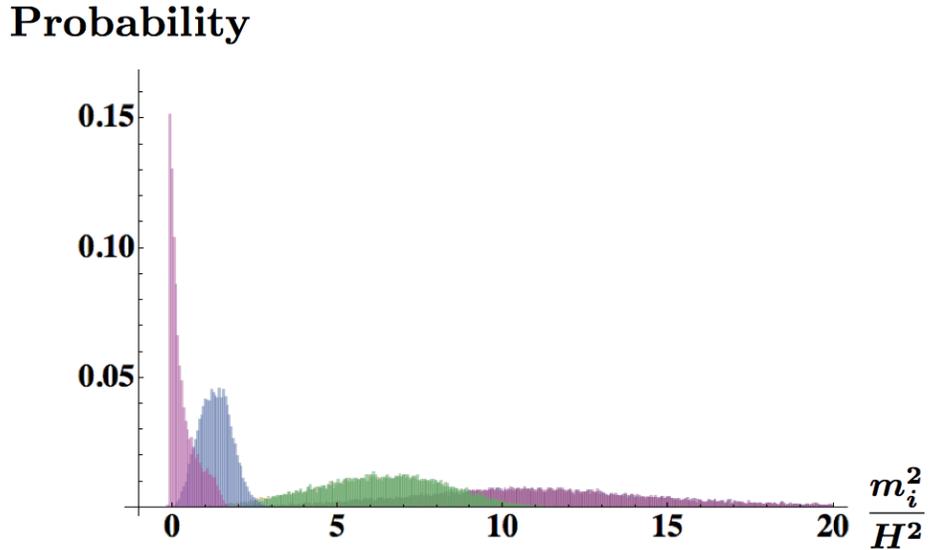


Figure 3.2: The mass spectra of the six scalar fields, in units of  $H^2$ . The left-most peak, which has support at tachyonic values, corresponds to the lightest (adiabatic) field  $\psi_1$ . The next peak corresponds to the second-lightest field  $\psi_2$ . The third peak corresponds to  $\psi_3$  and  $\psi_4$ , which are nearly degenerate in each realization, and the broad final peak similarly corresponds to  $\psi_5$  and  $\psi_6$ .

The mass spectrum of  $\psi_1$  has a sizable spike at mildly tachyonic values:

<sup>11</sup>In general the  $\psi_i$  are complicated combinations of the six natural coordinates on the conifold, though in trajectories that are primarily radial there is significant overlap between  $\psi_1$  and the radial coordinate  $r$ .

$m_1^2 \gtrsim -0.1H^2$ . The lower bound on the mass-squared is a consequence of conditioning on prolonged inflation: at a *generic* point in the field space, not along an inflationary trajectory,  $\psi_1$  would have a much stronger tachyonic instability. In contrast,  $\psi_2$  is almost never tachyonic: we found  $m_2^2 < 0$  in only 4 out of 18731 realizations. Moreover,  $\psi_3$  through  $\psi_6$  were never tachyonic in our ensemble.

An important property of the spectrum is that  $m_3^2 \approx m_4^2$  and  $m_5^2 \approx m_6^2$ , to an accuracy of a few percent. This is not a statistical statement: these eigenvalues are degenerate in each realization, not just after taking the ensemble average. The corresponding histograms are overlaid in Fig. 3.2.

Although we have evaluated the masses at Hubble exit, each mass slowly changes during the course of inflation, in a predictable way. We find that the four heaviest fields become fractionally more massive: in one e-fold around Hubble crossing, there is a change  $\delta m_a^2 \sim .02m_a^2$  for  $a = 3 \dots 6$ . In contrast, the two lightest fields diminish in mass, with  $\delta m_2^2 \sim -.01m_2^2$  and  $\delta m_1^2 \sim -0.18|m_1|^2$ .

The masses that are relevant for determining the evolution of the fluctuations consist of a standard part coming from the potential, namely the  $V_{,II}$  term in equation (3.7), as well as kinematic contributions involving time derivatives of background quantities, corresponding to the remaining terms in equation (3.7). We find that the kinematic contributions introduce corrections at the level of one part in  $10^4$ , and so can be neglected for practical purposes.

### 3.3.2 A random matrix model for the masses

We will now show that distinctive qualitative features of the mass spectrum can be understood using random matrix theory. The framework for this analysis is ‘random supergravity’, by which we mean an ensemble of four-dimensional  $\mathcal{N} = 1$  supergravity theories whose Kähler potential and superpotential are random functions, in a sense made precise in [94]. A matrix model that is slightly simpler than that of [94] will suffice for our purposes: we take the Hessian matrix  $\mathcal{H}$  to be of the form<sup>12</sup>

$$\mathcal{H} = \begin{pmatrix} A\bar{A} + B\bar{B} & C \\ \bar{C} & \bar{A}A + \bar{B}B \end{pmatrix}, \quad (3.19)$$

up to a shift proportional to the identity matrix, where  $A$ ,  $B$ , and  $C$  are  $3 \times 3$  complex symmetric matrices. We take the entries of  $A$ ,  $B$ , and  $C$  to be random complex numbers drawn from a normal distribution; as explained at length in [94], the spectrum of  $\mathcal{H}$  is essentially independent of the statistical properties of the matrix entries. For  $C = 0$  the spectrum is positive-definite and doubly degenerate, but  $C \neq 0$  breaks the degeneracies and permits negative eigenvalues.

By requiring inflation, one has effectively imposed a lower bound on  $m_1^2$ , which clearly affects the empirical spectrum shown in Fig. 3.2. We should therefore impose a corresponding restriction in the matrix model: we take  $m_1^2 \geq -0.1H^2$ . (The overall scale is arbitrary in the matrix model, and we set our units by matching the right tail of the rightmost peak in Fig. 3.2.) In Fig. 3.3 we show the results of simulations of the spectrum in this simple matrix model. By adjusting parameters — such as the weights assigned to  $A$ ,  $B$ , and  $C$ , or the relative variance of their entries — one can achieve reasonably good quantita-

---

<sup>12</sup>Cf. [94] for an explanation of the relationship between this model and the full Hessian matrix in supergravity. We thank T. Wrase for helpful discussions of this point.

tive modeling of the empirical spectrum, but it is not clear that this is physically meaningful. We have instead presented the results for the simplest case, with equal weights for all matrices, and equal variance for all entries, for which the qualitative agreement is already surprisingly good. Notice that the eigenvalues are approximately pairwise degenerate, which matches the empirical result for  $m_3^2 \approx m_4^2$  and  $m_5^2 \approx m_6^2$  but is a less accurate model of  $m_1^2$  and  $m_2^2$ . In contrast to the analysis of [94], the matrices in question are not large: they are  $3 \times 3$ , so that large  $N$  arguments (where  $N$  is the size of the matrix) are marginal at best. However, for  $C = 0$  and any  $N$ , the spectrum of  $\mathcal{H}$  is exactly doubly degenerate, and the approximate degeneracies in the mass matrix with  $C \neq 0$  hold at finite  $N$ .

### Probability

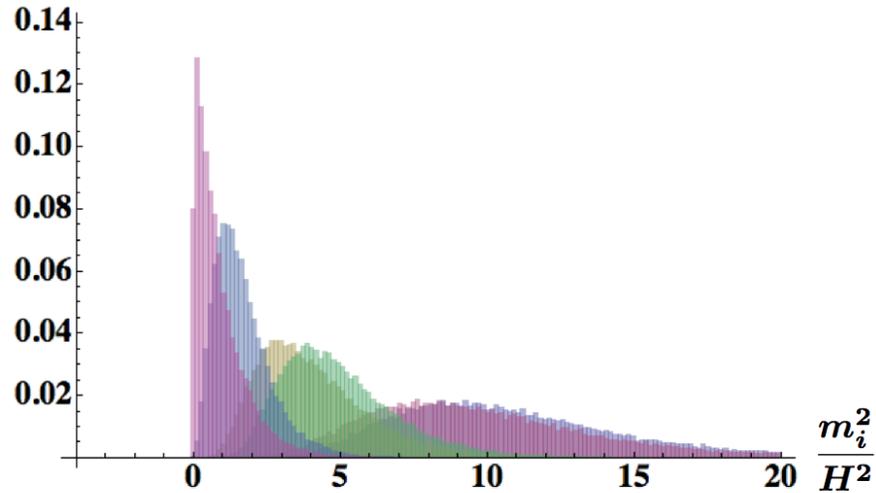


Figure 3.3: The mass spectrum simulated in the matrix model given in equation (3.19), cf. [94]. Comparing to Fig. 3.2, we see that the model is in good qualitative agreement with the results of simulations in the ensemble of inflaton potentials.

Another key feature of the empirical spectrum that is reproduced in the matrix model is the sharp left edge in the probability density for  $m_1^2$ . The re-

quirement of  $N_e \gtrsim 66$  e-folds of inflation effectively imposes a lower bound  $m_1^2 \gtrsim -0.1H^2$ . The steep pileup of the probability density near this edge is a characteristic feature of eigenvalue spectra of ensembles of random matrices subject to constraints on the smallest eigenvalues.

We conclude that qualitative features of the mass spectrum arising from the scalar potential on the conifold can be reproduced in the very general random matrix model of [94], or its simplified version (3.19), which governs any  $\mathcal{N} = 1$  supergravity theory for which the Kähler potential and superpotential are random functions to good approximation. Correspondingly, these features are plausibly properties of a general random supergravity theory, not consequences of the particular conifold setting of the present work. This strongly motivates using kindred matrix models to study much more general many-field inflationary scenarios.

To make a further observation about the scope of the matrix model, we briefly recall the method used in [6] to compute the scalar potential for a D3-brane on the conifold. The computation of [6] amounted to determining the most general solutions of ten-dimensional supergravity, in expansion around the Klebanov-Strassler solution, with certain asymptotics. As such, these solutions incorporate in full detail the structure of a conifold geometry attached to a stabilized compactification, but — being intrinsically ten-dimensional — lack any manifest four-dimensional  $\mathcal{N} = 1$  supersymmetry. In particular, the Kähler potential and superpotential of the four-dimensional theory arising upon dimensional reduction were not obtained directly in [6]. We therefore find it remarkable that a random matrix model based on four-dimensional random supergravity is in excellent agreement with our calculation of the mass spectrum

arising from the ten-dimensional results of [6].

A further quantity of significant interest is the number of scalar fields that are light enough to fluctuate during inflation. We will now show that an efficient measure of the number of fluctuating scalars is a weighted average that takes into account the reduced contributions of fields with masses  $m \rightarrow 3/2H$  that barely fluctuate. To determine the proper weighting, we consider the equation of motion for the perturbations. The canonically normalized perturbation  $v = aQ$  of a test scalar field of mass  $m$  in de Sitter spacetime obeys the equation of motion, in Fourier space,

$$v_k'' + \left( k^2 - \frac{1}{\tau^2} \left( \nu^2 - \frac{1}{4} \right) \right) v_k = 0, \quad (3.20)$$

where

$$\nu^2 \equiv \frac{9}{4} - \frac{m^2}{H^2}, \quad (3.21)$$

and ' denotes a derivative with respect to conformal time  $\tau$ . The solution of this equation with the appropriate Bunch-Davies behavior inside the Hubble radius reads

$$v_k(\tau) = \frac{\sqrt{\pi}}{2} e^{i(\nu+1/2)\pi/2} \sqrt{-\tau} H_\nu^{(1)}(-k\tau), \quad (3.22)$$

where  $H_\nu^{(1)}$  is the Hankel function of the first kind, of order  $\nu$ , and we use the convention

$$\nu = \sqrt{\nu^2} \quad \text{when } \nu^2 > 0, \quad (3.23)$$

$$\nu = i\sqrt{-\nu^2} \quad \text{when } \nu^2 < 0. \quad (3.24)$$

The power spectrum of fluctuations at Hubble crossing, i.e. for  $-k\tau_\star = 1$ , can then be computed as

$$\mathcal{P}_{Q_k, m^2}(\tau_\star) \equiv \frac{k^3}{2\pi^2} |Q_k(\tau_\star)|^2. \quad (3.25)$$

For each of our inflationary realizations, one can sum the six contributions of the form (3.25) corresponding to the six eigenvalues of the Hessian matrix determined in §3.3.1. We thus obtain an analytical estimate of the total power spectrum of scalar field fluctuations at Hubble exit,

$$\mathcal{P}_{\text{tot}} = \sum_{i=1}^6 \mathcal{P}_{Q_k, m_i^2}(\tau_\star). \quad (3.26)$$

In realizations with small bending of the trajectory (see §3.4.2), the estimate (3.26) agrees extremely well with a direct numerical calculation of the total power spectrum of field fluctuations at Hubble exit, with a precision of order 0.01%. In more general models, the estimate (3.26) is only qualitatively correct, and our full numerical treatment is necessary (see for instance §3.4.3).

The effective number of fluctuating fields,  $n_f$ , can then be defined by comparing the estimate (3.26) to the corresponding expression for a *massless* scalar field:

$$n_f \equiv \mathcal{P}_{\text{tot}} / \mathcal{P}_{Q_k, m^2=0}(\tau_\star). \quad (3.27)$$

In Fig. 3.4 we show a histogram of  $n_f$  in our ensemble of inflationary realizations. Remarkably, the histogram is sharply peaked, and  $n_f$  falls between 2 and 3 in about 99% of our realizations.

We caution the reader that the number of fields that are light enough to fluctuate,  $n_f$ , is in general different from the number of fields  $n_{\mathcal{R}}$  that contribute to the curvature perturbation: for example, if the background trajectory is straight, then entropic perturbations are not converted to curvature perturbations, and  $n_{\mathcal{R}} = 1$  for any  $n_f$ . The calculation above concerns the mass spectrum, and hence determines the statistical distribution of  $n_f$ . We will determine the distribution of  $n_{\mathcal{R}}$  in §3.4.3.

## Probability

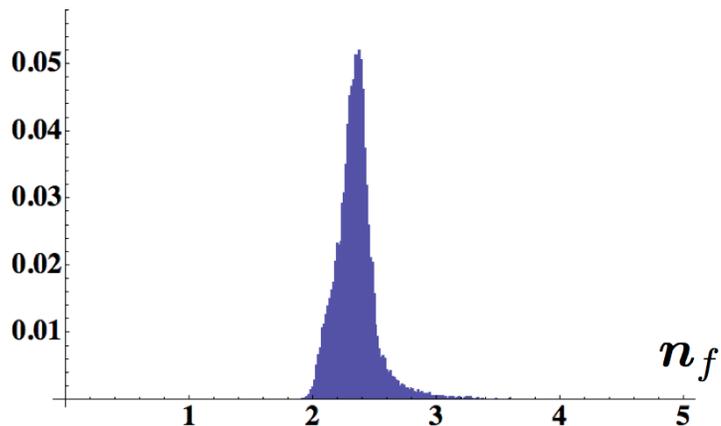


Figure 3.4: Histogram showing the relative probability of the number  $n_f$  of scalar fields that are light enough to fluctuate during inflation, cf. equation (3.27).

### 3.3.3 Violations of slow roll

A pivotal property of the scalar mass spectrum discussed in the preceding section is that the lightest field  $\psi_1$  has a mass  $m_1^2 \sim H^2$  in a large fraction of realizations yielding  $N_e \gtrsim 66$  e-folds of inflation. Thus, the slow roll approximation is only marginally applicable.

In Fig. 3.5 we show that the mass-squared of the adiabatic direction, evaluated 60 e-folds before the end of inflation, has a clear dependence on the total number of e-folds,  $N_e$ . Realizations with  $N_e \ll 100$  have  $m_\sigma^2/H^2 \gtrsim 1$ , while realizations with  $N_e \gg 100$  have<sup>13</sup>  $m_\sigma^2 \approx -0.1H^2$ . This is simply a consequence of the fact that inflation is occurring at an inflection point. Realizations yielding  $N_e < 120$  have the 60 e-fold mark above the inflection point, where the curva-

<sup>13</sup>It is straightforward to show analytically that 60 e-folds before the end of inflation in a single-field inflection point model yielding  $N_e \gg 100$ , the inflaton mass obeys  $m^2 \approx -\frac{H^2}{10}$ , in excellent agreement with our simulations.

ture of the potential is positive, while realizations yielding more inflation have the 60 e-fold mark slightly below the inflection point. See §3.2.2.

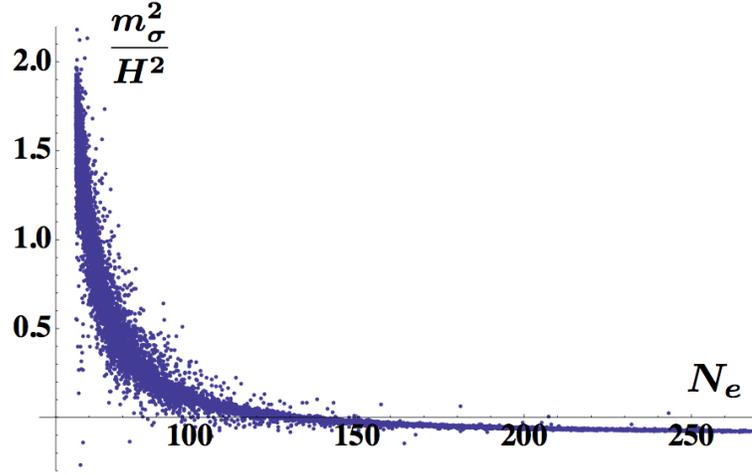


Figure 3.5: The mass-squared of the adiabatic fluctuation in units of  $H$ , evaluated 60 e-folds before the end of inflation, versus the total number of e-folds,  $N_e$ .

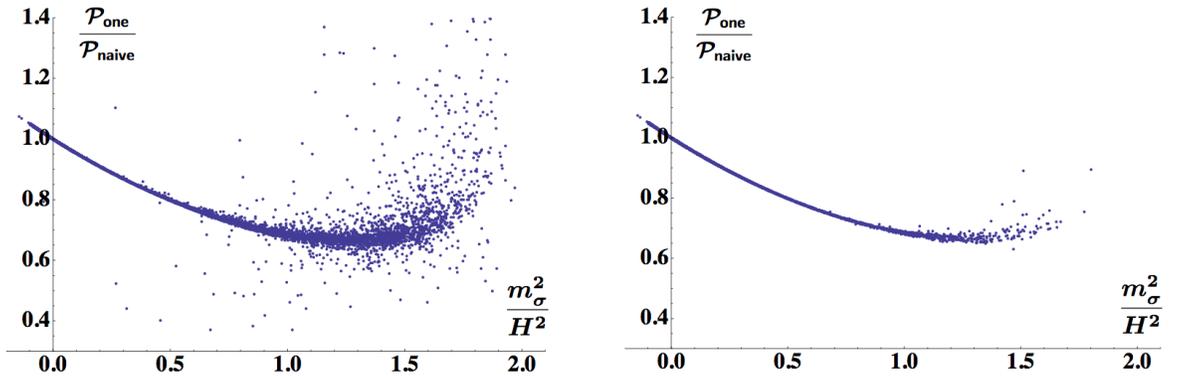


Figure 3.6: The ratio of the one-field power to the naive power, versus the adiabatic mass-squared in units of  $H$ . The left panel shows the entire ensemble, while the right panel is restricted to effectively single-field realizations (see §3.2.3).

Violations of slow roll provide significant corrections to the scalar power spectrum in our ensemble. In Fig. 3.6 we show the ratio of the one-field power

to the naive power, as a function of  $m_\sigma^2/H^2$ . We learn that increasing the mass of the adiabatic direction decreases the scalar power in a predictable manner.

The effects of slow roll violations on the tilt for effectively single-field models are shown in Fig. 3.7. Except for the handful of cases with  $m_\sigma^2/H^2 \gtrsim 1.3$ , the exact tilt is more blue than the naive tilt, by an amount that is strongly correlated with  $m_\sigma^2/H^2$ . Thus, slow roll violations tend to shift the spectrum to be slightly more blue.

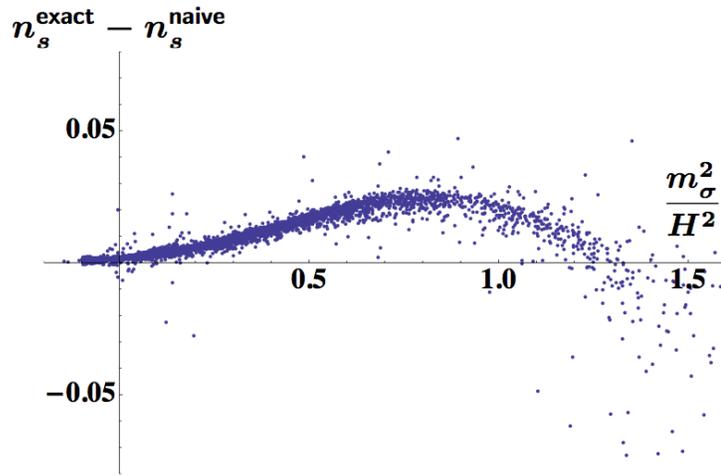


Figure 3.7: The difference between the exact and naive spectral tilts, versus the adiabatic mass-squared in units of  $H$ , for effectively single-field realizations (see §3.2.3).

### 3.4 Multifield effects

We now turn to our primary objective, the characterization of multifield contributions to the scalar power spectrum. To begin, in §3.4.1 we discuss the degree to which multifield inflation in our ensemble is predictive. Next, in §3.4.2 we quantify the degree of bending of the trajectory. In §3.4.3 we characterize the

frequency with which multifield and many-field effects arise, and in §3.4.4 we describe the consequences of imposing observational constraints on the tilt. We discuss the prospect of observable non-Gaussianities in §3.4.5.

### 3.4.1 Decay of entropic perturbations

Entropic perturbations are essential for super-Hubble evolution of the curvature perturbation: if all entropic modes become massive and decay at some point after Hubble exit, the curvature perturbation becomes constant, and one says that an *adiabatic limit* has been reached (see [95, 96, 97, 98] for recent discussions). Provided that an adiabatic limit is reached during the inflating phase, one can predict the curvature perturbation on observable angular scales without knowing the details of the end of inflation or of reheating. Conversely, failure to reach an adiabatic limit makes predictions contingent on an understanding of reheating.

The approach to an adiabatic limit clearly depends on the masses of the entropic modes: modes with  $m \gtrsim H$  decay quickly outside the Hubble radius, and those with  $m > \frac{3}{2}H$  do not oscillate at all. One might a priori expect that in our ensemble all six fields have masses of order  $H$ , with a substantial likelihood that more than one field is tachyonic. Our findings from §3.3.2 differ from this expectation in important details: there is at most<sup>14</sup> *one* tachyonic instability, the second-lightest field  $\psi_2$  has  $m_2^2 \sim H^2$ , and the heavier fields  $\psi_3 \dots \psi_6$  have masses considerably larger than  $H$ . Correspondingly, we expect that the entropic modes will decay during the course of inflation, so that an adiabatic

---

<sup>14</sup>Strictly speaking, we found two tachyons in a negligible fraction of trials, in 4 out of 18731 examples.

limit is reached before the end of inflation.

To see that this is the case, we examined a subset of realizations and verified that the total power of the five entropic perturbations decays exponentially. In 60 examples, the power in entropic modes at the end of inflation was never larger than  $10^{-10}$  times the adiabatic power. In Fig. 3.8 we show the characteristic decaying behavior for 19 examples.

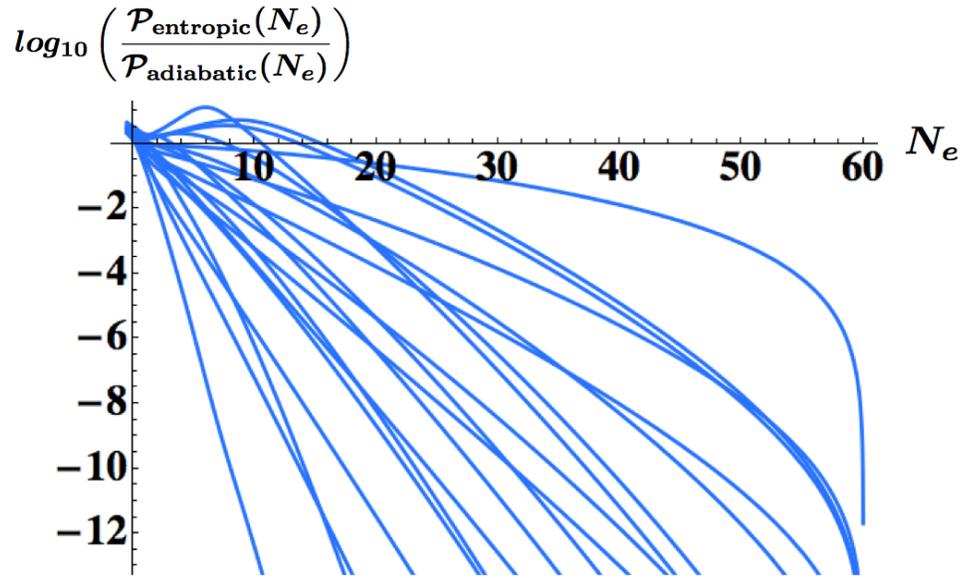


Figure 3.8: The curves show the rapid decay of the total power of the five entropic modes over the course of inflation, for 19 realizations. The decaying behavior depicted is representative of the ensemble: the lightest entropic mode generically has  $m^2 \gtrsim H^2$ , and was tachyonic at Hubble crossing in a negligible fraction of trials.

One possible subtlety is worth mentioning. In the argument above we have assumed that the lightest field  $\psi_1$  corresponds to the adiabatic direction  $\phi_\sigma$ , so that having  $m_2^2 \sim H^2$  and  $m_3^2, m_4^2, m_5^2, m_6^2 \gg m_2^2$  ensures that all five entropic modes are massive enough to decay quickly. There is good evidence for this identification: the histograms of  $m_1^2$  and of  $m_\sigma^2$  are nearly indistinguishable, but are quite

different from those of the remaining fields.

Somewhat different results concerning the approach to the adiabatic limit in a similar system have been reported<sup>15</sup> by Dias, Frazer, and Liddle [47], who find that the angular fields have masses that are very small compared to  $H$ . Correspondingly, no adiabatic limit is reached, and there is a problematic persistence of entropic modes. Based on the analysis of [6] and [48], or more generally on the grounds of naturalness in a theory with spontaneously broken supersymmetry, we would expect masses of order  $H$  for all six fields (modulo a suppression of the mass of the adiabatic direction resulting from conditioning on prolonged inflation), which is consistent with the results reported here, but does not appear consistent with [47].

### 3.4.2 Bending of the trajectory

Because the entropic perturbations generally decay rapidly outside the Hubble radius, they are relevant for the late-time curvature perturbation only if the inflationary trajectory bends shortly after they exit the Hubble radius, so that the entropic perturbations are rapidly converted to curvature perturbations. We therefore turn to characterizing the incidence of bending trajectories in our ensemble.

A very useful measure of turning is the parameter  $\eta_{\perp}$  defined in equation (3.13), which measures the acceleration of the trajectory transverse to the instantaneous velocity. Analytic methods that make a slow roll approximation around the time of Hubble crossing generally require  $\eta_{\perp} \ll 1$ , which can be related to the

---

<sup>15</sup>We thank Mafalda Dias and Jonathan Frazer for discussions of this point.

‘slow-turn approximation’ discussed in [90]. Correspondingly, effects requiring  $\eta_{\perp} \gtrsim 1$  at Hubble crossing are only partially understood. Our treatment makes no approximation, and indeed  $\eta_{\perp}$  can be quite large: see Fig. 3.9, which shows the evolution of  $\eta_{\perp}$  in the first 15 e-folds of inflation in a handful of representative examples. The gradual decay of  $\eta_{\perp}$  can be understood from the properties of the evolution near the inflection point: while the inflaton is spiraling down to the inflection point, turning is generic, but in the immediate vicinity of the inflection point, where prolonged inflation occurs, the trajectory is quite straight.

In the left panel of Fig. 3.10 we plot  $\eta_{\perp}$ , evaluated at Hubble crossing, versus the total number of e-folds,  $N_e$ . Evidently,  $\eta_{\perp}$  at Hubble crossing is quite small in scenarios giving rise to  $N_e \gg 60$  e-folds of inflation, but is significant in scenarios yielding less inflation. This is consistent with the picture described above in which there are large transient contributions to  $\eta_{\perp}$  that decay after prolonged inflation.

Another practical measure of the amount of turning is the ‘total turn’  $\Delta\theta$ , which we define to be the integral of  $\eta_{\perp}$  from six e-folds before Hubble crossing until the end of inflation:  $\Delta\theta \equiv \int_{HC-6}^{end} dN_e \eta_{\perp}$ .<sup>16</sup> From the right panel of Fig. 3.10 we see that the total turn is quite large in cases with  $N_e \lesssim 80$  e-folds of inflation. Notice that even in cases with arbitrarily many e-folds, a total turn of order  $\Delta\theta \sim 1/2$  remains likely. This is an effect from the end of inflation: as the inflaton falls off the inflection point, its trajectory very often bends (see [45] for a discussion of related points). We have verified this by checking that the restricted integral  $\int_{HC-6}^{end-10} dN_e \eta_{\perp}$  is small in cases with  $N_e \gg 60$  e-folds, even though  $\int_{HC-6}^{end} dN_e \eta_{\perp}$  is not.

---

<sup>16</sup>Turns occurring shortly before Hubble crossing can have an effect on the perturbations, and we therefore define the total turn as the integral of  $\eta_{\perp}$  from six e-folds before Hubble crossing to the end of inflation.

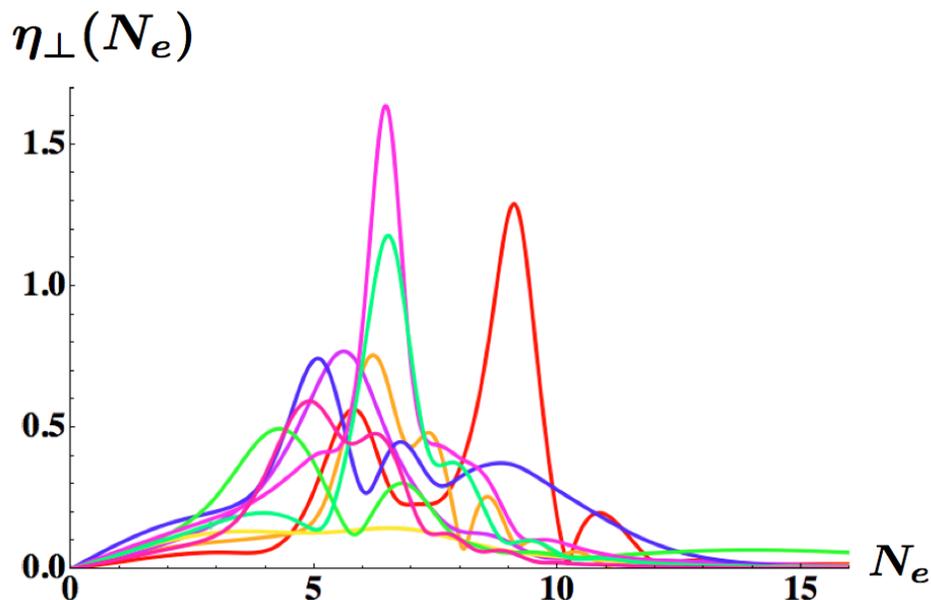


Figure 3.9: The curves show the rapid changes in  $\eta_{\perp}$  during the first 15 e-folds for 9 realizations of inflation. Note that  $N_e = 0$  refers to the start of inflation in each realization, rather than to the moment when the CMB crosses the Hubble radius, which occurs much later in many examples.

### 3.4.3 Multifield effects and many-field effects

Having understood the properties of the mass matrix and of the background evolution that influence the evolution of entropic perturbations, we are in a position to understand multifield contributions to the primordial perturbations.

One of the most interesting questions for a model of inflation with more than two fields is whether the primordial perturbations are effectively governed by only two fields, or instead have distinctive ‘many-field’ signatures (see [99] for a recent discussion). In a general  $N$ -field model of inflation, as we reviewed in §3.2.3, one of the  $N - 1$  entropic fluctuations, which we call the first entropic fluctuation, plays a distinguished role by instantaneously coupling to the adi-

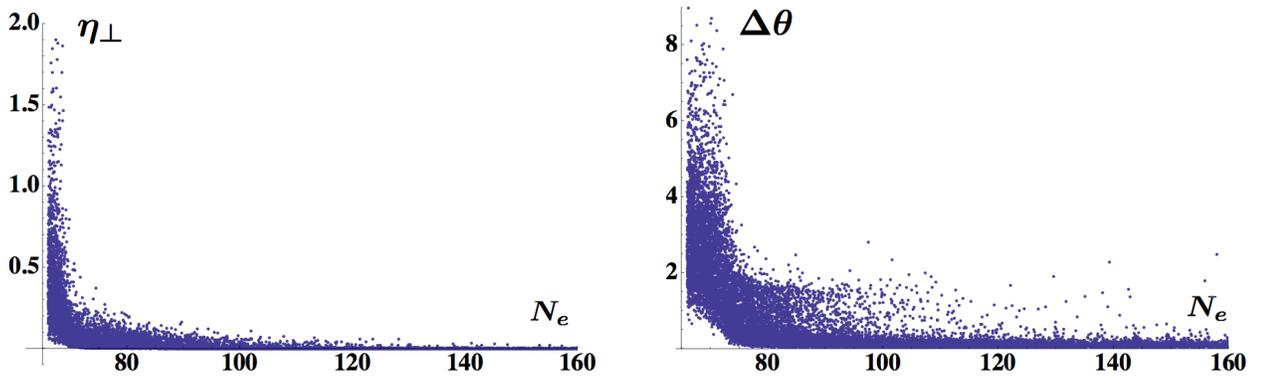


Figure 3.10: Left panel: the slow roll parameter  $\eta_{\perp}$  evaluated at Hubble crossing, versus the total number of e-folds,  $N_e$ . Right panel: the total turn of the trajectory, from six e-folds before Hubble crossing until the end of inflation,  $\Delta\theta \equiv \int_{HC-6}^{end} dN_e \eta_{\perp}$ , versus the total number of e-folds,  $N_e$ .

adiabatic perturbation. Most explicit studies of inflation with more than one field have taken  $N = 2$  for simplicity, in which case the entropic subspace is one-dimensional and this distinction is unnecessary. However, for our system there are five entropic modes, each of which can in principle contribute to the curvature perturbation.<sup>17</sup> We must therefore study not just the incidence of multi-field effects, but also the incidence of ‘many-field’<sup>18</sup> effects, in which the curvature perturbation is not dictated by the adiabatic and first entropic fluctuations alone, and instead receives contributions from the higher (second through fifth) entropic modes. The method for this analysis was described in §3.2.3: when  $\mathcal{P}_{\text{exact}} \neq \mathcal{P}_k$ , then at least  $k + 1$  modes contribute to the curvature perturbation.

<sup>17</sup>Although the second through fifth entropic modes do not couple instantaneously to the adiabatic perturbation, they do couple to the first entropic mode, which can then source the adiabatic perturbation when the trajectory bends.

<sup>18</sup>For the purpose of this discussion, because ‘multifield inflation’ means ‘inflation with two or more light fields’, we will use ‘many-field inflation’ to refer to ‘inflation with three or more light fields’. An important and challenging problem is to understand inflation with  $N \gg 1$  light fields, but we do not use ‘many’ in this sense in the present work.

## Illustrative examples

We begin by examining a collection of illustrative examples, shown in Fig. 3.11. For each example we display the power spectrum as a function of the number of e-folds since Hubble crossing, using the naive, one-field, two-field, and exact models.

In the upper left panel we show a striking example for which all four models give different results. Around Hubble crossing the various models give roughly comparable results for the curvature perturbation, but conversion of entropic perturbations to curvature perturbations after Hubble crossing causes the two-field and exact results to grow sharply. As  $\mathcal{P}_{\text{exact}} \neq \mathcal{P}_{\text{one}}$ , multifield effects cannot be ignored, but because also  $\mathcal{P}_{\text{exact}} \neq \mathcal{P}_{\text{two}}$ , we see that a two-field description is inadequate. Notice the scale: many-field contributions to the curvature perturbation dwarf the two-field contribution, which is itself non-negligible.

Next, in the upper right panel of Fig. 3.11 we exhibit a case in which  $\mathcal{P}_{\text{exact}} \approx \mathcal{P}_{\text{two}} \neq \mathcal{P}_{\text{one}}$ , so that two-field effects cannot be omitted, but many-field effects are negligible. The lower left panel shows an example in which  $\mathcal{P}_{\text{exact}} \neq \mathcal{P}_{\text{two}} \approx \mathcal{P}_{\text{one}}$ , for which many-field effects are significant but two-field effects are negligible. Finally, the lower right panel shows a rare example in which  $\mathcal{P}_{\text{exact}} \approx \mathcal{P}_{\text{two}} \neq \mathcal{P}_{\text{one}}$ , so that two-field but not many-field effects are important, and the resulting power spectrum is compatible with observational constraints on the tilt.

## The incidence of multifield effects

Equipped with a few examples of the possible phenomenology, we now proceed to a statistical analysis of the incidence of multifield effects in our ensemble.

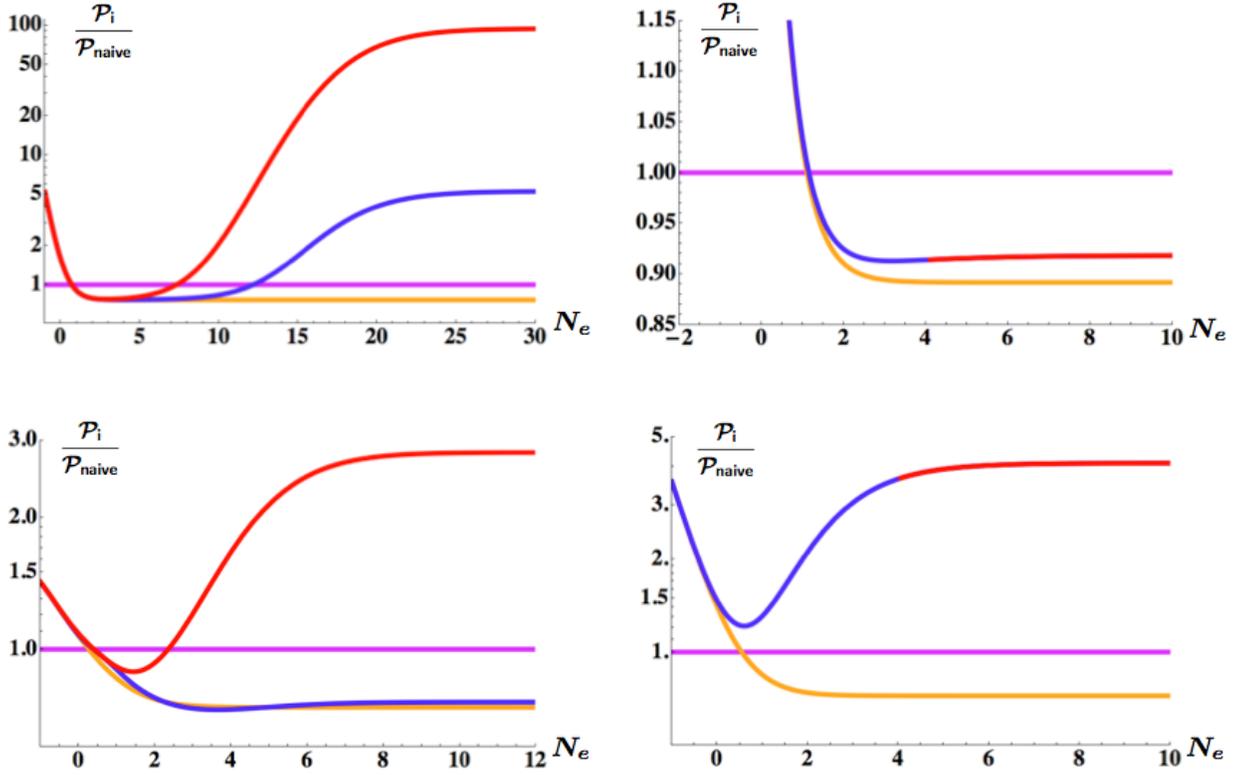


Figure 3.11: Scalar power spectra as functions of the number of e-folds since Hubble crossing, for four different realizations. The purple, orange, blue, and red lines correspond to the naive, one-field, two-field, and exact spectra, respectively. Lines with combined colors indicate that the corresponding spectra overlap to high accuracy. Upper left: many-field and two-field effects are both important. Upper right: two-field effects are present, but many-field effects are negligible. Lower right: two-field effects are present but many-field effects are negligible, in a rare example that is consistent with constraints on the tilt (see §3.4.4). Lower left: two-field effects are negligible, but many-field effects are large.

First, in Fig. 3.12 we show the cumulative probability that multifield contributions to the scalar power have a given size. The upper curve corresponds to  $\xi_{\text{exact/one}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{one}} - 1|$ . We find that approximately 30% of realizations have  $\xi_{\text{exact/one}} \geq .01$ , corresponding to a 1% correction to the power, and 10% of real-

izations have  $\xi_{\text{exact/one}} \geq 1$ , corresponding to a 100% correction to the power. The tail of the distribution is significant even at very large enhancements.

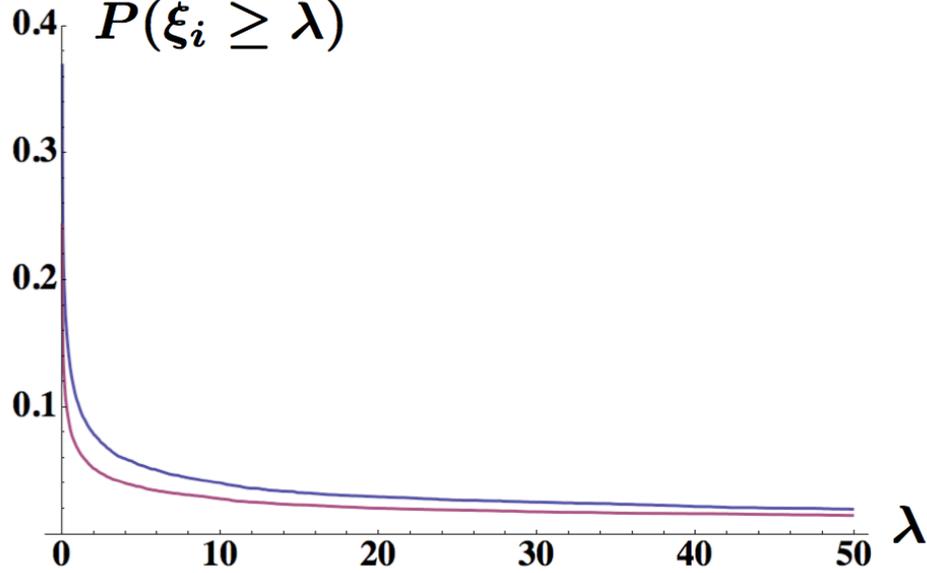


Figure 3.12: The relative probability  $P(\xi \geq \lambda)$  of different levels of corrections of the scalar power. The upper curve shows corrections due to multifield effects, corresponding to  $\xi_{\text{exact/one}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{one}} - 1|$ . The lower curve shows corrections due to many-field effects, corresponding to  $\xi_{\text{exact/two}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{two}} - 1|$ . Note that  $\lambda = 1$  corresponds to a 100% correction, and the tails extend to very large enhancements.

Next, in Fig. 3.13 we show the relationship between the multifield correction to the power spectrum and the total turn  $\Delta\theta$  of the inflationary trajectory. An important fact visible in Fig. 3.13 is that for any value of the enhancement in power due to multifield effects,  $\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{one}}$ , there is a minimum amount of total turn  $\Delta\theta$  necessary to effect such an enhancement. For total turning  $\Delta\theta$  exceeding a threshold value  $\Delta\theta_{\star} \approx 1.4$ , the enhancement in power can be very large. For practical purposes, the vertical range of Fig. 3.13 has been restricted; see Fig. 3.12 for an alternative representation of the incidence of large enhancements.

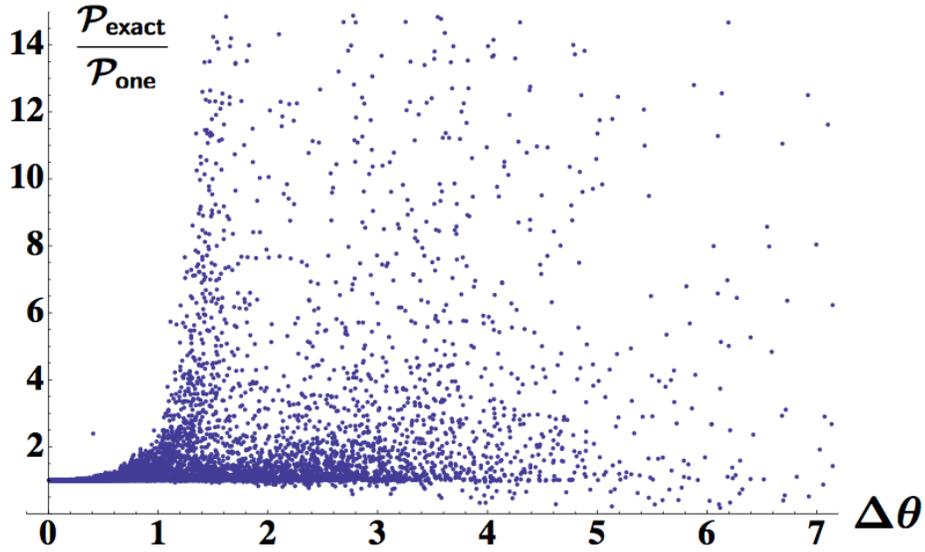


Figure 3.13: The ratio of the exact power to the one-field power, versus the total turn of the trajectory during inflation, for all realizations. Notice the clear threshold of total turning required to achieve a given ratio of multifield to single-field power.

### The incidence of many-field effects

We now turn to characterizing the incidence of many-field effects, characterized by  $\xi_{\text{exact/two}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{two}} - 1|$ . We remind the reader that the two-field truncation retains the instantaneous adiabatic and first entropic modes, so that  $\xi_{\text{exact/two}} > 0$  indicates that additional entropic modes contribute to the perturbations. For practical reasons, although we computed the exact scalar power for all 18731 realizations yielding at least 66 e-folds of inflation, we computed the two-field power only in the effectively multifield ensemble, for which  $\xi_{\text{exact/one}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{one}} - 1| \geq 0.01$ . The logic is that many-field effects will be significant only if multifield effects are significant, so that we can compute the frequency of many-field effects in the full ensemble without computing  $\mathcal{P}_{\text{two}}$  in

the effectively single-field subset.<sup>19</sup>

Many-field effects are only slightly less common than multifield effects in general: many-field corrections exceed 1% (i.e.,  $\xi_{\text{exact/two}} \geq 0.01$ ) in 18% of all realizations, and exceed 100% in 6% of realizations. Among effectively multifield models, many-field effects are commonplace:  $\xi_{\text{exact/two}} \geq 0.01$  in 62% of models with  $\xi_{\text{exact/one}} \geq 0.01$ . The lower curve in Fig. 3.12 indicates the cumulative probability that many-field contributions to the scalar power have a given size, as measured by  $\xi_{\text{exact/two}}$ . Comparing to the upper curve, which corresponds to  $\xi_{\text{exact/one}} \equiv |\mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{one}} - 1|$ , we see that when multifield effects are large, the likelihood of many-field effects grows dramatically. On the other hand, large many-field effects and large two-field effects are rarely present in the same model. The anticorrelation between two-field and many-field effects is clearly visible in Fig. 3.14, which primarily consists of two distinct populations, one with exclusively two-field effects (horizontal branch) and another with exclusively many-field effects (vertical branch). It is worth stressing the peculiarity of this latter population: based on a two-field description, one would wrongly conclude that these models have negligible multifield effects, while in fact only higher entropic modes affect the curvature perturbation.

### Destructive multifield effects

An interesting phenomenon visible in Fig. 3.13 and Fig. 3.14 is destructive interference, in which  $\mathcal{P}_{k+1} < \mathcal{P}_k$  for some  $k$ . Destructive two-field effects are

---

<sup>19</sup>In principle, one can envision a scenario in which  $\mathcal{P}_{\text{exact}} \approx \mathcal{P}_{\text{one}} \ll \mathcal{P}_{\text{two}}$ , so that two-field and many-field effects are both present but cancel out in the exact power (see §3.4.3 for a discussion of this point). Correspondingly, some scenarios with  $\xi_{\text{exact/one}} < 0.01$  could involve significant many-field effects. We have verified that this occurrence is very infrequent in our ensemble, occurring in less than 1% of realizations.

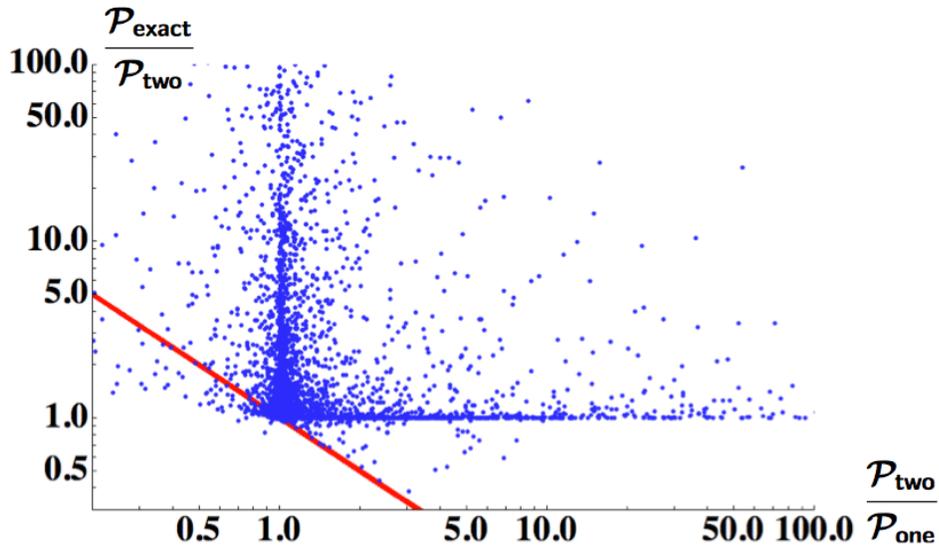


Figure 3.14: Many-field and two-field effects, for effectively multifield models (see §3.2.3). Two distinct populations are visible: the vertical branch has negligible two-field effects and significant many-field effects, while the horizontal branch has significant two-field effects and negligible many-field effects. Points below the red line have destructive multifield effects, i.e.  $\mathcal{P}_{\text{exact}} < \mathcal{P}_{\text{one}}$ , cf. §3.4.3.

not uncommon, but one can see from Fig. 3.14 that the majority of cases with  $\mathcal{P}_{\text{two}} < \mathcal{P}_{\text{one}}$  have  $\mathcal{P}_{\text{exact}} > \mathcal{P}_{\text{one}}$ , so that the net effect of all entropic modes is to add power. However, a handful of examples with  $\mathcal{P}_{\text{exact}} < \mathcal{P}_{\text{one}}$  are visible in Fig. 3.13.

In certain approximate treatments of the evolution of the perturbations in multifield inflation, the addition of entropic modes can only increase the power spectrum of the curvature perturbation  $\mathcal{R}$ , and the destructive multifield effects that we observe cannot arise. We will therefore pause to explain why destructive multifield effects are not forbidden, and what implicit assumptions lead to the conclusion that destructive multifield effects cannot arise. This discussion is

somewhat decoupled from the rest of the paper and can be omitted in a first reading.

Let us consider a two-field inflationary model for the sake of simplicity. In the exact description, a possible quantization scheme consists of writing

$$\hat{\mathcal{R}}_{\mathbf{k}}^{\text{two}} = \mathcal{R}_k^\sigma \hat{a}_{\mathbf{k}}^\sigma + (\mathcal{R}_k^\sigma)^* (\hat{a}_{-\mathbf{k}}^\sigma)^\dagger + \mathcal{R}_k^s \hat{a}_{\mathbf{k}}^s + (\mathcal{R}_k^s)^* (\hat{a}_{-\mathbf{k}}^s)^\dagger, \quad (3.28)$$

where the creation and annihilation operators satisfy

$$[\hat{a}_{\mathbf{k}}^I, (\hat{a}_{\mathbf{k}'}^J)^\dagger] = \delta^{IJ} \delta(\mathbf{k} - \mathbf{k}'), \quad (3.29)$$

with  $I, J \in (\sigma, s)$ . The connection to the method of numerical evolution described in §3.2.3 is as follows. The exact coupled equations of motion are solved twice, for two different initial conditions. In a first run, the adiabatic fluctuation begins in the Bunch-Davies vacuum, while the entropic perturbation is initially set to zero. The corresponding solution for  $\mathcal{R}$  leads to  $\mathcal{R}_k^\sigma$ . In a second run, the entropic fluctuation begins in the Bunch-Davies vacuum, while the adiabatic perturbation is initially set to zero. The corresponding solution for  $\mathcal{R}$  leads to  $\mathcal{R}_k^s$ . The total power spectrum of  $\mathcal{R}$  is then given by

$$\langle 0 | \hat{\mathcal{R}}_{\mathbf{k}}^{\text{two}} \hat{\mathcal{R}}_{\mathbf{k}'}^{\text{two}} | 0 \rangle = \delta(\mathbf{k} + \mathbf{k}') (|\mathcal{R}_k^\sigma|^2 + |\mathcal{R}_k^s|^2). \quad (3.30)$$

Thus, in the exact description, there are two contributions to the curvature perturbation that add in quadrature.

In the one-field description, on the other hand, one simply writes

$$\hat{\mathcal{R}}_{\mathbf{k}}^{\text{one}} = \mathcal{R}_k \hat{a}_{\mathbf{k}}^\sigma + (\mathcal{R}_k)^* (\hat{a}_{-\mathbf{k}}^\sigma)^\dagger, \quad (3.31)$$

and the entropic fluctuations are set to zero for all time, which leads to

$$\langle 0 | \hat{\mathcal{R}}_{\mathbf{k}}^{\text{one}} \hat{\mathcal{R}}_{\mathbf{k}'}^{\text{one}} | 0 \rangle = \delta(\mathbf{k} + \mathbf{k}') |\mathcal{R}_k|^2. \quad (3.32)$$

If  $\mathcal{R}_k^\sigma$  in equation (3.28) and  $\mathcal{R}_k$  in equation (3.31) were identical, then we would conclude that entropic perturbations affect the two-point function of the curvature perturbation exclusively through the term  $|\mathcal{R}_k^s|^2$  in equation (3.30), and this contribution is manifestly nonnegative.

In fact, however,  $\mathcal{R}_k^\sigma$  in equation (3.28) and  $\mathcal{R}_k$  in equation (3.31) are quite different. In the evolution that determines  $\mathcal{R}_k^\sigma$ , the entropic fluctuations are initially set to zero, but are not forced to be zero for all time: entropic fluctuations can be generated through coupling with the adiabatic mode, and can then back-react on the adiabatic mode itself. Thus, it can happen that  $|\mathcal{R}_k^\sigma|^2 < |\mathcal{R}_k|^2$ , and in turn it is logically possible — though not necessarily common — to have  $|\mathcal{R}_k^\sigma|^2 + |\mathcal{R}_k^s|^2 < |\mathcal{R}_k|^2$ , i.e.  $\mathcal{P}_{\text{two}} < \mathcal{P}_{\text{one}}$ .

Another perspective on this effect comes from considering the coupling between the adiabatic and entropic perturbations during Hubble crossing. A convenient assumption that is often utilized (but not always explicitly invoked) in the literature is that  $\eta_\perp \ll 1$  during the few e-folds of Hubble crossing, so that the adiabatic and entropic perturbations evolve independently during that time. In this approximation, adiabatic and entropic perturbations are uncorrelated soon after Hubble crossing, say at a time  $t_\star$  at which spatial gradients can already be neglected, so that one can write (again taking a two-field model for simplicity)

$$\hat{\mathcal{R}}_{\mathbf{k}}^{\text{two}} = \mathcal{R}_k^\sigma \hat{a}_{\mathbf{k}\star}^\sigma + (\mathcal{R}_k^\sigma)^* (\hat{a}_{-\mathbf{k}\star}^\sigma)^\dagger + \mathcal{R}_k^s \hat{a}_{\mathbf{k}\star}^s + (\mathcal{R}_k^s)^* (\hat{a}_{-\mathbf{k}\star}^s)^\dagger, \quad (3.33)$$

where  $\hat{a}_{\mathbf{k}\star}^\sigma$  and  $\hat{a}_{\mathbf{k}\star}^s$  satisfy the relation (3.29), and in particular are *independent*. We stress that  $\hat{a}_{\mathbf{k}\star}^\sigma$  and  $\hat{a}_{\mathbf{k}\star}^s$  are different from  $\hat{a}_{\mathbf{k}}^\sigma$  and  $\hat{a}_{\mathbf{k}}^s$  in equation (3.28): the former are defined such that the power spectrum of  $\hat{\mathcal{R}}_{\mathbf{k}}^{\text{two}}$  at  $t_\star$  coincides with  $|\mathcal{R}_k^\sigma(t_\star)|^2$ . Just as before,  $\mathcal{R}_k^\sigma(t \geq t_\star)$  is computed by solving the full system of equations, setting the entropic perturbations to zero at  $t = t_\star$ , while  $\mathcal{R}_k^s(t \geq t_\star)$  is

computed by solving the full system of equations, setting the adiabatic perturbations to zero at  $t = t_*$ . The power spectrum is then obtained as in equation (3.30). However, the assumption of decoupling during Hubble crossing results in one crucial difference, as we now explain. As is well known, although the entropic fluctuation can source the adiabatic fluctuation on super-Hubble scales, the adiabatic fluctuation does not source the entropic fluctuation in this regime [100]. Hence, if the entropic fluctuation and its time derivative are set to zero at time  $t_*$ , they will remain zero, and no backreaction on  $\mathcal{R}_k^\sigma(t \geq t_*)$  is possible. Correspondingly, if multifield effects are neglected until a time  $t_*$  after which the adiabatic mode cannot source the entropic mode — which is precisely what occurs when the slow turn approximation is made during Hubble crossing — then  $\mathcal{R}_k^\sigma$  coincides with the one-field quantity  $\mathcal{R}_k$  in (3.31). Because of the second term in (3.30), one then always obtains that, under these approximations,  $\mathcal{P}_{\text{two}} \geq \mathcal{P}_{\text{one}}$ . This corresponds, for example, to the prediction of the zeroth-order transfer functions formalism.

As should be clear from the discussion above, a necessary condition for destructive multifield effects is bending of the trajectory around the time of Hubble crossing. As a consistency check, we have verified that in our realizations with destructive multifield effects,  $\eta_\perp$  always has a large peak at Hubble crossing. However, bending at Hubble exit is not a sufficient condition: a large number of our realizations have  $\eta_\perp \gtrsim 1$  during Hubble crossing but do not display destructive multifield effects.

We should point out that although  $\mathcal{P}_{\text{exact}} \ll \mathcal{P}_{\text{one}}$  occurs in our ensemble,  $\mathcal{P}_{\text{exact}} \ll \mathcal{P}_{\text{naive}}$  does not: when  $\mathcal{P}_{\text{exact}} \ll \mathcal{P}_{\text{one}}$ , we typically find  $\mathcal{P}_{\text{naive}} \ll \mathcal{P}_{\text{one}}$  and  $\mathcal{P}_{\text{exact}} = O(\mathcal{P}_{\text{naive}})$ . Even so, examples with  $\mathcal{P}_{\text{exact}} < \mathcal{P}_{\text{naive}}$  are common: 61%

of all models have  $0.63 < \mathcal{P}_{\text{exact}}/\mathcal{P}_{\text{naive}} < 1$ . However, none of the realizations with  $\mathcal{P}_{\text{exact}} < \mathcal{P}_{\text{one}}$  or  $\mathcal{P}_{\text{exact}} < \mathcal{P}_{\text{naive}}$  has a scalar tilt in the observationally allowed region (see §3.4.4).

Finally, we remark that the amplitude of primordial gravitational waves, as measured by  $r$ , can be slightly larger than what the naive estimate  $r = 16\epsilon$  would suggest. In several examples we indeed find  $\mathcal{P}_{\mathcal{T}}/\mathcal{P}_{\text{exact}} > 16\epsilon$ , with  $\mathcal{P}_{\mathcal{T}}$  denoting the tensor power spectrum. It would be interesting to understand if it is therefore possible to violate the consistency relation  $r \leq -8n_{\mathcal{T}}$  [64, 101], but to determine this requires computing  $n_{\mathcal{T}}$  very precisely, which is challenging in models with  $r \ll 1$ .

### 3.4.4 Constraints from scale invariance

The results described in the preceding sections refer to the ensemble of realizations of inflation that give rise to 66 or more e-folds of expansion followed by a hybrid exit. Such a cosmic history serves to solve the horizon, flatness, and monopole problems, but is not necessarily consistent with observations of the CMB temperature anisotropies. In this section we study multifield corrections to the tilt (§3.4.4), and then describe the restricted ensemble of realizations consistent with WMAP7 constraints on the tilt (§3.4.4).

#### Multifield contributions to the tilt

We begin by characterizing multifield contributions to the tilt  $n_s$  of the scalar power spectrum. In Fig. 3.15 we show the correlation between the multifield

correction to the tilt, defined as  $\delta n_s \equiv n_s^{\text{exact}} - n_s^{\text{naive}}$ , and the naive tilt  $n_s^{\text{naive}}$ . In realizations with a distinctly blue naive spectrum, inflation is occurring well above the inflection point, and transients from the approach to the inflection point have not necessarily decayed at the time that the CMB exits the Hubble radius (see §3.2.2). Correspondingly, there is a larger likelihood of multifield effects, as evident in Fig. 3.15. We observe that multifield effects typically shift the spectrum toward scale invariance, but the size of the effect is rarely large enough to produce a model in the observational window denoted by the orange band. There are a few very interesting cases with large multifield effects that have tilts<sup>20</sup> consistent with WMAP7 at the  $2\sigma$  level, which we will discuss in more detail in §3.4.4.

The reader might object that  $\delta n_s \equiv n_s^{\text{exact}} - n_s^{\text{naive}}$  does not exclusively represent *multifield* effects: if  $n_s^{\text{one}} \neq n_s^{\text{naive}}$  then  $\delta n_s$  contains a contribution from strictly single-field slow-roll-violating effects. Slow-roll-violating effects, for effectively single-field trajectories, were already described in §3.3.3 (see Fig. 3.7). We have verified that for the full ensemble, including realizations with large multifield effects, slow roll violations generally make small ( $\delta n_s^{\text{SR}} \lesssim 0.05$ ) positive contributions to  $n_s$ , as in the single-field case of §3.3.3. In contrast, the multifield effects described above make a much larger negative contribution to  $n_s$ .

### Realizations consistent with constraints on the tilt

In this section we present results for the restricted ensemble of realizations that are consistent with WMAP7 constraints on the *tilt* of the scalar power spectrum, which constitute 21 % of our full ensemble. The overall amplitude of the scalar

---

<sup>20</sup>We have not determined whether the running of the scalar power spectrum further constrains these models.

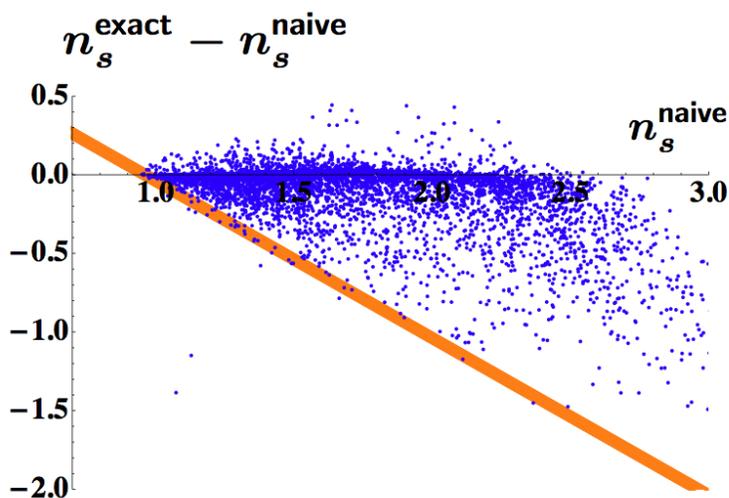


Figure 3.15: The difference  $\delta n_s$  between the exact and naive results for the tilt  $n_s$ , versus the naive tilt, for effectively multifield realizations (see §3.2.3). Notice that multifield contributions typically shift the spectrum toward scale invariance, and the more blue the naive spectrum, the larger the scatter of  $\delta n_s$ . The orange band corresponds to the window allowed at  $2\sigma$  by WMAP7.

power spectrum can be adjusted by changing the scale<sup>21</sup> of the potential, and we will assume that this has been done. We do not incorporate constraints on the running, but it could be interesting to do so.

The consequence of imposing constraints on the tilt is very striking, and easily understood by recalling the single-field inflection point model described in §3.2.2. A single-field inflection point model is compatible with constraints on the tilt only if the total number of e-folds is  $N_e \gtrsim 120$ , so that the observed CMB exits the Hubble radius when the inflaton is at or below the inflection point, at which point the spectrum becomes red.

<sup>21</sup>In the underlying microphysical model, this adjustment corresponds to changing an inflaton-independent contribution to the vacuum energy. The magnitude of the vacuum energy could in principle be correlated with other terms in the potential, but modeling such a correlation is beyond the scope of the present work.

It follows that for any realization consistent with observations, either (i)  $N_e > 120$ , and the tilt can be well-approximated by the single-field slow roll result, or (ii) the number of e-folds is constrained only by the solution of the horizon problem, and the tilt must differ substantially from the single-field slow roll result. Although case (ii) can occur, it is quite rare: multifield effects do tend to redden the spectrum, as noted in §3.4.4, but given how blue the naive spectrum is, the multifield effect is typically not large enough to bring the exact spectrum into agreement with observations (corresponding to points falling inside the orange band in Fig. 3.15). In case (i), which describes 98.7% of realizations consistent with observations, the 60 or more e-folds of inflation that precede the exit of the observed CMB generally suffice to damp out all transient effects from the approach to the inflection point (this is reflected in Fig. 3.16). Thus, although turning trajectories that imprint multifield effects in the curvature perturbations are present in a reasonable fraction (roughly 30%) of realizations of inflation, they only represent 1.3% of models with an observationally allowed tilt, in which these turns are generally complete long before the CMB exits the Hubble radius.

For the same reason, violations of slow roll are rare in realizations consistent with observations: 99% of models with an allowable tilt have  $-0.1 < m_\sigma^2/H^2 < -0.01$ . Similarly, trajectories with substantial bending are uncommon once constraints on the tilt are imposed. In the left panel of Fig. 3.16 we show  $\eta_\perp$  at Hubble crossing for the observationally allowed ensemble, which is to be contrasted with the left panel of Fig. 3.10. (In order to show the structure at small  $\eta_\perp$ , we have omitted a small number of points with  $\eta_\perp \sim 0.1$ .) The right panel of Fig. 3.16 shows the total turn of the trajectory, to be contrasted with the right panel of Fig. 3.10. Note that a handful of realizations allowed by constraints on

the tilt do have significant bending.

It is interesting to understand the characteristics of the realizations that have important multifield effects and are also consistent with constraints on the tilt. Perhaps surprisingly, these realizations do not display extraordinary background behavior. Of course, all such models have non-negligible bending of the trajectory around or after Hubble crossing, but the maximum value of  $\eta_{\perp}$  along the inflationary evolution need not be greater than  $O(0.1)$  to generate significant multifield effects. In fact, a majority display only a moderate degree of bending,  $\eta_{\perp} \sim O(0.05)$ , albeit sustained over ten or more e-folds. One distinctive difference compared to the general population is a substantially smaller value of the first entropic mass,  $m_{s_1}^2$ , at the time of Hubble crossing. Whereas  $m_{s_1}^2$  ranges from 0 to 10 in the full ensemble, typical values are of order 0.1 in observationally allowed realizations with important multifield effects. We show an example of such a realization in Fig. 3.17.

### 3.4.5 Non-Gaussianities

Although so far we have only discussed the scalar power spectrum, a major motivation for the study of multifield models is the prospect of a detectable primordial bispectrum. The D3-brane inflation scenario we have considered involves six fields that interact via Planck-suppressed couplings, leading to masses of order  $H$  for each field.<sup>22</sup> Our setting therefore corresponds to a microphysical realization of *quasi-single-field inflation*, introduced by Chen and Wang in [102]. Quasi-single-field inflation occupies the middle ground between single-field in-

---

<sup>22</sup>Recall from §3.3.1 that this is only a parametric estimate, and there is considerable additional structure in the scalar mass spectrum.

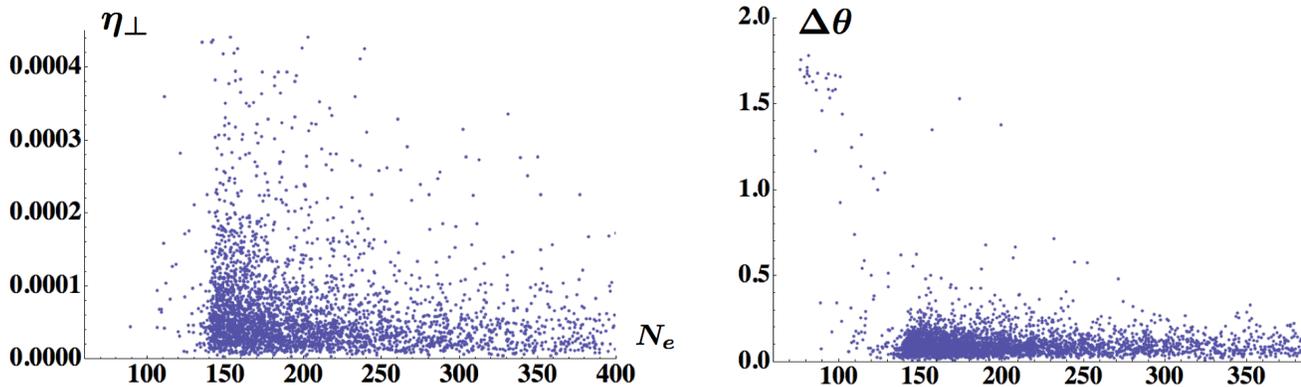


Figure 3.16: Left panel: the slow roll parameter  $\eta_{\perp}$  evaluated at Hubble crossing, versus the total number of e-folds,  $N_e$ , for models with tilt  $n_s$  consistent with observations. Right panel: the total turn of the trajectory, from six e-folds before Hubble crossing until the end of inflation,  $\Delta\theta \equiv \int_{HC-6}^{end} dN_e \eta_{\perp}$ , versus the total number of e-folds,  $N_e$ , for models with tilt  $n_s$  consistent with observations.

flation, in which any scalar fields other than the inflaton have masses  $m \gg H$ , and the simplest models of multifield inflation, in which all the fluctuating fields have masses  $m \ll H$ .

A striking feature of quasi-single-field inflation is the possibility of large non-Gaussianity, given apparently modest cubic couplings in the Lagrangian, and modest rates of turning, as measured by  $\eta_{\perp}$ . Specifically, there is a contribution to the bispectrum scaling as [102]

$$f_{NL} \sim \frac{1}{\sqrt{\mathcal{P}_{\mathcal{R}}}} \frac{V_{s_1 s_1 s_1}}{H} \eta_{\perp}^3, \quad (3.34)$$

where  $V_{s_1 s_1 s_1}$  denotes the third derivative of the potential with respect to the instantaneous first entropic direction. Evidently, if  $V_{s_1 s_1 s_1}$  is not very small compared to  $H$  then  $f_{NL}$  can be large, even for perturbatively small  $\eta_{\perp}$ .

We should point out that several important simplifying assumptions made

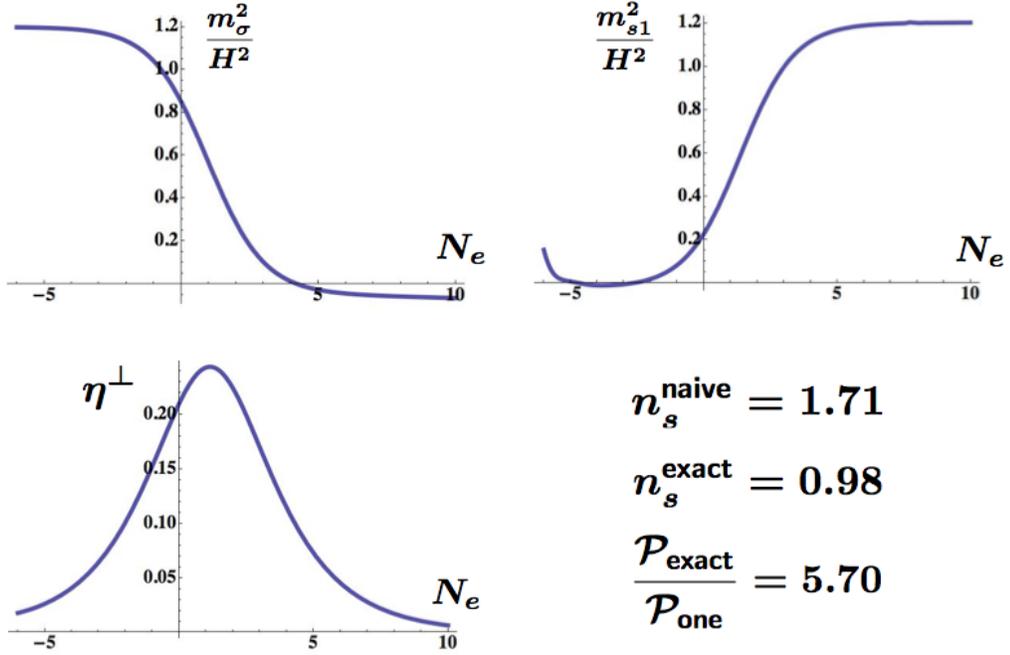


Figure 3.17: An example in which multifield effects — specifically, two-field effects — are large, and the power spectrum is consistent with observations. This realization also appears in the lower right panel of Fig. 3.11.

in [102] are not necessarily applicable in our setting. In particular,  $\eta_{\perp}$  and the first entropic mass squared,  $m_{s_1}^2$ , were assumed to be constant during inflation. Moreover, the authors of [102] considered a regime in which the multifield effects give subdominant contributions to the amplitude of the power spectrum. Finally, it was assumed in [102] that  $\eta_{\perp}$  was small enough to be used as a perturbative expansion parameter. This last point implies that the operator identified as leading in [102] may not give the dominant contribution to  $f_{NL}$  in our context. In the following, we will obtain a rough picture of the prospects for non-Gaussianity in our ensemble using the estimate (3.34). Going beyond the approximations above is an interesting problem for the future.

Although quasi-single-field inflation is a plausible proposal at the level of field theory, it is reasonable to ask whether the particular assumptions required to achieve detectable non-Gaussianity are natural in the effective theories arising from well-motivated microscopic constructions. For clarity, we distinguish two properties of the effective theory: first, that typical scalars — including the entropic fluctuations — have masses of order  $H$ ; and second, that the cubic couplings  $V_{s_1 s_1 s_1}$  are not too small<sup>23</sup> compared to  $H$ . There is very ample and long-established evidence for the first property in theories with spontaneously broken supersymmetry, as recently explained in detail in [46], and as noted above this is borne out in detail in our ensemble. However, the scale of the cubic couplings is more subtle, as we now explain.

Consider a D3-brane in a conifold region of a de Sitter compactification of string theory. Using a spurion superfield  $X$  to represent the source of supersymmetry breaking, so that  $|F_X|^2 = 3H^2 M_p^2$ , the scalar potential for the D3-brane includes contributions of the form [6]

$$V = \sum_{\Delta} c_{\Delta} \int d^4\theta X^{\dagger} X \left( \frac{\phi}{M_p} \right)^{\Delta} f_{\Delta}(\Psi) = \sum_{\Delta} 3c_{\Delta} H^2 M_p^2 \left( \frac{\phi}{M_p} \right)^{\Delta} f_{\Delta}(\Psi), \quad (3.35)$$

where  $\Delta$  is an operator dimension,  $c_{\Delta}$  is a Wilson coefficient that is expected to be of order unity, and  $\phi = \sqrt{T_3} r$ , with  $T_3$  the D3-brane tension, is the canonically normalized field representing radial motion. Here  $f_{\Delta}(\Psi)$  is a (known) function of the dimensionless angular coordinates  $\theta_1, \phi_1, \theta_2, \phi_2, \psi$ , collectively denoted  $\Psi$ .

It will be critical to relate the  $\Psi$  to the canonically normalized angular fields, which we denote by  $\varphi_a$ ,  $a = 1 \dots 5$ . The metric on the angular manifold  $T^{1,1}$  is not diagonal in the  $\theta_1, \phi_1, \theta_2, \phi_2, \psi$  basis, but one can perform a coordinate transformation to obtain independent dimensionless fields, which we may write as

---

<sup>23</sup>For  $V_{s_1 s_1 s_1} \gtrsim H$ , there is a risk of large radiative corrections to the masses. We thank Daniel Baumann and Daniel Green for instructive discussions of this point.

$\hat{\Psi}_a$ ,  $a = 1 \dots 5$ . The kinetic term for the  $\hat{\Psi}_a$  takes the form  $\mathcal{L} \supset \sum_{a=1}^5 \phi^2 (\partial \hat{\Psi}_a)^2$ , so at a given radial location  $\phi = \phi_\star$ , the canonically normalized angular fields are given by  $\varphi_a = \phi_\star \hat{\Psi}_a$ .

The cubic coupling for three angular fields therefore takes the schematic form

$$V_{\varphi_a \varphi_b \varphi_c} \sim \sum_{\Delta} c_{\Delta} \frac{H^2}{M_p} \left( \frac{\phi_\star}{M_p} \right)^{\Delta-3} f_{\Delta}^{abc}(\Psi), \quad (3.36)$$

where  $f_{\Delta}^{abc}$  denotes the corresponding derivative of  $f$ , which is generally of order unity. We observe that differentiating with respect to canonically normalized angular fields has introduced the factor  $(M_p/\phi_\star)^3$ . As many operators with  $\Delta < 3$  are present in the D3-brane Lagrangian, and the inflationary inflection point occurs at  $\phi_\star \ll M_p$ , the cubic couplings of the angular fields are therefore parametrically large compared to  $\frac{H^2}{M_p}$ . This is worth emphasizing: if all fields had entered the Lagrangian on the same footing, and the scalar potential took the schematic form<sup>24</sup>

$$V = H^2 M_p^2 P(\phi_1/M_p, \dots, \phi_N/M_p), \quad (3.37)$$

with  $P$  a general polynomial of  $N$  fields  $\phi_1/M_p, \dots, \phi_N/M_p$ , with Taylor coefficients of order unity, then the cubic couplings would scale as  $V''' \sim H^2/M_p$ . In the conifold (or any other Calabi-Yau cone) we find a potentially dramatic enhancement compared to this ‘democratic’ estimate.

The discussion so far has only addressed the couplings of the angular fields.

The cubic couplings of the radial field  $\phi$  include terms of the form

$$V_{\phi\phi\phi} \sim c_{\Delta} \Delta(\Delta-1)(\Delta-2) \frac{H^2}{M_p} \left( \frac{\phi_\star}{M_p} \right)^{\Delta-3}. \quad (3.38)$$

---

<sup>24</sup>This estimate relies on the assumption that the fields are moduli whose masses and cubic couplings vanish before supersymmetry breaking. For D3-branes in the conifold, and more generally for moduli in string compactifications, this assumption is appropriate, but in more general settings larger cubic couplings need not be unnatural. We thank Daniel Baumann for helpful comments on this point.

Thus, if the  $\Delta$  were all integers, we would have  $V_{\phi\phi\phi} \lesssim H^2/M_p$ , which is too small to contribute to detectable non-Gaussianity (see (3.40) below). However, the operator dimensions are not all integers in the field theory dual to the conifold. Inflection point inflation in this model typically results from a cancellation between a term with  $\Delta = 3/2$  and the well-known conformal coupling term with  $\Delta = 2$  (see [26] for a discussion), which can occur at small values of  $\phi$ . Differentiation of the term with  $\Delta = 3/2$  gives rise to a cubic coupling for  $\phi$  that is of the same order as the cubic couplings of the angular fields.

To summarize, in a natural effective theory describing moduli that obtain their potential from gravitational-strength coupling to a source of supersymmetry breaking, in which the scalar potential involves a generic polynomial of the form (3.37), involving only integer powers of the fields, with all fields entering on the same footing, the cubic couplings scale as  $V''' \sim H^2/M_p$ , and are far too small to generate detectable non-Gaussianity. We have seen that D-brane inflation in the conifold necessarily modifies this simple picture, in two ways: the operator dimensions are not all integers, and the radial and angular fields play different roles in the potential. The result is that the cubic couplings of all six fields are parametrically of order

$$V''' \sim \sum_{\Delta} \frac{H^2}{M_p} \left( \frac{\phi_{\star}}{M_p} \right)^{\Delta-3}. \quad (3.39)$$

Having addressed the parametric scalings of the cubic couplings, we turn to examining the range of numerical values that the quantities appearing in equation (3.34) attain in our ensemble. As a measure of the scale  $H$ , we use  $\epsilon$ : equation (3.34) can then be rewritten as

$$f_{NL} \sim 10 \sqrt{\epsilon} \frac{V_{s_1 s_1 s_1} M_p}{H^2} \eta_{\perp}^3. \quad (3.40)$$

The typical value is  $\epsilon \sim 10^{-12}$ , and there is a tail toward somewhat larger values: 7% of realizations have  $\epsilon > 10^{-8}$ , though we found no examples with  $\epsilon > 10^{-6}$ . The typical size of  $\eta_{\perp}$  is  $\sim 10^{-3}$ , while  $\eta_{\perp} < 1$  in 99% of realizations, and we found no example with  $\eta_{\perp} > 3$ .<sup>25</sup> The typical scale of the cubic coupling of the first entropic mode in our ensemble is  $V_{s_1 s_1 s_1} \sim 10^3 H^2 / M_p$ , and as a conservative estimate we take  $V_{s_1 s_1 s_1} \lesssim 10^4 H^2 / M_p$ . Assembling these results, we find that  $f_{NL} > 10$  in less than  $1\% \times 7\% = .07\%$  of all realizations.

The results above apply to the full ensemble, but non-Gaussianity is far less likely in the ensemble consistent with constraints on the tilt: 97 % of observationally allowed models have  $\eta_{\perp} \lesssim 10^{-3}$ , and 99% have  $\epsilon < 10^{-13}$ , so that there is a negligible prospect of detectable non-Gaussianity in such models.

It is straightforward to extend our analysis to the quartic couplings, and correspondingly to estimate the overall amplitude  $t_{NL}$  of the trispectrum. We rewrite the result [102]

$$t_{NL} \sim \max \left( \frac{1}{\mathcal{P}_{\mathcal{R}}} \left( \frac{V_{s_1 s_1 s_1}}{H} \right)^2 \eta_{\perp}^4, \frac{1}{\mathcal{P}_{\mathcal{R}}} V_{s_1 s_1 s_1 s_1} \eta_{\perp}^4 \right), \quad (3.41)$$

where  $V_{s_1 s_1 s_1 s_1}$  denotes the fourth derivative of the potential with respect to the instantaneous first entropic direction, as

$$t_{NL} \sim 100 \epsilon \eta_{\perp}^4 \max \left( \left( \frac{V_{s_1 s_1 s_1} M_p}{H^2} \right)^2, \frac{V_{s_1 s_1 s_1 s_1} M_p^2}{H^2} \right). \quad (3.42)$$

We find that typically  $V_{s_1 s_1 s_1} \lesssim 10^8 H^2 / M_p^2$ , from which we deduce that there is a very limited possibility of a detectable trispectrum ( $t_{NL} \gtrsim 1000$ ) in the full ensemble, but the corresponding probability is negligibly small in the ensemble consistent with constraints on the tilt.

---

<sup>25</sup>These statements refer to the value of  $\eta_{\perp}$  averaged from one e-fold before Hubble exit until one e-fold after Hubble exit, which can be considerably smaller than the maximum value of  $\eta_{\perp}$  during this period.

In conclusion, our results suggest that significant non-Gaussianity ( $f_{NL} \gtrsim 10$  and/or  $t_{NL} \gtrsim 1000$ ) on scales far outside the present Hubble radius is possible, albeit somewhat rare, in warped D-brane inflation; but correspondingly large non-Gaussianity is very rare on observable angular scales. Even so, a dedicated study of non-Gaussianity in D-brane inflation — or in some other well-motivated microphysical realization of quasi-single-field inflation — would be worthwhile.

### 3.5 Conclusions

We have determined the scalar power spectrum that results when inflation takes place in a random six-field potential. The potential in question governs the motion of a D3-brane in a conifold region of a string compactification, and was derived and extensively studied in [6]. The signal property for this work is not the string theory provenance of the inflaton action, but merely the fact that the action is natural: the terms in the potential correspond to Planck-suppressed operators with coefficients of order unity, and the masses are of order  $H$ , as expected for moduli with gravitational-strength couplings to a source of supersymmetry breaking. For this reason, we believe our analysis is representative of a general class of multifield potentials that are natural in the Wilsonian sense, even though aspects of the conifold geometry do influence our results.

The essence of our approach to computing the primordial perturbations was the comparison between numerical integration of the exact equations for the linearized perturbations, making no approximation, and an array of truncated models omitting one or more of the entropic modes. This allowed us to deter-

mine the extent to which various approximations and truncations — including the slow roll approximation, the single-field truncation, and the two-field truncation — capture the physics of a generic realization of inflation. As we made no slow roll or slow turn approximation, we were able to characterize phenomena that are common in our ensemble, but more rarely seen in analytic treatments.

We began by determining the spectrum of scalar masses using a random matrix model, building on [94], and demonstrated excellent agreement with simulations of the full potential. We found that at the time of Hubble exit, one entropic mode is typically light enough to fluctuate, but before the end of inflation an adiabatic limit is reached, and one can predict the late-time curvature perturbation without modeling the details of reheating.

Our results for the perturbations are usefully divided into characterizations of the ensemble of realizations of inflation that give rise to at least 60 e-folds of inflation, but are otherwise unconstrained, and characterizations of the ensemble of realizations of inflation that give rise to at least 60 e-folds of inflation, and are also consistent with observations of the CMB temperature anisotropies: conditioning on the approximate scale-invariance required by observations drastically changes the outcome.

For the ensemble of inflationary models that solve the horizon problem but are not required to give nearly scale-invariant density perturbations, multifield effects are often significant: multifield corrections to the spectrum are at least at the 1% level in roughly 30% of realizations, and exceed 100% in 10% of realizations. We also found that *many-field* contributions to the perturbations — by which we mean effects that cannot be described by the two-field truncation, which retains only the instantaneous adiabatic and first entropic modes — are

similarly common: many-field corrections exceed 1% in 18% of realizations, and exceed 100% in 6% of realizations. Most models with substantial multifield effects have either large two-field effects or large many-field effects, but not both. Finally, we observed that the exact scalar power can be smaller than the power calculated in the single-field truncation: this is a consequence of trajectories that turn quickly at the time of Hubble crossing.

Turning now to the restricted ensemble of realizations that are consistent with the WMAP7 constraints on the tilt of the scalar power spectrum, corresponding to 21% of the total ensemble, we find a qualitatively different picture. When inflation occurs near an approximate inflection point, as it does in the class of potentials studied here, the tilt depends on whether the observed CMB exits the Hubble radius before or after the inflaton descends past the inflection point. When the 60 e-fold mark occurs above the inflection point, the curvature of the potential is positive and the spectrum computed in the single-field slow roll approximation is blue, which is inconsistent with observations. Therefore, the allowable potentials are those in which 60 or more e-folds occur below the inflection point. In such potentials, there is a prolonged phase of inflation before the observed CMB exits the Hubble radius: namely, the inflation occurring above the inflection point. This phase generally suffices to damp out all transients from the onset of inflation, *including the curving trajectories that can give rise to multifield effects in the perturbations*. For this simple reason, although multifield contributions to the perturbations are commonplace in random inflection point models, for more than 98% of models with an observationally allowed tilt these contributions arise on angular scales that are far outside the present Hubble radius. We conclude that for inflection point models of the sort studied here, multifield effects in the observable perturbations are uncommon in realizations

consistent with present limits on scale-invariance. We stress that the multifield effects themselves are *not* responsible for the problematic deviations from scale invariance: in fact, the spectrum calculated in the single-field approximation is blue, and multifield effects shift it toward scale invariance. Instead, multifield effects generally arise from transients from the onset of inflation, and at the onset of inflation at an inflection point, the dominant single-field component of the spectrum is often, but not always, unacceptably blue.

The ensemble we have studied provides a concrete microphysical realization of quasi-single-field inflation. The couplings in the inflaton action arise from Planck-suppressed operators, and the cubic and quartic interactions of the entropic modes, although enhanced by contributions from an operator with dimension  $\Delta = 3/2$ , as well as by the intrinsic asymmetry between the radial and angular directions of the conifold, are only occasionally sufficiently large to produce detectable non-Gaussianity along the lines envisioned in [102]. More generally, the suppression of transients resulting from constraints on the tilt ensures, as explained above, that in a majority of realizations consistent with observations of the spectrum, the single-field slow roll approximation is valid. Of course, the perturbations are Gaussian to good approximation in all such cases. Understanding the generality of this finding in well-motivated effective theories is an important problem for the future.

We expect that our methods, and some aspects of our findings, have broad applicability. Our present understanding of string compactifications suggests considering ensembles of effective theories with numerous moduli and spontaneously broken supersymmetry. In this work we have characterized the cosmological signatures of one such ensemble. We have seen that when inflation

occurs at an inflection point in a many-field potential, then even though it is common for multiple fields to fluctuate, and ultimately to contribute to the curvature perturbations, these contributions are most often on unobservably large angular scales. The signatures of the model are therefore often indistinguishable from those of a single-field inflection point scenario. Nevertheless, it remains to be seen which other universality classes of multifield inflation may exist, and it is reasonable to anticipate very different conclusions when the number of fields is large.

CHAPTER 4  
OSCILLATIONS IN THE CMB FROM AXION MONODROMY  
INFLATION

## 4.1 Introduction

Inflation [1, 2, 3] is a successful paradigm for describing the early universe, but it is sensitive to the physics of the ultraviolet completion of gravity. This motivates pursuing realizations of inflation in string theory, a candidate theory of quantum gravity. Considerable progress has been made on this problem in recent years, so much so that the most pressing task, particularly in view of upcoming CMB experiments, is to learn how to distinguish various incarnations of inflation in string theory from each other and from related models constructed directly in quantum field theory.

Fortunately, the additional constraints inherent in realizing inflation in an ultraviolet-complete framework can leave imprints in the low-energy Lagrangian, and hence ultimately in the cosmological observables. In favorable cases, a given class of models may make distinctive predictions for a variety of correlated observables, allowing one to exclude this class of models given adequate data.

One decisive observable for probing inflation is the tensor-to-scalar ratio,  $r$ . A promising class of string inflation models producing a detectable tensor signature are those involving *monodromy* [10], in which the potential energy is not periodic under transport around an angular direction in the configuration space. The first examples [10] involved monodromy under transport of a

wrapped D-brane in a nilmanifold, and a subsequent class of examples invoked monodromy in the direction of a closed string axion [11].

The axion monodromy inflation scenario of [11] is falsifiable on the basis of its tensor signature,  $r \approx 0.07$ . However, primordial tensor perturbations have not been detected at present, while the temperature anisotropies arising from scalar perturbations have been mapped in great detail [12]. One could therefore hope to constrain axion monodromy inflation more effectively by understanding the signatures that it produces in the scalar power spectrum and bispectrum. Characterizing these signatures is the subject of the present paper.

As we shall explain, the potential in axion monodromy inflation is approximately linear, but periodically modulated: each circuit of the loop in configuration space can provide a bump on top of the otherwise linear potential. Modulations of the inflaton potential with suitable frequency and amplitude can yield two striking signatures: periodic undulations in the spectrum of the scalar perturbations, and resonant enhancement [103] of the bispectrum. Let us stress that the presence of some degree of modulations of the potential is automatic, and is an example of the situation described above in which traces of ultraviolet physics remain in the low-energy Lagrangian. We do not introduce modulations in order to make the scalar perturbations more interesting. However, it is important to examine the typical amplitude and frequency of modulations in models that are under good microphysical control, in order to ascertain whether well-motivated models produce signatures that can be detected in practice.

To achieve this, we first investigate in detail the realization of axion monodromy inflation in string theory. We compute the axion decay constants in terms of compactification data, we assess the importance of higher-derivative

terms, and we estimate the amplitude of modulations for the case of Euclidean D1-brane contributions to the Kähler potential. We also identify a potentially-important contribution to the inflaton potential, arising from backreaction in the compact space, and we present a model-building solution that suppresses this contribution.

We find that detectable modulations of the scalar power spectrum and bispectrum are possible in models that are consistent with all current data and that are under good microphysical control. In fact, we find substantial parameter ranges that are excluded not by microphysics, but by observational constraints on modulations of the scalar power spectrum.

The organization of this paper is as follows. We begin in §4.2 by describing the classical evolution of the homogeneous background in axion monodromy inflation with a modulated linear potential. We then solve, in §4.3, the Mukhanov-Sasaki equation governing the evolution of scalar perturbations, giving an analytical result for the spectrum in terms of the frequency and amplitude of the modulations of the potential. Next, we briefly discuss the bispectrum and express the amplitude of the non-Gaussianity in terms of the model parameters. We then present, in §4.4, an analysis of the constraints imposed on axion monodromy inflation by the WMAP5 data (for prior work constraining similar oscillatory power spectra, see *e.g.* [104, 105, 106, 107, 108, 109, 110]). Then, in §4.5 and §4.6, we present a comprehensive analysis of the constraints imposed by the requirements of computability and of microphysical consistency, including validity of the string loop and  $\alpha'$  perturbation expansions, successful moduli stabilization, and bounds on higher-derivative terms. In §4.7 we combine the observational and theoretical constraints, with results presented in

figure 4.7.

### 4.1.1 Review of axion monodromy inflation

In this section we will briefly review the motivation for axion monodromy inflation, as well as the most salient phenomenological features. We will postpone until §4.5 a more comprehensive discussion of the realization of this model in string theory.

Inflation is sensitive to Planck-scale physics: contributions to the effective action arising from integrating out degrees of freedom with masses as large as the Planck scale play a critical role in determining the background evolution, and hence the observable spectrum of perturbations (see [111] for a review of this issue). A central problem in inflationary model-building is establishing knowledge of Planck-suppressed terms in the effective action with accuracy sufficient for making predictions. The most elegant solution to this problem is to provide a symmetry that forbids such Planck-suppressed contributions. Because invoking such a symmetry amounts to forbidding couplings of the inflaton to Planck-scale degrees of freedom, it is important to understand this issue in an ultraviolet-complete theory, such as string theory.

One promising mechanism for inflation in string theory involves the shift symmetry of an axion. Axions are numerous in string compactifications and generally enjoy continuous shift symmetries  $a \rightarrow a + \text{constant}$  that are valid to all orders in perturbation theory, but are broken by nonperturbative effects to discrete shifts  $a \rightarrow a + 1$ . As noted in [11], the shift symmetries of axions descending from two-forms are also broken by suitable space-filling fivebranes

(D5-branes or NS5-branes) wrapping two-cycles in the compact space.

In axion monodromy inflation [11], an NS5-brane wrapped on a two-cycle  $\Sigma$  breaks the shift symmetry of the Ramond-Ramond two-form potential  $C_2$ , inducing a potential that is asymptotically linear in the corresponding canonically normalized field  $\phi$ ,

$$V = \mu^3 \phi, \tag{4.1}$$

with  $\mu$  a constant mass scale. Inflation begins with a large expectation value for the inflaton,  $\phi \propto \int_{\Sigma} C_2 \gg 1$ , and proceeds as this expectation value diminishes; note that the NS5-brane, like any D-branes that may be present in the compactification, remains fixed in place during inflation. As argued in [11], this gives rise to a natural model of inflation, with the residual shift symmetry of the axion protecting the potential from problematic corrections that are endemic in string inflation scenarios.

In this paper we perform a careful analysis of the consequences of nonperturbative effects for the axion monodromy scenario. Such effects are generically present: specifically, Euclidean D-branes make periodic contributions to the potential in most realizations of axion monodromy inflation. However, the size of these contributions is model-dependent. It was shown in [11] that there exist classes of examples in which nonperturbative effects are practically negligible, but we expect – as explained in detail in §4.6.5 – that in generic configurations, periodic terms in the potential make small, but not necessarily negligible, contributions to the slow roll parameters.

Therefore, it is of interest to understand the consequences of small periodic modulations of the inflaton potential in axion monodromy inflation. In this paper we address this question in two ways: first, in §4.2-§4.4, by studying a

phenomenological potential that captures the essential effects; and second, in §4.5 and §4.6, by investigating the ranges of the phenomenological parameters that satisfy all known microphysical consistency requirements dictated by the structure of string compactifications in which axion monodromy inflation can be realized.

## 4.2 Background Evolution

In this section we will study the background evolution of the inflaton in the presence of small periodic modulations of the potential. We will focus on modulations in axion monodromy inflation with a linear potential, but our derivations are easily modified to account for other models with a modulated potential. We will denote the size of the modulation by  $\Lambda^4$ , and write our potential as in [11],

$$V(\phi) = \mu^3 \phi + \Lambda^4 \cos\left(\frac{\phi}{f}\right) = \mu^3 \left[ \phi + bf \cos\left(\frac{\phi}{f}\right) \right], \quad (4.2)$$

where we defined the parameter  $b \equiv \frac{\Lambda^4}{\mu^3 f}$ . The equation of motion for the inflaton is then

$$\ddot{\phi} + 3H\dot{\phi} + \mu^3 - \mu^3 b \sin\left(\frac{\phi}{f}\right) = 0. \quad (4.3)$$

To solve (4.3), we begin with two approximations. Monotonicity of the potential requires<sup>1</sup>  $b < 1$ , and as we will see in §4.4, for the case  $b < 1$  observational constraints in fact imply  $b \ll 1$ . This suggests treating the oscillatory term in

---

<sup>1</sup>The case of non-monotonic potentials may also be interesting. On the one hand, for sufficiently large  $b > 1$ , it may be possible to realize chain inflation [112, 113, 114] in our model. In this scenario, the inflaton would tunnel from minimum to minimum, with the universe expanding by less than one third of an e-fold per tunneling event. This requires a more careful analysis, and we will leave this for future studies. On the other hand, for  $b \gtrsim 1$  the model essentially turns into a small-field model of inflation because the inflaton gets trapped at the peaks for a large number of e-folds. It seems hard to distinguish this from other models of small field inflation, but it may be interesting to take a closer look at this as well.

the potential as a perturbation. Furthermore, the COBE normalization implies that  $\phi \gg M_p$  during the era when the modes that are observable in the cosmic microwave background exit the horizon. This allows us to drop terms of higher order in  $M_p/\phi$ .

Under these conditions, it is straightforward to solve for the evolution of the homogeneous background. Expanding the field as  $\phi = \phi_0 + b\phi_1 + O(b^2)$ , the equations of motion of zeroth and first order in  $b$  become

$$\dot{\phi}_0 = -\sqrt{\frac{\mu^3}{3\phi_0}}, \quad (4.4)$$

$$\ddot{\phi}_1 + \sqrt{3\mu^3\phi_0}\dot{\phi}_1 - \frac{\mu^3}{2\phi_0}\phi_1 = \mu^3 \sin\left(\frac{\phi_0}{f}\right), \quad (4.5)$$

where we have neglected terms of higher order in  $M_p/\phi$  and we have made use of the slow roll approximation for  $\phi_0$ .<sup>2</sup> Using equation (4.4), we can rewrite equation (4.5) with  $\phi_0$  as an independent variable instead of  $t$ , yielding

$$\phi_1'' - 3\phi_0\phi_1' - \frac{3}{2}\phi_1 = 3\phi_0 \sin\left(\frac{\phi_0}{f}\right). \quad (4.6)$$

where primes denote derivatives with respect to  $\phi_0$ . For the period of interest, in which the modes now visible in the CMB exit the horizon, it is a good approximation to neglect the motion of  $\phi_0$  everywhere except in the driving term. The inhomogeneous solution is then given by

$$\phi_1(t) = f \frac{6f\phi_*}{(2+3f^2)^2 + 36f^2\phi_*^2} \left[ -(2+3f^2) \sin\left(\frac{\phi_0(t)}{f}\right) + 6f\phi_* \cos\left(\frac{\phi_0(t)}{f}\right) \right], \quad (4.7)$$

where  $\phi_*$  denotes the value of the field  $\phi_0$  at the time at which the pivot scale  $k_*$  exits the horizon. Assuming 60 e-foldings of inflation, this happens around  $\phi_* \simeq$

---

<sup>2</sup>In approximating  $\sin(\phi/f) \simeq \sin(\phi_0/f)$  on the right hand side of (4.5), we have assumed not only that  $b \ll 1$  but also that  $b\phi_1/f \ll 1$ . As we will see from the solution (4.8),  $\phi_1$  is of order  $f^2\phi_*$ . Hence the mild assumption  $b f \phi_* \ll 1$  justifies this approximation.

$11M_p$ . For decay constants  $f$  obeying  $f \gtrsim M_p/10$ , there is less than one oscillation in the range of modes that are observable in the cosmic microwave background, leading to an uninteresting modulation with very long wavelength. We will thus make the additional assumption that  $f \ll M_p$ . Assuming that  $\phi_0 \gg M_p$  and  $f \ll 1$ , using the slow roll approximation for  $\phi_0(t)$ , and working to first order in  $b$ , the solution thus becomes

$$\phi(t) = \phi_0(t) + b\phi_1(t) = \phi_0(t) + bf \frac{3f\phi_*}{1 + (3f\phi_*)^2} \left[ -\sin\left(\frac{\phi_0(t)}{f}\right) + 3f\phi_* \cos\left(\frac{\phi_0(t)}{f}\right) \right], \quad (4.8)$$

with  $\phi_0(t)$  given by

$$\phi_0(t) = \left[ \phi_*^{3/2} - \frac{\sqrt{3}}{2} \mu^{3/2} (t - t_*) \right]^{2/3}. \quad (4.9)$$

In the absence of oscillations, *i.e.* for  $b = 0$ , axion monodromy provides a model of large field inflation that is easily studied using the slow roll expansion. Assuming for concreteness that the CMB scales left the horizon 60 e-foldings before the end of inflation, we are interested in the perturbations around  $\phi_* \simeq 11M_p$ . After imposing the COBE normalization, one finds that CMB perturbations are produced at a scale  $V^{1/4} \simeq 7 \cdot 10^{-3} M_p \simeq 1.7 \cdot 10^{16}$  GeV with a spectral tilt  $n_s \simeq 0.975$  and a tensor-to-scalar ratio  $r \simeq 0.07$ . For reference, the Hubble constant during inflation is then  $H \simeq 2.8 \cdot 10^{-5} M_p \simeq 6.8 \cdot 10^{13}$  GeV.

One can then ask what happens once the oscillations are switched on, *i.e.* when  $b \neq 0$ . It turns out that the effect on the number of e-foldings is negligible as long as  $b \ll 1$ . Hence the inflationary scale is well-approximated by the slow roll analysis. On the other hand, the detailed properties of the perturbations are very different from the slow roll case and cannot be calculated in that expansion. We turn to this issue in the next section.

### 4.3 Spectrum of Scalar Perturbations

Having understood the background evolution, we are now in a position to calculate the power spectrum in axion monodromy inflation. One might be tempted to do this by brute-force numerical calculation, but we find it more instructive to have an analytic result. We will show that under the same assumptions made in calculating the background evolution, *i.e.* slow roll for  $\phi_0(t)$ ,  $\phi_0 \gg M_p$ ,  $f \ll M_p$ , and to first order in  $b$ , the scalar power spectrum is of the form

$$\Delta_{\mathcal{R}}^2(k) = \Delta_{\mathcal{R}}^2(k_*) \left(\frac{k}{k_*}\right)^{n_s-1} \left[1 + \delta n_s \cos\left(\frac{\phi_k}{f}\right)\right] \approx \Delta_{\mathcal{R}}^2\left(\frac{k}{k_*}\right)^{n_s-1 + \frac{\delta n_s}{\ln(k/k_*)} \cos\left(\frac{\phi_k}{f}\right)}, \quad (4.10)$$

where the quantity  $\Delta_{\mathcal{R}}^2(k_*)$  parameterizes the strength of the scalar perturbations and will be introduced in detail in the next subsection. The second equality is valid as long as  $\delta n_s \ll 1$ , and  $\delta n_s$  is given by

$$\delta n_s = \frac{12b}{\sqrt{(1 + (3f\phi_*)^2)}} \sqrt{\frac{\pi}{8} \coth\left(\frac{\pi}{2f\phi_*}\right)} f\phi_*, \quad (4.11)$$

where

$$\phi_k = \sqrt{\phi_*^2 - 2 \ln k/k_*} \simeq \phi_* - \frac{\ln k/k_*}{\phi_*} \quad (4.12)$$

is the value of the scalar field at the time when the mode with comoving momentum  $k$  exits the horizon.

In §4.3.1 we will give a derivation of this result that makes no further approximations. In §4.3.2 and §4.3.3 we will present two additional derivations of (4.10) that are valid only as long as  $f\phi_* \ll 1$  but that lead to a better understanding of the relevant physical effects behind the power spectrum (4.10). Let us at this point briefly summarize the scales that will be relevant for our discussion in the next subsections.

Given the potential (4.2), the time frequency of the oscillations of the inflaton is  $\omega = \dot{\phi}/f$ . This is also the time frequency of the oscillations of the background. Perturbations around this background can be quantized in terms of the solutions of the Mukhanov-Sasaki equation, assuming an asymptotic Bunch-Davies vacuum. Every perturbation mode with comoving momentum  $k$  oscillates with a time frequency  $k/a$  that is redshifted by the expansion of the universe until the mode exits the horizon and freezes when  $k = aH$ .

Then, if  $H < \omega < M_p$ , every mode will at a certain time resonate with the background, as stressed by Chen, Easther, and Lim in [103]. Using the slow roll equation of motion and the COBE normalization,

$$3H\dot{\phi} \simeq -V'(\phi), \quad \dot{\phi}^2 \simeq \frac{2}{3}\epsilon V, \quad V \simeq 5 \cdot 10^{-7} \epsilon M_p^4, \quad (4.13)$$

the requirement  $H < \omega < M_p$  can be re-expressed as

$$\frac{\omega}{H} \simeq \frac{M_p^2}{\phi f} \simeq \sqrt{2\epsilon} \frac{M_p}{f} > 1, \quad (4.14)$$

$$\frac{\omega}{M_p} \simeq \sqrt{\frac{2\epsilon V}{3}} \frac{1}{f M_{pl}} < 1, \quad (4.15)$$

hence defining a range of values for the axion decay constant  $f$  for which resonances occur. Using  $\sqrt{2\epsilon} \simeq M_p/\phi_* \simeq .09$ , we obtain  $2.4 \cdot 10^{-6} < \frac{f}{M_{pl}} < 0.09$ . We will show in §4.5 and §4.6 that  $f$  falls in this range in a class of microphysically well-controlled examples.

Going beyond our approximations, the model also predicts a small amount of running of the scalar spectral index, of order  $10^{-4}$ , from terms of higher order in the  $M_p/\phi$  expansion. Furthermore,  $\delta n_s$  develops a very mild momentum dependence. We will neglect these effects because these will most likely not be observable in current or near-future CMB experiments.

### 4.3.1 Analytic solution of the Mukhanov-Sasaki equation

We begin our study of the spectrum by choosing a gauge such that the scalar field is unperturbed,  $\delta\phi(\mathbf{x}, t) = 0$ , and the scalar perturbations in the spatial part of the metric take the form

$$\delta g_{ij}(\mathbf{x}, t) = 2a(t)^2 \mathcal{R}(\mathbf{x}, t) \delta_{ij}. \quad (4.16)$$

The quantity  $\mathcal{R}(\mathbf{x}, t)$  is a gauge-invariant quantity and in the case of single-field inflation is conserved outside the horizon. It is closely related to the scalar curvature of the spatial slices, but we will not need its precise geometric interpretation at this point.

The translational invariance of the background and thus the equations of motion governing the time evolution of the perturbations make it convenient to look for solutions of the linearized Einstein equations in Fourier space. One defines

$$\mathcal{R}(\mathbf{x}, t) = \int \frac{d^3k}{(2\pi)^{3/2}} \left[ \mathcal{R}_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} \alpha(\mathbf{k}) + \mathcal{R}_k(t)^* e^{-i\mathbf{k}\cdot\mathbf{x}} \alpha^*(\mathbf{k}) \right], \quad (4.17)$$

where  $\mathbf{k}$  is the comoving momentum, and  $k$  is its magnitude. The rotational invariance of the background ensures that  $\mathcal{R}_k(t)$  can depend only on the magnitude of the comoving momentum but not on its direction. Directional dependence can only be contained in the stochastic parameter  $\alpha(\mathbf{k})$  that parameterizes the initial conditions and is normalized so that

$$\langle \alpha(\mathbf{k}) \alpha^*(\mathbf{k}') \rangle = \delta(\mathbf{k} - \mathbf{k}'), \quad (4.18)$$

where the average denotes the average over all possible histories. With this ansatz, the Einstein equations turn into an ordinary differential equation, the

Mukhanov-Sasaki equation, governing the time evolution of  $\mathcal{R}_k(t)$ . We will use it in the form<sup>3</sup>

$$\frac{d^2\mathcal{R}_k}{dx^2} - \frac{2(1+2\epsilon+\delta)}{x} \frac{d\mathcal{R}_k}{dx} + \mathcal{R}_k = 0, \quad (4.19)$$

where  $x \equiv -k\tau$ , with the conformal time  $\tau$  given as usual by  $\tau \equiv \int^t \frac{dt'}{a(t')}$ . Outside the horizon, *i.e.* for  $x \ll 1$  or equivalently  $k/a \ll H$ , the quantity  $\mathcal{R}_k(x)$  approaches a constant which we denote by  $\mathcal{R}_k^{(o)}$ . In terms of  $\mathcal{R}_k^{(o)}$  we define the primordial power spectrum for the scalar modes as

$$|\mathcal{R}_k^{(o)}|^2 = 2\pi^2 \frac{\Delta_{\mathcal{R}}^2(k)}{k^3}. \quad (4.20)$$

To evaluate this quantity, it will again turn out to be sufficient to solve to first order in  $b$ . We therefore expand the slow roll parameters,

$$\epsilon = \epsilon_0 + \epsilon_1 + \mathcal{O}(b^2), \quad (4.21)$$

$$\delta = \delta_0 + \delta_1 + \mathcal{O}(b^2). \quad (4.22)$$

For the background solution (4.8), the first-order terms are given by

$$\epsilon_1 = -\frac{3bf}{\phi_*[1+(3f\phi_*)^2]} \left[ \cos\left(\frac{\phi_0}{f}\right) + (3f\phi_*) \sin\left(\frac{\phi_0}{f}\right) \right], \quad (4.23)$$

$$\delta_1 = -\frac{3b}{[1+(3f\phi_*)^2]} \left[ \sin\left(\frac{\phi_0}{f}\right) - (3f\phi_*) \cos\left(\frac{\phi_0}{f}\right) \right]. \quad (4.24)$$

We now consider an ansatz of the form

$$\mathcal{R}_k = \mathcal{R}_{k,0}^{(o)} \left[ i \sqrt{\frac{\pi}{2}} x^{y_0} H_{y_0}^{(1)}(x) + g(x) \right]. \quad (4.25)$$

---

<sup>3</sup>We use the same definitions for the slow roll parameters as in [115], *i.e.*  $\epsilon \equiv -\frac{\dot{H}}{H^2}$ ,  $\delta \equiv \frac{\ddot{H}}{2H\dot{H}}$ .  $\delta$  is related to the Hubble slow-roll parameters  $\eta \equiv \dot{\epsilon}/\epsilon H$  by  $\delta = \eta/2 - \epsilon$ . The other slow-roll parameters that are sometimes used are  $\epsilon_V \equiv (V'/V)^2/2$  and  $\eta_V \equiv V''/V$ . When the slow roll expansion is valid they are related to the Hubble slow-roll parameters by  $\epsilon_V = \epsilon$  and  $\eta_V = 4\epsilon - \eta$ .

Here the index  $\nu_0$  on the Hankel function,  $H_{\nu_0}^{(1)}(x)$ , is given by  $\nu_0 = \frac{3}{2} + 2\epsilon_0 + \delta_0$ ,  $g(x)$  is a perturbation of order  $b$ , and  $\mathcal{R}_{k,0}^{(o)}$  is the value of  $\mathcal{R}_k(t)$  outside the horizon in the absence of modulations, *i.e.* for  $b = 0$ . To be explicit, it is given by<sup>4</sup>

$$\mathcal{R}_{k,0}^{(o)} = \mp i \sqrt{\frac{\mu^3 \phi_k^3}{6}} \frac{1}{k^{3/2}}, \quad (4.26)$$

where  $\phi_k \approx \phi_* - \frac{\ln k/k_*}{\phi_*}$  once again is the value of the scalar field at the time the mode with comoving momentum  $k$  exits the horizon. The quantity of interest to first order in  $b$  is then

$$|\mathcal{R}_k^{(o)}|^2 = |\mathcal{R}_{k,0}^{(o)}|^2 [1 + 2 \operatorname{Re} g(0)] \approx |\mathcal{R}_{k,0}^{(o)}|^2 e^{2 \operatorname{Re} g(0)} = |\mathcal{R}_{k,0}^{(o)}|^2 \left(\frac{k}{k_*}\right)^{\frac{2 \operatorname{Re} g(0)}{\ln(k/k_*)}}. \quad (4.27)$$

Our ansatz automatically solves the equation of order  $b^0$ . To first order in  $b$  and in the slow roll parameters, the Mukhanov-Sasaki equation leads to an equation for  $g(x)$  of the form

$$\frac{d^2 g}{dx^2} - \frac{2}{x} \frac{dg}{dx} + g = 2e^{ix}(2\epsilon_1 + \delta_1). \quad (4.28)$$

In writing this equation, we have dropped terms of order  $\mathcal{O}(b\epsilon_0, b\delta_0)$ , which amounts to setting  $\nu_0 = 3/2$ . Next, we notice that  $\epsilon_1$  is suppressed relative to  $\delta_1$  by a factor  $\frac{f}{\phi_*}$ . Since we are interested in the regime  $\frac{f}{\phi_*} \ll 1$ , we can thus drop the term proportional to  $\epsilon_1$  on the right hand side of equation (4.28). Furthermore, it turns out to be convenient to rewrite  $\delta_1$  using trigonometric identities. Ignoring an unimportant phase, one finds

$$\delta_1 = -\frac{3b}{\sqrt{1 + (3f\phi_*)^2}} \cos\left(\frac{\phi_0}{f}\right). \quad (4.29)$$

It will be convenient to write  $\phi_0(x)$  as  $\phi_0(x) = \phi_* - \frac{\ln(k/k_*)}{\phi_*} + \frac{\ln x}{\phi_*} = \phi_k + \frac{\ln x}{\phi_*}$ . Introducing  $r(x) \equiv \operatorname{Re}(g(x))$ , equation (4.28) becomes

$$\frac{d^2 r}{dx^2} - \frac{2}{x} \frac{dr}{dx} + r = -\frac{6b}{\sqrt{1 + (3f\phi_*)^2}} \cos(x) \cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right). \quad (4.30)$$

<sup>4</sup>As mentioned earlier, we will ignore the running of the scalar spectral index, but it may be worth pointing out that the information about the running is contained in this formula.

The solution to this equation can be found *e.g.* using Green's functions. We are particularly interested in the inhomogeneous solution at late times, *i.e.* in the limit of vanishing  $x$ . Using more trigonometric identities, we find that the solution in this limit can be brought into the form

$$r(0) = \frac{6b|\mathcal{I}(f\phi_*)|}{\sqrt{1 + (3f\phi_*)^2}} \cos\left(\frac{\phi_k}{f} + \beta(f\phi_*)\right), \quad (4.31)$$

where  $\beta(f\phi_*)$  is an unimportant phase that we will ignore, and  $\mathcal{I}$  is the integral

$$\mathcal{I}(f\phi_*) = \frac{\pi}{2} \int_0^\infty dx J_{\frac{3}{2}}(x) J_{-\frac{1}{2}}(x) x^{\frac{i}{f\phi_*}}. \quad (4.32)$$

Written in this form, the integral can be recognized as a Weber-Schafheitlin integral and can be done analytically (see *e.g.* [116]). One finds

$$|\mathcal{I}| = \sqrt{\frac{\pi}{8} \coth\left(\frac{\pi}{2f\phi_*}\right)} f\phi_*. \quad (4.33)$$

Combining equations (4.27), (4.31) and (4.33), we finally obtain an expression for  $\delta n_s$ ,

$$\delta n_s = \frac{2r(0)}{\cos\left(\frac{\phi_k}{f}\right)} = \frac{12b}{\sqrt{1 + (3f\phi_*)^2}} \sqrt{\frac{\pi}{8} \coth\left(\frac{\pi}{2f\phi_*}\right)} f\phi_*. \quad (4.34)$$

Once again, this derivation is valid to first order in  $b$  and assumes slow roll for  $\phi_0(t)$ ,  $\phi_0 \gg M_p$ , and  $f \ll M_p$ . In particular, it makes no use of an  $f\phi_* \ll 1$  expansion, although this approximation will be needed in the derivations in §4.3.2 and §4.3.3. A comparison between our analytical result for  $\delta n_s$  as a function of  $f\phi_*$  for a fixed value of  $b$  and the result of a numerical calculation using a slight modification of the code described in [117] is shown in Figure 4.1.

### 4.3.2 Saddle-point approximation

As we have seen in the last subsection, it is possible to calculate the power spectrum analytically to first order in  $b$ , assuming slow roll for  $\phi_0(t)$ ,  $\phi_0 \gg M_p$ , and

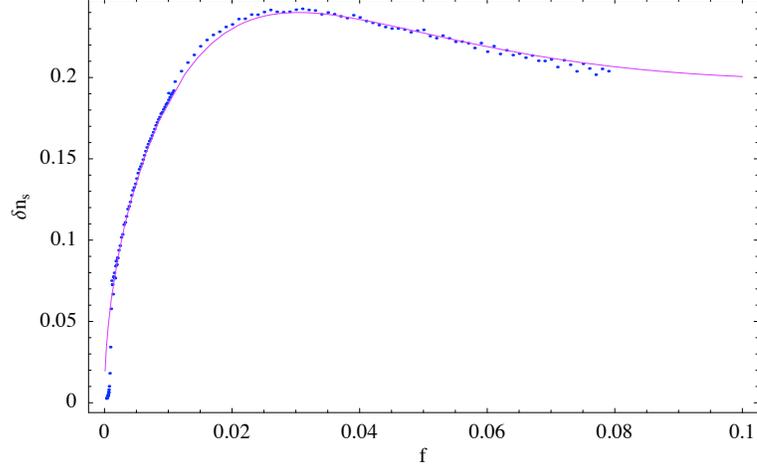


Figure 4.1: The solid line is the analytical result for  $\delta n_s$  as a function of  $f$ , for  $b = 0.08$ , while the dots are the numerical result obtained from an adaptation of the code used in [117].

$f \ll M_p$ , but the derivation sheds little light on the physics behind the results. To get a better understanding, it is instructive to look at the integral (4.32) more explicitly. For this purpose, it is convenient to separate  $\mathcal{I}$  into its real and imaginary parts,  $\mathcal{I} = \mathcal{I}_c + i\mathcal{I}_s$ , with

$$\mathcal{I}_c = \int_0^\infty dx \frac{(\sin x - x \cos x) \cos x}{x^2} \cos\left(\frac{\ln x}{f\phi_*}\right), \quad (4.35)$$

$$\mathcal{I}_s = \int_0^\infty dx \frac{(\sin x - x \cos x) \cos x}{x^2} \sin\left(\frac{\ln x}{f\phi_*}\right). \quad (4.36)$$

For ranges of the axion decay constant such that  $f\phi_* \ll 1$ , these integrals can be done in a stationary phase approximation. Using trigonometric identities to rewrite the products of trigonometric functions appearing in the integrands into sums of trigonometric functions with combined arguments, one finds that the stationary phase occurs at  $\bar{x} = \frac{1}{2f\phi_*}$ . Expanding around the stationary point and

performing the integral as usual, one finds to leading order in  $f\phi_*$

$$\mathcal{I}_c = \sqrt{\frac{\pi}{8}f\phi_*} \sin\left[\frac{1 + \ln(2f\phi_*)}{f\phi_*} - \frac{\pi}{4}\right], \quad (4.37)$$

$$\mathcal{I}_s = \sqrt{\frac{\pi}{8}f\phi_*} \cos\left[\frac{1 + \ln(2f\phi_*)}{f\phi_*} - \frac{\pi}{4}\right], \quad (4.38)$$

which leads to

$$|\mathcal{I}| = \sqrt{\mathcal{I}_c^2 + \mathcal{I}_s^2} = \sqrt{\frac{\pi}{8}f\phi_*}. \quad (4.39)$$

This agrees with our previous result, equation (4.33), as long as  $f\phi_* \ll 1$ . We have not only reproduced our earlier results, however: we also learn that at least for small  $f\phi_*$ , the integral is dominated by a period of time around  $\bar{\tau} = -\frac{1}{2kf\phi_*}$ . Up to the factor of two in the denominator, this corresponds to the period when the frequency of the oscillations of the scalar field background equals the frequency of the oscillations of a mode with comoving momentum  $k$ .<sup>5</sup> The stationary phase approximation thus captures a resonance between the oscillations of the background and the oscillations of the fluctuations, and is good as long as  $f\phi_* \ll 1$ , *i.e.* as long as the resonance occurs while the mode is still well inside the horizon. One might suspect that this has an interpretation in terms of particle production, and we shall make this more precise in what follows.

Recall that our ansatz for  $\mathcal{R}_k$  was given in (4.25), where  $g(x)$  is the solution of the equation

$$\frac{d^2g}{dx^2} - \frac{2}{x} \frac{dg}{dx} + g = 2e^{ix}\delta_1, \quad (4.40)$$

with  $\delta_1$  again given by

$$\delta_1 = -\frac{3b}{\sqrt{1 + (3f\phi_*)^2}} \cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right), \quad (4.41)$$

---

<sup>5</sup>This factor of two can be understood from momentum conservation, as will become clear in §4.3.3.

and initial conditions given by  $\lim_{x \rightarrow \infty} g(x) = 0$  and  $\lim_{x \rightarrow \infty} g'(x) = 0$ . As we have just learned, the effect of the driving term can be ignored long after the resonance has occurred, *i.e.* for  $x \ll \frac{1}{2f\phi_*}$ .<sup>6</sup> This implies that at late times,  $g(x)$  must be a solution of the homogeneous equation which can be written as

$$g(x) = c_k^{(+)} \left( i \sqrt{\frac{\pi}{2}} x^{\frac{3}{2}} H_{3/2}^{(1)}(x) \right) + c_k^{(-)} \left( -i \sqrt{\frac{\pi}{2}} x^{\frac{3}{2}} H_{3/2}^{(2)}(x) \right), \quad (4.42)$$

where  $c_k^{(\pm)}$  are momentum dependent coefficients. The solution for equation (4.40) can also be written explicitly as

$$g(x) = (x \cos x - \sin x) \int_x^\infty \frac{2e^{iy}(\cos y + y \sin y)}{y^2} \delta_1 + (\cos x + x \sin x) \int_x^\infty \frac{2e^{iy}(\sin y - y \cos y)}{y^2} \delta_1. \quad (4.43)$$

For  $x \ll \frac{1}{2f\phi_*}$  we can take the lower limit in the integrals to zero and this can be brought into the form

$$g(x) = \frac{1}{2}(\mathcal{I}^2 + i\mathcal{I}^1) \left( i \sqrt{\frac{\pi}{2}} x^{\frac{3}{2}} H_{3/2}^{(1)}(x) \right) + \frac{1}{2}(\mathcal{I}^2 - i\mathcal{I}^1) \left( -i \sqrt{\frac{\pi}{2}} x^{\frac{3}{2}} H_{3/2}^{(2)}(x) \right), \quad (4.44)$$

where the integrals  $\mathcal{I}^1$  and  $\mathcal{I}^2$  are given by

$$\mathcal{I}^1 = -\frac{6b}{\sqrt{1 + (3f\phi_*)^2}} \int_0^\infty \frac{e^{iy}(\cos y + y \sin y)}{y^2} \cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right), \quad (4.45)$$

$$\mathcal{I}^2 = -\frac{6b}{\sqrt{1 + (3f\phi_*)^2}} \int_0^\infty \frac{e^{iy}(\sin y - y \cos y)}{y^2} \cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right). \quad (4.46)$$

In the saddle point approximation these evaluate to

$$\mathcal{I}^1 = i\mathcal{I}^2 = -\frac{6b}{\sqrt{(1 + (3f\phi_*)^2)}} \sqrt{\frac{\pi}{8}} f\phi_* e^{-i\left(\frac{\phi_k}{f} - \frac{1 + \ln 2f\phi_*}{f\phi_*} + \frac{\pi}{4}\right)}. \quad (4.47)$$

---

<sup>6</sup>One should note that this is not because the driving term goes to zero, but because its frequency becomes too high for the system to keep up with it.

Combining equations (4.25), (4.44), and (4.47), we finally find that the curvature perturbation for  $x \ll \frac{1}{2f\phi_*}$  takes the form

$$\mathcal{R}_k = \mathcal{R}_{k,0}^{(o)} \left( i \sqrt{\frac{\pi}{2}} x^{y_0} H_{\nu_0}^{(1)}(x) - c_k^{(-)} i \sqrt{\frac{\pi}{2}} x^{y_0} H_{\nu_0}^{(2)}(x) \right), \quad (4.48)$$

with  $c_k^{(-)}$  given, up to an unimportant momentum-independent overall phase, by

$$c_k^{(-)} = \frac{6b}{\sqrt{(1 + (3f\phi_*)^2)}} \sqrt{\frac{\pi}{8}} f\phi_* e^{-i(\frac{\phi_k}{f})}. \quad (4.49)$$

One might now interpret the coefficient  $c_k^{(-)}$  of the negative frequency mode as a Bogoliubov coefficient that measures the amount of particles with comoving momentum  $k$  being produced while this mode is in resonance with the background. It seems hard to make this precise as one really is comparing mode solutions of different backgrounds rather than mode solutions of different asymptotically Minkowski regions in the same background.

Equation (4.48) also shows that instead of starting in the Bunch-Davies state and then following the mode through the resonance, one may start the evolution after the resonance has occurred but use a state that is different from the Bunch-Davies state, which is similar to what is considered in [104, 105, 106, 118, 119]. The departure from the Bunch-Davies state is of course quantified by  $c_k^{(-)}$ .

### 4.3.3 Particle production and deviations from the Bunch-Davies state

Here we will deal with a conceptual question that generically arises in inflationary models with oscillations in the scalar potential. Driven by the background

motion of the inflaton, the oscillating contributions constitute a time-oscillating perturbation to the Hamiltonian of the system. Now, perturbations oscillating in time will generically induce transitions, in our case from the original vacuum state to some excited states. This implies that the vacuum state of the full system will deviate from the Bunch-Davies vacuum of the homogeneous background inflationary evolution. We will now estimate the resulting quantity of particle production and relate the result to the derivation of the scalar power spectrum given in the preceding sections.

To lowest order the oscillating perturbation is given by

$$\Delta H^{(2)} = \frac{1}{2} V''(\phi_0(t)) \cdot \delta\phi^2, \quad (4.50)$$

implying that the lowest-order transitions will be from the vacuum  $|0\rangle$  to two-particle states  $|\vec{p}, -\vec{p}\rangle$ . As the physical momentum  $\vec{p} = \vec{k}/a$  corresponding to a given comoving momentum  $\vec{k}$  is exponentially decaying in the inflationary regime, any two-particle state with given comoving momentum  $|\vec{k}, -\vec{k}\rangle$  will be in resonance with the oscillating perturbation only for a short period of time which we will have to estimate in due course.

In transforming the Hamiltonian of the fluctuations into Fourier space

$$H[\delta\phi_{\vec{p}}] = \frac{1}{2} \delta\dot{\phi}_{\vec{p}}^2 + \frac{1}{2} \vec{p}^2 \delta\phi_{\vec{p}}^2 + \frac{1}{2} V''(\phi_0(t)) \cdot \delta\phi_{\vec{p}}^2, \quad (4.51)$$

we find that the system takes the form of a perturbed harmonic oscillator with eigenfrequency  $\omega_p = p \equiv |\vec{p}|$  for each momentum mode  $\delta_{\vec{p}}$  separately,

$$H[\delta\phi_p] = \frac{1}{2} \delta\dot{\phi}_p^2 + \frac{1}{2} \omega_p^2 \delta\phi_p^2 + \underbrace{\frac{1}{2} V''(\phi_0(t)) \cdot \delta\phi_p^2}_{\Delta H^{(2)}[\delta\phi_p]}. \quad (4.52)$$

Now we compare this to the perturbed harmonic oscillator in one-

dimensional quantum mechanics,

$$H = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega_0^2 x^2 + \frac{1}{2}\delta\omega(t)^2 x^2. \quad (4.53)$$

In going to dimensionless variables  $q, p$  we can write this as

$$H = \frac{1}{2}\omega_0 (p^2 + q^2) + \underbrace{\frac{\delta\omega(t)^2}{2\omega_0}}_{\Delta H^{(2)}} q^2, \quad (4.54)$$

where for our case of a periodic perturbation periodic with frequency  $\omega$  we have

$$\delta\omega(t)^2 = \delta\omega^2 \cos(\omega t). \quad (4.55)$$

We want to determine the time-dependent transition matrix element in time-dependent perturbation theory for a periodic perturbation. To do so, we first write the perturbation in standard form for time-dependent perturbation theory as

$$\Delta H^{(2)}(t) = \frac{\delta\omega(t)^2}{2\omega_0} q^2 = \frac{\delta\omega^2}{4\omega_0} q^2 (e^{i\omega t} + e^{-i\omega t}) \equiv F (e^{i\omega t} + e^{-i\omega t}), \quad (4.56)$$

in the notation of equations (40.1) through (40.9) of [120]. The Hamiltonian and the transition matrix elements can be written in terms of creation and annihilation operators  $a$  and  $a^\dagger$  using  $q = (a + a^\dagger)/\sqrt{2}$  and  $p = -i(a - a^\dagger)/\sqrt{2}$ . Then, canonical quantization of the unperturbed part yields a discrete spectrum  $|n\rangle$  of eigenstates with energy spectrum  $E_n = \omega_0(n + 1/2)$ .

If we compare this with our actual case above, we see that for each momentum mode  $\delta\phi_{\vec{p}}, q$  and  $p$  are replaced by appropriate dimensionless fields  $\delta\varphi_{\vec{p}}$  and  $\Pi_{\delta\varphi_{\vec{p}}}$ . In complete analogy to the simple quantum mechanical oscillator, there will be a tower of discrete states  $|n\rangle_p$  with energies  $E_{n,p} = \omega_p(n + 1/2) = p(n + 1/2)$ . In particular,  $|2\rangle_p$  labels the two-particle state  $|p, -p\rangle$  which has energy difference  $\Delta E_{2,p} = 2\omega_p = 2p$  with respect to the ground state. We thus have for the

perturbation in our actual case

$$\Delta H^{(2)}(t) = \frac{\delta\omega(t)^2}{2\omega_p} \delta\varphi_{\vec{p}}^2 = \frac{\delta\omega^2}{4\omega_p} \delta\varphi_{\vec{p}}^2 (e^{i\omega t} + e^{-i\omega t}) \equiv F (e^{i\omega t} + e^{-i\omega t}). \quad (4.57)$$

For the transition matrix element one then finds

$$\langle p, -p | \Delta H^{(2)} | 0 \rangle = F_{20} (e^{i\omega t} + e^{-i\omega t}) \quad \text{with} \quad F_{20} = \frac{\delta\omega^2}{4\omega_p} \langle 0 | \frac{a^2}{\sqrt{2}} \left( \frac{a + a^\dagger}{\sqrt{2}} \right)^2 | 0 \rangle = \frac{\delta\omega^2}{4\sqrt{2}\omega_p} \quad (4.58)$$

Here we have used that

$$|p, -p\rangle = \frac{(a^\dagger)^2}{\sqrt{2}} |0\rangle, \quad (4.59)$$

and

$$\langle 0 | a^2 (a + a^\dagger)^2 | 0 \rangle = \langle 0 | a^2 (a^\dagger)^2 | 0 \rangle = 2. \quad (4.60)$$

If the energy of the two-particle state  $E_{2p} = 2k/a$  were not too close to the perturbation frequency  $\omega$ , we could use time-dependent perturbation theory with the above matrix element and obtain the first order transition probability

$P_{0 \rightarrow 2k}$ ,

$$\begin{aligned} P_{0 \rightarrow 2k} &= \left| -i \int dt' \langle k, -k | \Delta H^{(2)}(t') | 0 \rangle e^{i\omega_{20}t'} \right|^2 \\ &= 2|F_{20}|^2 \frac{\left(\frac{2k}{a}\right)^2 + \omega^2 + \left[\left(\frac{2k}{a}\right)^2 - \omega^2\right] \cos(2\omega t)}{\left[\left(\frac{2k}{a}\right)^2 - \omega^2\right]^2} \\ &= \frac{\delta\omega^4}{16(k/a)^2} \frac{\left(\frac{2k}{a}\right)^2 + \omega^2 + \left[\left(\frac{2k}{a}\right)^2 - \omega^2\right] \cos(2\omega t)}{\left[\left(\frac{2k}{a}\right)^2 - \omega^2\right]^2}, \quad (4.61) \end{aligned}$$

where  $\omega_{20} = E_{2,p=k/a} - E_{0,p=k/a} = 2k/a$ . This gives the resonance line feature characteristic of transition processes.

However, as for any given  $k$  the physical momentum and frequency  $k/a$  will decrease extremely rapidly with  $1/a$ , we can approximate the amount of transition happening in the short time interval  $\Delta t_{res}$  during which the two-particle

state of given  $k$  is in near-resonance  $\omega \approx 2k/a$ . Close to resonance, time-dependent perturbation theory breaks down (visible in the singularity of the above result for  $\omega = 2k/a$ ); however, for periodic perturbations one can solve the Schrödinger equation of the coupled two-state system exactly [120]. One finds that on resonance the transition probability is

$$P_{0 \rightarrow 2k} = \frac{1}{2} [1 - \cos(2\Omega t)] = \frac{1}{2} \left[ 1 - \cos \left( \frac{\delta\omega^2}{2\sqrt{2}k/a} t \right) \right], \text{ where } \Omega \equiv F_{20}. \quad (4.62)$$

That is, near resonance the system effectively oscillates with frequency  $2\Omega = 2 \frac{\delta\omega^2}{4\sqrt{2}\omega_p} = \frac{\delta\omega^2}{2\sqrt{2}k/a}$  between the vacuum and the two-particle state.

We now have to estimate the time  $\Delta t_{res}$  during which a two-particle state of comoving momentum  $k$  stays in near-resonance. We will follow the analysis in [103] and look at the interference terms induced between the  $\cos(\omega t)$  perturbation and the  $\exp(i\omega_{02}t)$  periodicity of the interaction matrix element in (4.61). We note that, on the one hand, the two-particle state with frequency  $2k/a = \omega - \Delta\omega$  stays in resonance with the perturbation with frequency  $\omega$  only for a time roughly estimated to be (for the relative phase shifting from  $-\pi$  to  $\pi$ )

$$\Delta t_1 \sim \frac{2\pi}{\Delta\omega}. \quad (4.63)$$

On the other hand, in the inflating universe it takes very roughly a time

$$\Delta t_2 \sim \frac{2\Delta\omega}{\omega H} \quad (4.64)$$

to change the frequency of the two-particle state from, say,  $\omega + \Delta\omega$  to  $\omega - \Delta\omega$ . Equating the two provides us with the effective duration of near-resonance,

$$\Delta t_{res} \equiv \Delta t_1 = \Delta t_2 \sim 2 \sqrt{\pi \frac{H}{\omega} H^{-1}}. \quad (4.65)$$

Plugging this into the above transition result and remembering that near resonance  $k/a \approx \omega/2$ , we get

$$P_{0 \rightarrow 2k} \simeq \frac{1}{2} \left[ 1 - \cos \left( \sqrt{2} \frac{\delta\omega^2}{\omega} \sqrt{\pi \frac{H}{\omega} H^{-1}} \right) \right]. \quad (4.66)$$

Now, in our case above we see that  $p = k/a$  in terms of comoving momenta  $k$ , and further

$$\delta\omega(t)^2 = V''(\phi_0(t)) = \frac{\Lambda^4}{f^2} \cos\left(\frac{\phi_0(t)}{f}\right) = \delta\omega^2 \cos(\omega t) \quad , \quad \omega = \frac{H}{f\phi_0} \quad . \quad (4.67)$$

Noting that in our scenario of interest we have  $H < \omega$  and that  $\delta\omega \ll H$ , we can expand the argument of the cosine around zero. If we then plug in the microscopic definitions of the quantities  $\delta\omega^2 = \Lambda^4/f^2$  and  $\omega = H/(f\phi)$ , we get

$$P_{0 \rightarrow 2k} \simeq \frac{\pi}{2} \frac{\Lambda^8}{f^4 \omega^2 H^2} \cdot \frac{H}{\omega} = \frac{\pi}{2} \frac{\Lambda^8 f \phi_*^3}{f^2 H^4} = \frac{\pi}{2} \frac{9\Lambda^8 f \phi_*^3}{f^2 \mu^6 \phi_*^2} = \frac{36\pi}{8} b^2 f \phi_* \quad , \quad (4.68)$$

where  $\phi_* \simeq 11M_p$  denotes the vev of the inflaton field around 60 e-foldings before the end of inflation.

Next, because  $P_{0 \rightarrow 2k}$  characterizes the transition probability to the two-particle states, it may be related to the negative frequency Bogoliubov coefficient  $c^{(-)}$  that relates the out-vacuum to the in-vacuum. Specifically, the out-vacuum is specified by the modes

$$u_k(out) = c^{(-)} u_{-k} + c^{(+)} u_k \quad , \quad (4.69)$$

whereas the original Bunch-Davies in-vacuum had modes

$$u_k(in) = u_k \quad , \quad c^{(+)}(in) = 1 \quad . \quad (4.70)$$

We therefore find that

$$|c^{(-)}| \simeq \sqrt{P_{0 \rightarrow 2k}} = 6b \sqrt{\frac{\pi}{8} f \phi_*} \quad . \quad (4.71)$$

In comparing these results with the general treatment of the Mukhanov-Sasaki equation above, we see by looking at (4.48) and (4.27) that we can identify

$$u_k = i \sqrt{\frac{\pi}{2}} x^{y_0} H_{v_0}^{(1)}(x) \quad (4.72)$$

$$u_{-k} = -i \sqrt{\frac{\pi}{2}} x^{y_0} H_{v_0}^{(2)}(x) \quad (4.73)$$

and thus from (4.27) we conclude that

$$\delta n_s = \frac{2\text{Re } g(x)}{\cos\left(\frac{\phi_k}{f}\right)} \underset{x \rightarrow 0}{=} \frac{2\text{Re}(c^{(+)}c^{(-)})}{\cos\left(\frac{\phi_k}{f}\right)} \simeq 2|c^{(-)}| \simeq 12b \sqrt{\frac{\pi}{8} f \phi_*} \quad (4.74)$$

which agrees with the general result (4.34) in the appropriate limit  $\omega > H$  and  $\delta\omega \ll H$ , corresponding to  $f\phi_* < 1$ , where  $\coth(\pi/2f\phi_*) \rightarrow 1$ .

Note that in calculating the transition probability we lose information about the phase of the transition matrix element as given in (4.58). Therefore, if we estimate the population coefficient  $c^{(-)}$  from  $\sqrt{P_{0 \rightarrow 2k}}$ , we get only an estimate for  $|c^{(-)}|$  without the phase information. A more complete derivation using the full information in the transition matrix element should also yield the information about the phase as derived in the previous subsection.

Thus, we see that in the regime of rapid oscillations,  $f\phi_* < 1$ , the induced  $\delta n_s$  is due to a time-localized deviation from the Bunch-Davies state, which may be interpreted as being due to resonant bursts of particle production happening well before a given mode leaves the horizon during inflation.

### 4.3.4 Bispectrum of scalar perturbations

We start by reviewing how resonance can drive the production of large non-Gaussianity during inflation, as proposed in [103]. We then present an estimate for the size of the non-Gaussianity for the model (4.2).

The three-point function can be calculated as [121]

$$\langle \mathcal{R}(\tau, \mathbf{k}_1) \mathcal{R}(\tau, \mathbf{k}_2) \mathcal{R}(\tau, \mathbf{k}_3) \rangle = -i \int_{\tau_0}^{\tau} \langle [\mathcal{R}(\tau, \mathbf{k}_1) \mathcal{R}(\tau, \mathbf{k}_2) \mathcal{R}(\tau, \mathbf{k}_3), H_I(\tau')] \rangle a d\tau', \quad (4.75)$$

where  $H_I$  is the interacting part of the Hamiltonian.  $H_I$  was calculated for a

generic potential (see e.g. [103, 121]) at cubic order in the perturbations; it takes the form

$$H_I = - \int d^3x \left[ a\epsilon^2 \mathcal{R}\mathcal{R}'^2 + a\epsilon^2 \mathcal{R}(\partial\mathcal{R})^2 - 2\epsilon\mathcal{R}'(\partial\mathcal{R})(\partial\chi) + \frac{a}{2}\epsilon\eta'\mathcal{R}^2\mathcal{R}' + \frac{\epsilon}{2a}(\partial\mathcal{R})(\partial\chi)(\partial^2\chi) + \frac{\epsilon}{4a}(\partial^2\mathcal{R})(\partial\chi)^2 \right], \quad (4.76)$$

where  $\partial$  denote space derivatives,

$$\chi \equiv a^2\epsilon\partial^{-2}\dot{\mathcal{R}}, \quad (4.77)$$

and we used the Hubble slow-roll parameter  $\eta \equiv \dot{\epsilon}/(\epsilon H) = 2(\epsilon + \delta)$  because formulas in this subsection are simpler in terms of  $\eta$  than in terms of  $\delta$ .

We would like to stress that (4.76) is exact for arbitrary values of the slow roll parameters  $\epsilon$  and  $\eta$ . Substituting  $H_I$  into (4.75) produces six terms, plus an additional term coming from a field redefinition. For the modulated linear potential (4.2),  $\epsilon$  is small, as in standard slow roll inflation. On the other hand, contrary to the standard slow-roll approximation,  $\eta$  can be much larger than  $\epsilon^2$ . This suggests that the leading term comes from the  $\epsilon\eta$  term in the Hamiltonian.<sup>7</sup> Hence we have [103, 107]

$$\langle \mathcal{R}(t, \mathbf{k}_1)\mathcal{R}(t, \mathbf{k}_2)\mathcal{R}(t, \mathbf{k}_3) \rangle \simeq i \left( \prod_i u_i(\tau_{end}) \right) \times \int_{-\infty}^{\tau_{end}} d\tau \epsilon \eta' a^2 \left( u_1^*(\tau)u_2^*(\tau) \frac{d}{d\tau} u_3^*(\tau) + \text{sym} \right) \delta^3(\mathbf{K})(2\pi)^3 + c.c.. \quad (4.79)$$

As in [103], we parameterize the non-Gaussianity as

$$\langle \mathcal{R}(\tau, \mathbf{k}_1)\mathcal{R}(\tau, \mathbf{k}_2)\mathcal{R}(\tau, \mathbf{k}_3) \rangle \equiv \frac{G(k_1, k_2, k_3)}{(k_1 k_2 k_3)^3} \delta^3(\mathbf{K}) \Delta_{\mathcal{R}}^4(2\pi)^7, \quad (4.80)$$

<sup>7</sup>In (3.9) of [121] this term was written as

$$\frac{\dot{\phi}^2}{\dot{\rho}^2} e^{3\rho} \dot{\mathcal{R}} \mathcal{R}^2 \frac{d}{dt} \left( \frac{\ddot{\phi}}{2\dot{\phi}\dot{\rho}} + \frac{\dot{\phi}^2}{4\dot{\rho}^2} \right), \quad (4.78)$$

which can be reduced to the term in (4.76) using  $H' = -\dot{\phi}^2/2$ .

where  $\mathbf{K} = \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$ . We take as an ansatz for the shape of the non-Gaussianity for our modulated linear potential

$$\frac{G(k_1, k_2, k_3)}{k_1 k_2 k_3} = f_{res} \sin\left(\frac{2}{\phi f} \ln K + \text{phase}\right) \quad (4.81)$$

Following [103] and comparing (4.79), (4.80) and (4.81), we obtain the estimate

$$f_{res} \simeq \frac{3 \dot{\eta}_1}{8H \sqrt{\phi f}}, \quad (4.82)$$

where we have again used the notation  $\eta = \eta_0 + b\eta_1 + \dots$ . Using the background solution obtained in §4.2, it is straightforward to find

$$\dot{\eta}_1 \simeq 2\dot{\delta}_1 \simeq -\sqrt{\frac{\mu^3}{3\phi_* f[1 + (3f\phi_*)^2]}} \left[ \cos\left(\frac{\phi_0}{f}\right) + (3f\phi_*) \sin\left(\frac{\phi_0}{f}\right) \right]. \quad (4.83)$$

It is not hard to convince oneself that in the region of parameter space where  $f_{res} > 1$  and  $b \ll 1$ , the second term in (4.83) is always negligible, *i.e.*  $3f\phi \ll 1$ . Hence our estimate for the non-Gaussianity is

$$f_{res} \simeq \frac{9b}{4(f\phi)^{3/2}} = \frac{9}{4} b \left(\frac{\omega}{H}\right)^{3/2}. \quad (4.84)$$

where we remind the reader that  $\omega = \dot{\phi}/f$ . As we will often refer to this equation, let us pause and comment on it. The resonant non-Gaussianity vanishes when the modulation is switched off, *i.e.* for  $b = 0$ . It is inversely proportional to some power of  $f$  (depending on which quantity is held fixed). Hence the smaller the axion decay constant  $f$ , the larger the non-Gaussianity. On the other hand, as we will see in §4.5, there are theoretical lower (as well as upper) bounds on  $f$ , so that the non-Gaussian signal cannot be made arbitrarily large.

No complete analysis of the observational constraints on resonant non-Gaussianity has been performed to date (however, see [122]), and such an analysis is beyond the scope of the present work. Based on a rough comparison with

known shapes of non-Gaussianity, we estimate that  $f_{res} \gtrsim 200$  might be at the borderline of being excluded by the current data, while  $f_{res} \lesssim 1$  would be difficult to detect in the next generation of experiments. A comprehensive analysis of the detectability of resonant non-Gaussianity is a very interesting topic for future research.

#### 4.4 Observational Constraints

In the last section, we derived the theoretical predictions of axion monodromy inflation for the primordial power spectrum. We will now use these predictions to compare the model with the five-year WMAP data [12]. While the data in principle allows for a variety of statistics to be extracted, we will limit ourselves to the most fundamental one, the angular power spectrum. The reason for this is that the data is not now adequate for the polarization data or the three-point correlations to place meaningful additional constraints on the model. This will change as soon as the Planck data becomes available, and will be an interesting problem especially given the unusual shape of the non-Gaussianities the model predicts.

For the benefit of the less cosmologically-inclined reader, we now briefly summarize the basic observables relevant to our analysis. In the ideal scenario, in which a full-sky map is available, the temperature of the cosmic microwave background as a function of the position in the sky can be expanded in spherical harmonics as

$$T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}) \quad (4.85)$$

The theoretical counterparts of these measured expansion coefficients, which

we will denote  $a_{\ell m}^{th}$ , should be thought of as random variables satisfying a (possibly only nearly) Gaussian distribution. Each realization of these coefficients corresponds to a possible history of the universe. In the Gaussian case, all the information about the theory is contained in the two-point correlations of these, as the odd n-point functions vanish, and the even n-point functions are sums of products of the two-point functions. Assuming an isotropic background, the two-point correlations must take the form

$$\langle a_{\ell m}^{th} a_{\ell' m'}^{th*} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'} , \quad (4.86)$$

where the brackets denote an average over all possible histories or equivalently (by the ergodic theorem) all possible positions. The  $a_{\ell m}^{th}$ 's themselves, being random variables encoding initial conditions, cannot be predicted from a given cosmological model, and only the multipole coefficients,  $C_\ell$ , encoding their correlations are of interest. These multipole coefficients  $C_\ell$  can be estimated from the measured expansion coefficients  $a_{\ell m}$  via

$$C_\ell^{\text{sky}} = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}|^2 . \quad (4.87)$$

For noiseless, full-sky CMB data, these provide an unbiased estimate of the true power spectrum in the sense that the average of the analogously defined quantity for the  $a_{\ell m}^{th}$ 's satisfies

$$\langle C_\ell^{\text{sky,th}} \rangle \equiv \frac{1}{2\ell + 1} \left\langle \sum_m |a_{\ell m}^{th}|^2 \right\rangle = C_\ell . \quad (4.88)$$

Since there are only  $2\ell + 1$  modes per  $\ell$ , even for the ideal noiseless full-sky map the estimate of the multipole coefficient has the cosmic variance uncertainty

$$\left\langle \left( \frac{C_\ell^{\text{sky,th}} - C_\ell}{C_\ell} \right)^2 \right\rangle = \frac{2}{2\ell + 1} . \quad (4.89)$$

In a more realistic setting with noise and sky cuts, this estimator is no longer unbiased and more sophisticated estimators have to be used. The current state

of the art is to use a pixel-based maximum likelihood estimator for low  $\ell$  (specifically, for  $\ell \leq 32$ ), and a pseudo- $C_\ell$  estimator for higher  $\ell$ . For details we refer the reader to [123] and references therein.

After this quick review of the basic relevant quantities, let us describe our analysis. We work on a grid of model parameters. For each point on the grid, we compute the theoretical angular power spectrum with the publicly-available CAMB code [124, 125]<sup>8</sup>, using the primordial power spectrum derived in the previous section in the form

$$\Delta_{\mathcal{R}}^2(k) = \Delta_{\mathcal{R}}^2(k_*) \left( \frac{k}{k_*} \right)^{n_s - 1 + \frac{\delta n_s}{\ln(k/k_*)} \cos\left(\frac{\phi_k}{f} + \Delta\varphi\right)}. \quad (4.90)$$

The likelihood for a given theoretical power spectrum is calculated with a modified version of the WMAP five-year likelihood code that is now available on the LAMBDA webpage [126]. The power spectrum in our model contains additional parameters beyond those of the WMAP five-year  $\Lambda$ CDM fit (namely,  $\{\Omega_B h^2, \Omega_c h^2, \Omega_\Lambda, \tau, n_s, \Delta_{\mathcal{R}}^2\}$ ) and the marginalization parameter  $\{A_{SZ}\}$ . The additional parameters are  $\delta n_s$ ,  $f$  and a phase  $\Delta\varphi$ . This phase parameterizes both our uncertainty in the number of e-folds needed, which originates in our poor understanding of reheating, and a microscopically determined phase offset in the sinusoidal modulation of the scalar potential arising in the string theory construction.

We fix the value of the scalar spectral index  $n_s = 0.975$ . As in any model of large-field inflation, the spectral index is a prediction of the model that depends only on the physics of reheating and, correspondingly, on the total amount of inflation since the observable modes exited the horizon. The value we choose

---

<sup>8</sup>Of course, we modify the CAMB code to calculate all the multipole coefficients rather than calculating some and interpolating.

corresponds to the situation in which the pivot scale exits the horizon 60 e-folds before the end of inflation. The results turn out to be fairly independent of the precise value chosen for the scalar spectral index and we could have chosen the value corresponding to any number of e-folds between 50 and 60. We fix  $\{\Omega_c h^2, \Omega_\Lambda, \tau, A_{SZ}\}$  to the WMAP five-year best-fit values for the  $\Lambda$ CDM fit. We allow  $f, \delta n_s, \Omega_B h^2, \Delta\varphi$  to vary on the grid, and we also marginalize over the scalar amplitude  $\{\Delta_{\mathcal{R}}^2\}$  in the likelihood code. To obtain Figure 4.2, we thus marginalize over  $\{\Omega_B h^2, \Delta_{\mathcal{R}}^2\}$  and over the unknown phase  $\Delta\varphi$ , while we fix  $\{\Omega_c h^2, \Omega_\Lambda, \tau, A_{SZ}\}$ , as we expect at most mild degeneracies between these parameters and the primordial ones.

The grid consists of 16 equidistantly spaced points in  $\Omega_B h^2$  between  $\Omega_B h^2 \approx 0.0212$  and  $\Omega_B h^2 \approx 0.0266$ , 128 equidistantly spaced points in  $\delta n_s$  between  $\delta n_s = 0$  and  $\delta n_s = 0.44$ , 512 logarithmically spaced points in the axion decay constant  $f$  between  $f = 9 \times 10^{-5}$  and  $f = 10^{-1}$ , as well as 32 points for the phase  $\Delta\varphi$  between  $\Delta\varphi = -\pi$  and  $\Delta\varphi = \pi$ . This leads to a grid with a total of 33,554,432 points. The analysis was run on 64 of the compute nodes of the Ranger supercomputer at the Texas Advanced Computing Center. The compute nodes are SunBlade x6420 blades, and each of the nodes provides four AMD Opteron Quad-Core 64-bit processors with a core frequency of 2.3 GHz.

The resulting 68% and 95% contours in the  $\delta n_s - f$  plane are shown in the left plot of Figure 4.2. To convert the resulting observational constraints on  $\delta n_s$  as a function of  $f$  into constraints on the microscopic parameter  $bf$  as a function of  $f$ , we make use of equation (4.34). The resulting 68% and 95% contours in the  $bf-f$  plane are shown in the right plot of Figure 4.2. Roughly, the results can be summarized as  $bf \lesssim 10^{-4}$  for  $f \lesssim 0.01$  at 95% confidence level. Our best

fit point is at a rather small value of the axion decay constant,  $f = 6.67 \times 10^{-4}$ , and a rather large amplitude for the oscillations,  $\delta n_s = 0.17$ . The fit improves by  $\Delta\chi^2 \simeq 11$  over the fit in the absence of oscillations. The corresponding angular power spectrum is shown in Figure 4.3.

The improvement can be traced to a better fit to the data around the first peak. We would like to stress, however, that we do not take this as an indication of oscillations in the observed angular power spectrum. Similar spikes in the likelihood function occur quite generally when fitting an oscillatory model to toy data generated with the conventional power spectrum without any oscillations, because the oscillations fit some features in the noise. The polarization data could provide a cross check, but we find that it is presently not good enough to do so in a meaningful way.

Let us say a few words motivating the necessity of marginalizing over  $\Omega_B h^2$  and  $\Delta\varphi$ . There is a known degeneracy in the angular power spectrum between  $\Omega_B h^2$  and  $n_s$ , as changing  $\Omega_B h^2$  changes the ratio of the power in the first and second acoustic peaks, which to some extent can be undone by changing the spectral tilt  $n_s$ . In our case we do not vary  $n_s$ , but we add a sinusoidal contribution to the standard power spectrum. It is intuitively clear that by doing so we can change the ratio of power in the first and second acoustic peak by choosing the right oscillation frequency (controlled by  $f$ ) and phase  $\Delta\varphi$ , leading to a degeneracy between  $\Omega_B h^2$  and  $\delta n_s$  at least for a certain range of  $f$ .

The most straightforward way to demonstrate this degeneracy between  $\Omega_B h^2$  and  $\delta n_s$  arising for certain ‘resonant’ values of  $f$  is to present a likelihood plot in the  $\Omega_B$ - $\delta n_s$  plane for a value of  $f$  for which the degeneracy is clearly visible. An example is shown in the plot on the left side of Figure 4.4. It shows that

marginalizing over  $\Omega_b h^2$  is necessary to obtain correct exclusion contours on  $\delta n_s$  and  $f$ . That marginalization over the phase is necessary can easily be seen from a likelihood plot in the  $\delta n_s$ - $\Delta\phi$  plane. This is shown in the plot on the right side of Figure 4.4.

We have also performed a Markov chain Monte Carlo analysis for the model using the publicly available CosmoMC code [127], [128]. While the Monte Carlo has the advantage that it is less computationally intensive than a grid when varying all cosmological parameters, the likelihood function for oscillatory models turns out to be rather spiky, making the Monte Carlo hard to set up, because the chains tend to get trapped in the spikes.

To some extent this can be overcome by taking out the problematic regions or increasing the temperature of the Monte Carlo. When run on parts of the parameter space where the Monte Carlo runs reliably, we found agreement with the grid-based results shown above.

The most problematic direction to sample is that of the axion decay constant,  $f$ . We show the result of one of our chains for  $f = 0.01$  in Figure 4.5. The plot shows marginalized one-dimensional distributions and two-dimensional 68% and 95% confidence level limits for the most important  $\Lambda$ CDM parameters as well as  $\delta n_s$  and  $\Delta\phi$ . In the Monte Carlo, we sampled the parameters  $\delta n_s$ ,  $\Delta\phi$ , all parameters of the  $\Lambda$ CDM except the scalar spectral index, as well as the Sunyaev-Zel'dovich amplitude.

## 4.5 Microphysics of Axion Monodromy Inflation

In §1.1 we briefly reviewed the properties of axion monodromy inflation, focusing on the description in effective field theory. For a general characterization of the signatures of the scenario, the phenomenological model of §1.1 was sufficient. However, the phenomenological parameters  $f, \mu, \Lambda$  are in principle derivable from the data of a string compactification, and as such they obey non-trivial microscopic constraints: the ranges and correlations of these parameters are restricted by microphysics.

We should therefore determine the values of the phenomenological parameters allowed in consistent, computable string compactifications. We will begin by reviewing the string theory origin of axion monodromy inflation, both to set notation and to highlight the properties most relevant in constraining the parameters  $f, \mu, \Lambda$ . For concreteness we will restrict our attention to a specific realization of the scenario, in O3-O7 orientifolds of type IIB string theory, with the Kähler moduli stabilized by nonperturbative effects. Our considerations could be generalized to other compactifications, but the numerical results would differ.

### 4.5.1 Axions in string theory

Let us first review the origin of the relevant axions. Our conventions and notation are summarized in appendix B.1. Consider type IIB string theory compactified on an orientifold of a Calabi-Yau threefold  $X$ . Let the forms  $\omega^J$  be a basis of the cohomology  $H^2(X, \mathbb{Z})$ , normalized such that  $\int_{\Sigma_I} \omega^J = \delta_I^J (2\pi)^2 \alpha'$ , where  $\Sigma_I$

are a basis of the dual homology  $H_2(X, \mathbb{Z})$ . The RR two-form  $C_2$  gives rise to a four-dimensional axion via the ansatz<sup>9</sup>

$$C_2 = \frac{1}{2\pi} c_I(x) \omega^I, \quad (4.91)$$

where  $x$  is a four-dimensional spacetime coordinate. The ten-dimensional Einstein-frame action [129] that follows is

$$\int d^{10}x \frac{g_s \sqrt{-g_E}}{2(2\pi)^7 \alpha'^4} |dC_2|^2 = \int d^{10}x \frac{g_s \sqrt{-g_E}}{12(2\pi)^9 \alpha'^4} g_E^{\mu\nu} \partial_\mu c_I \partial_\nu c_J \omega_{ij}^I \omega_{i'j'}^J g_E^{ii'} g_E^{jj'}. \quad (4.92)$$

Notice that the axions only have derivative couplings, and hence enjoy a continuous shift symmetry at the level of the classical action. In §4.5.2 we will recall the origin of this symmetry and explain how it persists to all orders in perturbation theory and is broken by nonperturbative effects.

Upon dimensional reduction, one finds a relation between the four-dimensional reduced Planck mass  $M_p$  and  $\alpha'$ ,

$$\alpha' M_p^2 = \frac{\mathcal{V}_E}{\pi}, \quad (4.93)$$

where  $\mathcal{V}_E$  is the Einstein-frame (dimensionless) volume of the Calabi-Yau  $X$  measured in units of  $l_s \equiv 2\pi \sqrt{\alpha'}$ . The decay constant of the canonically normalized axion is then

$$\frac{f^2}{M_p^2} = \frac{g_s}{48\pi^2 \mathcal{V}_E} \left[ \frac{\int \omega \wedge * \omega}{(2\pi)^6 \alpha'^3} \right]. \quad (4.94)$$

The present definition of the axion decay constant differs by a factor of  $2\pi$  from that in [11], *i.e.*  $f^{\text{here}} = 2\pi f^{\text{there}}$ . As a consequence our canonically normalized axion has periodicity  $2\pi f$ , consistent with (4.2).

---

<sup>9</sup>The factor of  $2\pi$  is introduced so that the four-dimensional axions  $c_I$  have periodicity  $2\pi$ , as can be seen via S-duality from the world-sheet coupling  $i \int B_2 / (2\pi\alpha')$ . Notice that in our conventions  $C_2$  and  $\omega^I$  have the dimensions of length-squared, while  $c_I$  are dimensionless.

## 4.5.2 Dimensional reduction and moduli stabilization

### Four-dimensional data of O3-O7 orientifolds

Now we consider how to stabilize the compactification in a setup that will allow inflation. We focus on the KKLT scenario for moduli stabilization [36]. We assume that the complex structure moduli, the dilaton, and any open string moduli have been stabilized at a higher scale, and we concentrate on the remaining closed string moduli (specifically, the remaining moduli are those descending from hypermultiplets). The  $\mathcal{N} = 1$  supersymmetric four-dimensional theory resulting from dimensional reduction of type IIB orientifolds was worked out in detail in [130]. We are interested in orientifold actions under which the holomorphic three-form  $\Omega$  of the Calabi-Yau manifold is odd, so that the fixed-point loci are O3-planes and O7-planes. The cohomology decomposes into eigenspaces of the orientifold action,

$$H^{(r,s)} = H_+^{r,s} \oplus H_-^{r,s}. \quad (4.95)$$

We therefore divide the basis  $\omega_A, A = 1, \dots, h^{1,1}$  into  $\omega_\alpha, \alpha = 1, \dots, h_+^{1,1}$  and  $\omega_a, a = 1, \dots, h_-^{1,1}$ . Working out the sign of the orientifold action on the physical fields, one finds that the two-forms  $C_2$  and  $B_2$  are odd, and should be expanded in terms of the  $\omega_a$ . Grimm and Louis [130] have derived the Kähler coordinates on the corresponding moduli space, *i.e.* the proper complex combinations of fields that appear as the lowest components of chiral multiplets:<sup>10</sup>

$$G^a \equiv \frac{1}{2\pi} \left( c^a - i \frac{b^a}{g_s} \right) \quad (4.96)$$

$$T_\alpha \equiv i\rho_\alpha + \frac{1}{2} c_{\alpha\beta\gamma} v^\beta v^\gamma + \frac{g_s}{4} c_{abc} G^b (G - \bar{G})^c \quad (4.97)$$

<sup>10</sup>We use the same notation as [130] with two exceptions: we rescale  $T_\alpha$  as  $T_\alpha^{here} = (2/3)T_\alpha^{there}$ , and we add a factor of  $(2\pi)^{-1}$  in the definition of  $G^a$  such that the fields  $c^a$  and  $b^a$  have periodicity  $2\pi$ . See appendix B.1 for more details on our conventions.

where  $\rho_\alpha$  comes from the RR four-form  $C_4$  integrated over some orientifold-even four-cycle  $\Sigma_\alpha$ , with  $\alpha = 1, \dots, h_+^{1,1}$ ; and  $c^a$  and  $b^a$  come from the RR and NS-NS two-forms  $C_2$  and  $B_2$  integrated over some orientifold odd two-cycle  $\Sigma_a$  with  $a = 1, \dots, h_-^{1,1}$ . The tree-level Kähler potential is given by<sup>11</sup>

$$K = \log\left(\frac{g_s}{2}\right) - 2 \log \mathcal{V}_E \quad (4.98)$$

where the (dimensionless) Einstein-frame volume  $\mathcal{V}_E$  of the Calabi-Yau manifold is defined in (B.4). The dependence of this Kähler potential on the multiplets (4.96) and (4.97) cannot be written down explicitly for a generic choice of the intersection numbers  $c_{IJK}$ . The implicit dependence is given by writing the (Einstein-frame) volume in terms of two-cycle volumes  $v^\alpha$

$$K = \log\left(\frac{g_s}{2}\right) - 2 \log \left[ \frac{1}{6} c_{\alpha\beta\gamma} v^\alpha(T, G) v^\beta(T, G) v^\gamma(T, G) \right]. \quad (4.99)$$

Then one has to solve (4.97) for  $v^\alpha$  and substitute the result into the above Kähler potential. The Kähler potential is a function of  $\mathcal{V}_E$ , and hence is a function of  $v^\alpha$ , and in turn of  $\tau_\alpha \equiv \text{Re } T_\alpha$  and  $\text{Im } G$ , but does not depend on  $\text{Re } G$  and  $\text{Im } T_\alpha$  (as can be seen by taking the real part of (4.97)). One might be tempted to conclude that  $c$  enjoys a shift symmetry but that  $b$  does not, but, as we will explain in §4.5.2, both fields have shift symmetries.

The tree-level superpotential  $W_0$  does not depend on the multiplets (4.96) and (4.97). In fact it depends on the complex structure moduli and the dilaton, which we assume have already been stabilized by fluxes. Therefore we will take  $W_0$  to be a discretely tunable constant.

---

<sup>11</sup>We assume that the axio-dilaton  $\tau = C_0 + ie^{-\phi}$  is already stabilized by fluxes at  $\tau = i/g_s$  and we write down the dilaton-dependent part of the Kähler potential only to keep track of factors of  $g_s$ .

## Nonperturbative stabilization of the Kähler moduli

Let us now proceed to consider nonperturbative effects. We follow the KKLT strategy [36] for the construction of a de Sitter vacuum. We assume that each four-cycle  $T_\alpha$  is wrapped either by a Euclidean D3-brane or by a stack of D7-branes giving rise to a four-dimensional gauge theory that undergoes gaugino condensation.<sup>12</sup> This results in the following four-dimensional superpotential:

$$W = W_0 + \sum_{\alpha=1}^{h_+^{1,1}} A_\alpha e^{-a_\alpha T_\alpha}, \quad (4.100)$$

where  $A_\alpha$  will be treated as constants, as they depend on the complex structure moduli, which we have assumed to be stabilized;  $a_\alpha \equiv 2\pi/N_\alpha$ , with  $N_\alpha$  the number of D7-branes in the stack; and  $N_\alpha = 1$  for the case of a Euclidean D3-brane. We can find a supersymmetric minimum by solving for the vanishing of all the F-terms: for the  $h_+^{1,1}$  even Kähler moduli via

$$\begin{aligned} 0 &= D_\alpha W \equiv \partial_{T_\alpha} W + W \partial_{T_\alpha} K = -A_\alpha a_\alpha e^{-a_\alpha T_\alpha} - 2W \frac{\partial_{T_\alpha} \mathcal{V}_E}{\mathcal{V}_E} \\ &= -A_\alpha a_\alpha e^{-a_\alpha T_\alpha} - W \frac{v^\alpha}{2\mathcal{V}_E}, \end{aligned} \quad (4.101)$$

and for the  $h_-^{1,1}$  odd moduli via

$$0 = D_a W \equiv \partial_{G^a} W + W \partial_{G^a} K = -iW \frac{c_{\alpha ac} v^\alpha b^c}{4\pi \mathcal{V}_E} \quad (4.102)$$

where in both cases in the last step we used the chain rule and the definitions of  $G^a$  and  $T_\alpha$  in terms of two-cycle volumes  $v^\alpha$ . The condition (4.101) is simplified if we first solve for  $\text{Im } T_\alpha$ , which gives

$$a_\alpha \text{Im } T_\alpha = \theta_{A_\alpha} - \theta_{W_0} + k_\alpha \pi, \quad k_\alpha \in \mathbb{Z}. \quad (4.103)$$

<sup>12</sup>In general, Euclidean D3-branes or D7-branes will wrap some linear combinations  $\tilde{T}_\alpha$  of the cycles appearing in (4.97), rather than the basis cycles  $T_\alpha$  themselves, but for simplicity we will suppress this issue.

Then we are left with the set of real equations for each  $\alpha$ ,

$$(\pm 1)_\alpha |A_\alpha| a_\alpha e^{-a_\alpha \tau_\alpha} = \partial_{T_\alpha} K \left( |W_0| + \sum_\beta (\pm 1)_\beta |A_\beta| e^{-a_\beta \tau_\beta} \right), \quad (4.104)$$

where  $(\pm 1)$  depends on the value of  $k$  in (4.103). As long as the orientifold-even four-cycle Kähler moduli are defined as in (4.97), then  $\partial_{T_\alpha} K = -v^\alpha / (2\mathcal{V}_E) < 0$  for every  $\alpha$ . Now we prove that in (4.104) the minus sign has to be chosen for every  $\alpha$  in order to have a supersymmetric solution. First we notice that the sign of the right hand side does not depend on  $\alpha$ , so  $k_\alpha$  and hence  $(\pm 1)_\alpha$  have to be the same for every  $\alpha$ . If we choose the positive sign in (4.104), the quantity in brackets in the right hand side is manifestly positive. Then the two sides of the equation have opposite signs and no (compact) solution exists. To summarize, the minimization of  $\text{Im } T_\alpha$  boils down to taking all  $A_\alpha$  real and negative and  $W_0$  real and positive or the other way around.<sup>13</sup>

Concerning (4.102), an obvious solution is given by  $b^a = 0$  for every  $a$ . As argued in [11], an inflationary model with a  $b$ -type axion as the inflaton will generically suffer from an eta problem, and we will therefore focus on a  $c$ -type axion.

## Nonperturbative breaking of axionic shift symmetries

Axionic shift symmetries are central to this paper, so we will now explain how they originate and how they are ultimately broken by nonperturbative effects.

First, let us recall the classic result [131, 132] establishing the shift symmetry to

---

<sup>13</sup>We notice that if one chooses as Kähler variable a linear combination of the  $T_\alpha$  defined in (4.97), as is done *e.g.* in the large volume scenario with Swiss-cheese Calabi-Yau manifolds, then the sign of  $\partial_{T_\alpha} K$  can depend on  $\alpha$ . In this case, the minimization of  $\text{Im } T_\alpha$  boils down to taking  $W_0$  real and positive and  $A_\alpha$  real with  $\text{Sign}(A_\alpha) = \text{Sign}(\partial_{T_\alpha} K)$ , up to multiplying  $W$  by an overall phase.

all orders in perturbation theory. Consider the axion  $b = \int_{\Sigma} B/(2\pi\alpha')$ , where  $B$  is the NS-NS two-form potential and  $\Sigma$  is a two-cycle in the Calabi-Yau manifold. The vertex operator representing the coupling of  $b$  to the string worldsheet is [131]

$$V(k) = \frac{1}{2\pi\alpha'} \int_{\Sigma} d^2\xi \exp(ik \cdot X(\xi)) \epsilon^{\alpha\beta} \partial_{\alpha} X^{\mu} \partial_{\beta} X^{\nu} B_{\mu\nu}(X). \quad (4.105)$$

At zero momentum, this coupling is seen to be a total derivative in the worldsheet theory. Therefore, the axion  $b$  can only have derivative couplings (which vanish at zero momentum), to any order in sigma-model perturbation theory. Notice that the genus of the worldsheet did not enter in this argument, so the axion shift symmetry is also valid to all orders in string perturbation theory.

This argument fails in the presence of worldsheet boundaries (*i.e.*, D-branes), and also fails once worldsheet instantons, or D-brane instantons, are included. In axion monodromy inflation, both sorts of breaking play an important role, as we shall now explain.

First, the introduction of an NS5-brane wrapping a curve  $\Sigma_a$  creates a monodromy for the axion  $c_a$ , spoiling its shift symmetry and inducing an asymptotically linear potential [11]. Specifically, the potential induced by the Born-Infeld action of the NS5-brane (obtained by S-dualizing the Born-Infeld action of a D5-brane) is

$$V(c_a) = \frac{\epsilon}{g_s(2\pi)^5 \alpha'^2} \sqrt{\ell^4 + (2\pi g_s c_a)^2}, \quad (4.106)$$

where  $\ell \sqrt{\alpha'}$  is the size of  $\Sigma_a$  and  $\epsilon$  captures the possibility of suppression due to warping. For  $c_a \gg 1$ , this potential is linear in  $c_a$ , or in the corresponding canonically normalized field, which we denoted by  $\phi$  in the preceding sections. Let us remark that the square root form of the potential can be important at the end of inflation and also makes a small change in the number of e-foldings

produced for given parameter values, so that in a model that includes a specific scenario for reheating, the square root structure should be incorporated as well. As we have not invoked a concrete reheating scenario, for our purposes the linear potential suffices, but one must still bear in mind that this form is not valid for small  $\phi$ .

As we will explain in detail, the D-brane instantons involved in moduli stabilization introduce sinusoidal modulations to the linear potential. We will work exclusively in a regime in which the breaking by wrapped branes dominates over the nonperturbative breaking, although we remark in passing that the complementary regime might be interesting for realizing models involving repeated tunneling.

The breaking of the  $b$  shift symmetry by Euclidean D-branes (or by gaugino condensation on D7-branes) is slightly subtle, so we will address it briefly. As we remarked above,  $b$  appears quadratically in the classical Kähler potential, which seems to contradict the statement that it enjoys a shift symmetry at the perturbative level in the absence of boundaries. However, there is no contradiction: the shift symmetry of  $b$  is true at constant two-cycle volumes  $v$  and *not* at constant four-cycle volumes  $T$ . To see this, suppose that there is a single Kähler modulus  $T$ , so that the superpotential is of the form (4.100) with  $h_+^{1,1} = 1$ . The Kähler potential is then [130]

$$K = -3M_p^2 \log(T + \bar{T} - db^2), \quad (4.107)$$

with  $d$  a constant. In the absence of a nonperturbative superpotential, a suitable simultaneous shift of  $T + \bar{T}$  and  $b$  is a symmetry of the scalar potential of this system; under such a shift, the two-cycle volumes  $v$  are invariant. However, this symmetry is spoiled by the nonperturbative term in  $W$ , because the

superpotential and the scalar potential are no longer invariant. Therefore, in a scenario in which the four-cycle volumes are stabilized nonperturbatively, the  $b$  axion receives a mass in a stabilized vacuum.

At this stage the mass-squared  $m_b^2$  of  $b$  is proportional to the vacuum energy and hence is negative in the supersymmetric AdS minimum. The minimum of the potential will be the final point of the inflationary dynamics, and hence we would like it to have a very small *positive* cosmological constant to be consistent with the current accelerated expansion of the universe. Thus, we need to include an uplifting term. In the uplifted minimum,  $m_b^2 \propto V_{dS} > 0$ . This relation is the origin of the eta problem that was found in [11] choosing  $b$  as an axion: for a generic uplifting,<sup>14</sup>  $V''(b) \sim V(b)$ , so that  $\eta \sim \mathcal{O}(1)$  and slow roll inflation does not take place. This is completely analogous to the eta problem of D-brane inflation found in [5] and can be intuitively understood in the same way. Here we will take  $b = 0$  as the stabilized value<sup>15</sup> of  $b$  and concentrate on  $c$  as a candidate inflaton.

Let us now turn to consider  $c$ , which does not appear in the Kähler potential or superpotential at any order in perturbation theory. To assess  $c$  as an inflaton, one should determine the leading nonperturbative effects, either in the superpotential or in the Kähler potential, that do introduce a potential for  $c$ , *i.e.* one should identify the leading breaking of the shift symmetry. Euclidean D3-branes carrying vanishing D1-brane charge do not induce a potential for  $c$ , but Euclidean D3-branes supporting worldvolume fluxes (and hence nonvanishing D1-brane charge) give rise to a dependence on  $c$ , via the Chern-Simons coupling

---

<sup>14</sup>Notice that the proportionality constant in  $m_b^2 \propto V_{dS}$  depends on the volume-dependence of the uplifting term and could be made small for particular choices of the latter as proposed in [35].

<sup>15</sup>It is easy to check that  $b = 0$  is still the stabilized value after the inclusion of nonperturbative corrections to the Kähler potential, *cf.* §4.6.5.

$\int F_2 \wedge C_2$ . As observed in [11], it follows that when the Kähler moduli are stabilized by Euclidean D3-branes,  $c$  receives a mass in the stabilized vacuum: one must sum over Euclidean D-brane contributions to the superpotential, including summing over the amount  $n = \int F_2$  of magnetization, and this generically introduces an eta problem for  $c$ . The solution, as explained in [11], is to stabilize the Kähler moduli via gaugino condensation on D7-branes, which leads to an exponentially smaller (and hence negligible) mass for  $c$ .

### 4.5.3 Axion decay constants in string theory

We now turn to the important task of expressing the axion decay constant,  $f$ , in terms of the data of a compactification. As we reviewed in §4.5.1, the decay constant of an axion  $C_2 = c(x)\omega/(2\pi)$  is given by

$$\frac{f^2}{M_p^2} = \frac{g_s}{48\pi^2 \mathcal{V}_E} \left[ \frac{\int \omega \wedge * \omega}{(2\pi)^6 \alpha'^3} \right], \quad (4.108)$$

so that the primary task is to compute the norm  $\int \omega \wedge * \omega$ . (This problem has been studied in a wide range of examples in [133].) We will first recall, in §4.5.3, how to express the axion kinetic term, and hence also the axion decay constant, in terms of  $\mathcal{N} = 1$  data. This will lead us to a simple expression for the decay constant in terms of intersection numbers of the Calabi-Yau. We will then propose a class of models in which the decay constant is rather small, motivated by the fact that with other parameters held fixed, decreasing  $f$  increases the amplitude of the resonant non-Gaussianity. Next, in §4.5.3, we will present a concrete example that illustrates the geometry of a configuration that leads to small  $f$ .

## Decay constants in terms of $\mathcal{N} = 1$ data

In §4.5.2 we have reviewed, following [130], the four-dimensional  $\mathcal{N} = 1$  description of Type IIB O3-O7 orientifolds. The multiplets relevant for us are the orientifold-odd chiral multiplets  $G^a$  and the orientifold-even chiral multiplets  $T_\alpha$ . The tree-level Kähler potential given in (4.99) determines the kinetic terms for  $G^a$  and hence the decay constants of the axions  $b^a$  and  $c^a$ . First let us notice that the Kähler metric in the space of the chiral multiplets  $T_\alpha$  and  $G^a$  factorizes in two blocks,  $K_{T_\alpha \bar{T}_\beta}$  and  $K_{G^a \bar{G}^b}$ . The reason is that off-diagonal terms such as  $K_{T_\alpha \bar{G}^a}$  are proportional to intersection numbers  $c_{\alpha\beta a}$  with one odd index and two even indices, which are forbidden by the orientifold action [130]. We are interested in one particular mode from among the  $G^a$ , which we will denote by  $G^-$ ;  $\Sigma_-$  is then the orientifold-odd two-cycle that supports our candidate inflaton  $c^-$ . We now choose a basis for  $G^a$  such that  $K_{G^a \bar{G}^b}$  is block diagonal with a  $1 \times 1$  block  $K_{G^- \bar{G}^-}$ . The kinetic term for  $c^-$  is then given by

$$-\frac{1}{2}f^2 (\partial c^-)^2 = M_p^2 K_{G^- \bar{G}^-} \frac{1}{(2\pi)^2} (\partial c^-)^2 \subset M_p^2 K_{G^- \bar{G}^-} |\partial G^-|^2, \quad (4.109)$$

where

$$K_{G^- \bar{G}^-} = \frac{\partial^2 K(G, T)}{\partial G^- \partial \bar{G}^-} = -g_s \frac{c_{\alpha--} v^\alpha}{4\mathcal{V}_E}, \quad (4.110)$$

and we used

$$c_{\alpha\beta\gamma} v^\beta v^\gamma = 2\tau_\alpha + g_s c_{\alpha bc} \text{Im } G^b \text{Im } G^c. \quad (4.111)$$

Hence we can express the decay constant of the axion  $c^-$  as

$$\frac{f^2}{M_p^2} = \frac{g_s}{8\pi^2} \frac{c_{\alpha--} v^\alpha}{\mathcal{V}_E}. \quad (4.112)$$

As promised, we have expressed the norm  $\int \omega \wedge * \omega$  in terms of the intersection numbers

$$\frac{\int \omega \wedge * \omega}{(2\pi)^6 \alpha'^3} = \frac{2}{3} c_{\alpha--} v^\alpha. \quad (4.113)$$

In §4.6.4 we will discuss the constraints that follow from the result (4.112). First, in the following subsection we provide some geometrical intuition for (4.112).

### An example: a complex plane of fixed points

An instructive example arises from considering an orbifold that is locally  $\mathbb{C}^2/\mathbb{Z}_2 \times \mathbb{C}$ , *i.e.* an Eguchi-Hanson space fibered over a base  $\Sigma$  of complex dimension one. Let  $\omega$  be the two-form dual to the blowup cycle of the orbifold, and let  $\Sigma$  be the two-manifold of fixed points, *i.e.* the base over which the Eguchi-Hanson space is fibered.<sup>16</sup> We are interested in the decay constant of the axion  $C_2 = \frac{1}{2\pi}c(x)\omega$ , so we must compute  $\int \omega \wedge *\omega$ . In the local approximation, this is straightforward, as we shall see. However, far from the fixed-point locus, the fiber may deviate substantially from the Eguchi-Hanson geometry, in a complicated and model-dependent way, and moreover the fixed-point locus  $\Sigma$  may be embedded in the compact space in a nontrivial manner. Happily, the integral  $\int \omega \wedge *\omega$  has its primary support near the fixed-point locus, where the local approximation is excellent.

We recall, following the useful summary in Appendix B of [134], that the Eguchi-Hanson space has a unique homology two-cycle of radius  $a/2$ , where  $r = a$  defines the location of the coordinate singularity; here  $r$  is the standard radial coordinate. The two-form  $\omega$  corresponding to this cycle may be written

$$\omega = 2\pi\alpha' \frac{a^2}{r^2} \left( \frac{dr}{r} \wedge d\psi + \cos\theta \frac{dr}{r} \wedge d\phi + \frac{1}{2} \sin\theta d\theta \wedge d\phi \right) \quad (4.114)$$

in terms of  $r$  and the angular coordinates  $\psi, \theta, \phi$ . By observing that  $*_4\omega = -\omega$

---

<sup>16</sup>Concretely, we are imagining that  $\Sigma$  extends into a warped throat region, and that an NS5-brane wraps the blowup cycle at a particular location in the throat. The warping is invoked in order to suppress the energy density of the wrapped NS5-brane. See [11] for further details, and for an example of a suitable orbifold action in a Klebanov-Strassler throat.

and that  $\int \omega \wedge \omega = -(2\pi)^4 \alpha'^2 / 2$ , one finds

$$\frac{\int_{EH} \omega \wedge *_4 \omega}{(2\pi)^4 \alpha'^2} = \frac{1}{2}. \quad (4.115)$$

Clearly, given the form of  $\omega$ , this integral has its support in a region  $a \leq r \lesssim f_{ew} \times a$ . This justifies the local approximation as long as the compact space has a radius that is large compared to  $a$ . Next, we observe that

$$\frac{\int \omega \wedge *_6 \omega}{(2\pi)^6 \alpha'^3} = \frac{\int_{EH} d^4 x \omega \wedge *_4 \omega}{(2\pi)^4 \alpha'^2} \times \frac{\int_{\Sigma} \sqrt{g}}{(2\pi)^2 \alpha'} = \frac{1}{2} \text{Vol}(\Sigma) \quad (4.116)$$

Substituting this in (4.108), we recover the parametric scaling of §4.5.3.

## 4.6 Microscopic Constraints

We now turn to determining the ranges of our phenomenological parameters that are allowed in a consistent and computable microphysical model.

Let us first remark that, as usual in string theory model building, computability imposes stringent constraints on the compactification parameters. Because large-field inflation involves substantial energy densities and requires correspondingly steep moduli barriers, the compact space needs to be reasonably small, so that the Kaluza-Klein scale and the (necessarily lower) scale of moduli masses can be large enough to prevent runaway moduli evolution. Clearly, one must then carefully check that the compactification is still large enough for the supergravity approximation to be valid; furthermore, backreaction of the inflationary energy on the compact space is a serious issue, particularly when this space is not large in string units. Incorporating these requirements then leads to severe restrictions on the allowed values of the decay constant  $f$ .

We will begin in §4.6.1 by considering the constraints from computability, then give, in §4.6.2, a qualitative description of the constraints from backreaction, deferring details to appendix B. Next, in §4.6.3, we will verify that a two-derivative action suffices to describe this system. This is not obvious, as rapid oscillations in the potential could enhance the importance of generic higher-derivative terms; however, we will show that the specific terms emerging from string theory are negligible in our solution. We then apply these constraints in §4.6.4 to determine the range of the decay constant  $f$ . Finally, we estimate the size  $bf$  of the modulations; as this is rather model-dependent, in §4.6.5, we will restrict our attention to a specific example in which a periodic contribution is generated by Euclidean D1-brane corrections to the Kähler potential.

### 4.6.1 Constraints from computability

In this subsection we will list several constraints coming from the consistency of the string theory setup. We will first require the validity of the string and  $\alpha'$  perturbation expansions, and the validity of neglecting higher-order corrections to the nonperturbative superpotential, and then we will require that the inflaton potential does not destabilize the compactification.

First of all we require the validity of string perturbation theory, *i.e.* we require  $g_s \ll 1$ . We must also ensure the validity of the  $\alpha'$  expansion. To do this including a reasonable *estimate* of numerical factors such as  $2\pi$ , it is convenient to use worldsheet instantons as a proxy for perturbative  $\alpha'$  corrections, because the normalization is easily determined. To get the correct coefficient, we start from the string-frame ten-dimensional metric  $g_{string}$  and impose that the world-

sheet instanton action obeys  $e^{-S_{ws}} \lesssim e^{-2}$ , or

$$\frac{1}{2\pi\alpha'} \int \sqrt{g_{string}} \gtrsim 2, \quad (4.117)$$

which using  $g_{string} = g_{Einstein} \sqrt{g_s}$  is converted to Einstein frame

$$2 < \frac{\sqrt{g_s}}{2\pi\alpha'} \int_{\Sigma^\alpha} J = \sqrt{g_s} v^\alpha 2\pi \Rightarrow v^\alpha > \frac{1}{\pi \sqrt{g_s}}, \quad (4.118)$$

and we used that  $\int_{\Sigma^\alpha} \omega^\beta = (2\pi)^2 \alpha' \delta_\alpha^\beta$ .

As we invoked nonperturbative corrections to the superpotential, we must also require that any further superpotential corrections, *e.g.* from multi-instantons, are negligible. For this purpose it suffices to impose

$$e^{-a_\alpha T_\alpha} < e^{-2} \ll 1 \Rightarrow \tau_\alpha > \frac{N_\alpha}{\pi}. \quad (4.119)$$

Additional constraints come from the moduli stabilization process. To use the single-field inflationary analysis we have developed in §4.2 and §4.3, we need to require that the uplifted minimum is only slightly perturbed by the inflationary dynamics. In particular, the linear potential that we have represented as  $\mu^3 \phi$  actually depends on the compactification volume, and hence shifts the minimized value of the volume. In four-dimensional Einstein frame, the leading term in the inflaton potential is

$$V(\phi, \mathcal{V}_E) \approx \left( \frac{\langle \mathcal{V}_E \rangle}{\mathcal{V}_E} \right)^2 \mu^3 \phi \quad (4.120)$$

where  $\langle \mathcal{V}_E \rangle$  is the expectation value of the volume. To ensure that the resulting contribution to the potential for the volume is unimportant, we will insist that the inflaton potential induced by the NS5-brane,  $V(\phi)$ , is smaller than the moduli potential  $\mathcal{U}_{mod}$ .

At the supersymmetric minimum we have

$$V_{AdS} = -\frac{g_s}{2} \frac{3|W|^2}{\mathcal{V}_E^2}. \quad (4.121)$$

Without specifying the details of the uplifting mechanism, we assume that an uplifting to a small and positive cosmological constant is possible, and that the height of the potential barrier  $\mathcal{U}_{mod}$  that separates the uplifted minimum from decompactification is of the same order as  $\mathcal{U}_{mod} \sim |V_{AdS}|$ . Now, the COBE normalization tells us that

$$V(\phi_{CMB}) = \epsilon \left( 0.027 M_p \right)^4 \simeq 2.4 \cdot 10^{-9} M_p^4. \quad (4.122)$$

Hence we obtain the constraint

$$\frac{g_s}{2} \frac{3|W|^2}{\mathcal{V}_E^2} = |V_{AdS}| \simeq \mathcal{U}_{mod} \gg 2.4 \cdot 10^{-9} M_p^4. \quad (4.123)$$

To extract a useful form of the above constraints, let us substitute for  $W$  the solution of any of the equations (4.101)

$$W = +|A_\alpha| a_\alpha e^{-a_\alpha \tau_\alpha} \frac{2\mathcal{V}_E}{v^\alpha}, \quad (4.124)$$

with no sum over  $\alpha$ . We will also assume  $|A_\alpha| \sim 1$  (see [11] for a discussion of this point). After some manipulations we find

$$\tau_\alpha \ll -\frac{N_\alpha}{2\pi} \log \left( N_\alpha 10^{-5} \frac{v^\alpha}{\pi \sqrt{g_s}} \right), \quad (4.125)$$

again with no summation over  $\alpha$ . Finally, we should limit the number of D7-branes in each stack; although there plausibly exist examples with  $N_\alpha$  quite large, we will impose  $N_\alpha \leq 50$ . This gives us

$$\tau_\alpha \ll 73 - 8 \log \left( \frac{v^\alpha \pi \sqrt{g_s}}{2g_s} \right). \quad (4.126)$$

We notice that  $v^\alpha (\pi \sqrt{g_s}) > 1$  was the condition in (4.118) that enabled us to neglect  $\alpha'$  corrections, so that as long as  $g_s \leq 0.5$  the second term on the right hand side of (4.126) is negative.

## 4.6.2 Constraints from backreaction on the geometry

Another important constraint comes from the requirement that the backreaction of the inflationary energy density on the compact space is small. In this section we will give a qualitative description of the problem and will briefly sketch a model-building solution; the interested reader is referred to appendix B for a more complete treatment.

At the time that the CMB perturbations are produced, the inflaton has a large vev in Planck units,  $\phi \sim 11M_p$ , corresponding to a configuration of the two-form potential threading the two-cycle  $\Sigma_-$  of the form

$$\frac{1}{(2\pi)^2\alpha'} \int_{\Sigma_-} C_2 \equiv N_w = \frac{\phi}{2\pi f} \gg 1 \quad (4.127)$$

In the absence of an NS5-brane wrapping  $\Sigma$ , there would be no energy stored in this configuration, as  $C_2$  enjoys a shift symmetry. However, inflation is driven by the substantial energy stored in this system by the Born-Infeld action of the wrapped NS5-brane. Moreover, there is a corresponding D3-brane charge induced by the Chern-Simons coupling  $\int C_2 \wedge C_4$ . Note that the net induced D3-brane charge in the total compactification is zero, as required by Gauss's law, because we have arranged for an additional, tadpole-canceling NS5-brane that wraps a distant cycle  $\Sigma'_-$  homologous to  $\Sigma_-$ , but does so with opposite orientation. Therefore, the Chern-Simons coupling induces a dipole configuration of D3-brane charge, with  $F_5$  flux lines stretching from  $\Sigma_-$  to  $\Sigma'_-$ .

It is essential to ensure that the inflationary energy, which is effectively localized in the compact space in the vicinity of the wrapped NS5-brane, does not substantially correct the remainder of the compact geometry. Heuristically, one can imagine that the increased tension of the NS5-brane, as well as the induced

charge, is represented by  $N_w$  D3-branes dissolved in the NS5-brane. We must therefore estimate the effect of  $N_w$  D3-branes in a warped throat (recall that we have situated each wrapped NS5-brane in a warped region in order to suppress its energy density below the string scale, as required *e.g.* by the COBE normalization). Clearly, this backreaction will be reasonably small if  $N_w \ll N$ , with  $N$  the D3-brane charge of the background throat.

However, we must be careful about the effect of even a modest distortion of the geometry on the moduli stabilization and therefore on the four-dimensional potential. Let us first recall that in scenarios of D3-brane inflation in nonperturbatively-stabilized vacua, even a single D3-brane moving slowly in a throat can affect the warp factor, and correspondingly the warped volumes of four-cycles bearing nonperturbative effects, to such a degree that this interaction is the leading contribution to the inflaton potential [17, 135].

This sensitivity originates in two facts: first, D3-branes perturb the warped metric in a manner that is not suppressed by the background warp factor at the location of the D3-branes, because D3-branes are BPS with respect to a throat generated by D3-brane charge, and hence their contributions to the metric may simply be superposed on the background. Second, nonperturbative effects on a four-cycle are exponentially sensitive to changes in the four-cycle volume. Both these facts appear threatening for a situation such as ours in which the moduli are stabilized nonperturbatively and substantial D3-brane charge is induced in a throat: one can anticipate that as inflation proceeds and the D3-brane charge diminishes, the four-cycle volume changes, leading to an unanticipated, and possibly steep, contribution to the inflaton potential.

To understand this concretely, we will first consider a simpler system: an

anti-D3-brane in a warped throat generated by  $N$  D3-branes, or equivalently a warped throat generated by  $N - 1$  D3-branes, together with a brane-antibrane pair. Furthermore, from the result of [136] one learns that at long distances, the effect of the brane-antibrane pair on the supergravity solution is strongly suppressed by the warp factor at the location of the pair, *i.e.* at the tip of the throat. In contrast, the effects of D3-branes are not suppressed in this manner. Therefore, for the purpose of computing perturbations to the bulk compact space, we may replace an anti-D3-brane in a warped throat generated by  $N$  D3-branes with a warped throat generated by  $N - 1$  D3-branes, up to exponentially small corrections.

Equipped with this approximation, we may represent the configuration of interest as follows: two warped throats, carrying the charge of  $N_1, N_2$  D3-branes respectively, are perturbed to  $N_1 + N_w, N_2 - N_w$  by the inclusion of the NS5-brane in (say) the first throat, and the anti-NS5-brane in the second throat. Here we are ignoring the warping-suppressed correction indicated above, and we are approximating the NS5-branes by the D3-brane charge and tension that they carry, which is an excellent approximation for  $N_w \gg 1$ . Other effects due to the NS5-brane that do not depend on its induced D3-brane charge, *i.e.* on its world-volume flux, are independent of the inflaton and hence do not correct its potential. One can now easily see that the volume of a four-cycle at a generic location in the compact space will be corrected by the inclusion of the NS5-branes. If the four-cycle happens to enter one or both throats, the change in the volume is easily computed, and is seen to be substantial (*cf.* appendix B).

To control this problem, we situate the NS5-brane and the anti-NS5-brane, together with the family of homologous cycles connecting them, in a single

warped region. The idea is that from the bulk of the compact space, the NS5-brane configuration will appear to be a distant dipole whose net effect, integrated over a four-cycle, averages out to be small. This setup allows us to parametrically suppress the backreaction by a small factor given by the ratio of the dipole length, *i.e.* the distance between two NS5-branes, to the distance between the NS5-branes and the four-cycle in question. This small factor comes in addition to the suppression by the small ratio  $N_w/N$ .<sup>17</sup>

In appendix B.2 we give more details about the above setup. We show, through two explicit models of increasing complexity, the robustness of the above suppression mechanism.

### 4.6.3 Constraints from higher-derivative terms

The analysis presented thus far has used the two-derivative action, which is an approximation with a limited range of validity. In general, one expects an infinite series of higher-derivative terms, possibly including multiple derivatives as well as powers of the first derivative. Our background solution involves rapid oscillations, so it is reasonable to ask whether these high frequencies enhance the role of higher-derivative terms and render the two-derivative approximation invalid. To check this, one should evaluate the higher-derivative terms on the solution and compare to the two-derivative action. We will now show that the two-derivative approximation is valid in the scalar sector; analogous considerations apply to the gravitational action.

---

<sup>17</sup>A further suppression can be achieved with a carefully-chosen embedding of the four-cycle, *e.g.* one that is symmetric with respect to the two NS5-branes. However, this requires fine-tuning, whereas the dipole suppression on which we have focused is parametric.

Rather than write down the most general higher-derivative corrections to the scalar sector, we give here the terms that end up being present in the string theory examples. In string theory, we can directly compute the leading higher-derivative terms in the action for  $b$ , extending the result to  $c$  using S-duality. To get the leading terms, one considers the  $\alpha'^3$  corrections to the effective action due to Gross and Sloan [137] (at the four-point level) and Kehagias and Partouche [138] (up to the eight-point level). These corrections are of the same lineage as the famous Riemann<sup>4</sup> term, but involve NS-NS three-form flux. This yields corrections to the axion kinetic terms. Following [138], the ten-dimensional Einstein-frame action including the leading ( $\alpha'^3$ ) corrections is

$$S_{10D,E} = \frac{1}{(2\pi)^7 \alpha'^4} \int d^{10}x \sqrt{g_E} \left( R_E - \frac{1}{12g_s} H_{KMN} H^{KMN} + \frac{\zeta(3)}{3 \times 2^6 g_s^{-3/2}} \alpha'^3 \bar{R}^4 + \dots \right) \quad (4.128)$$

where

$$\bar{R}_{MN}{}^{PQ} = R_{MN}{}^{PQ} + \frac{1}{2} g_s^{-1/2} \nabla_{[M} H_{N]}{}^{PQ} - \frac{1}{4} g_s^{-1} H_{[M}{}^{CIP} H_{N]C}{}^{QI} + \dots \quad (4.129)$$

and the square brackets are defined without the combinatorial factor 1/2 in front. Hence, the terms that are relevant for our axion at order  $\alpha'^3$  are proportional to  $H^8$  and  $(\nabla H)^4$ . To estimate the importance of these terms, we will consider a special case in which the internal space is a  $T^2 \times T^4$ , with the NS-NS two-form field only along the  $T^2$  directions 8 and 9, *i.e.*  $B_{89} = -B_{98} = b$ . Furthermore, since the background dynamics involves large frequencies but not large spatial gradients, we are primarily interested in terms containing only time derivatives, and can therefore take  $b$  to be homogeneous in the noncompact spatial directions. In this special case, making use of (2.13) in [137], and using S-duality to determine the action for  $c$  from that for  $b$ , we find that after dimensional reduction the corrected action for  $c$  is

$$S_{4D} = \int d^4x \left[ -M_p^2 \frac{g_s}{2} \dot{c}^2 g^{88} g^{99} + \frac{\zeta(3)}{2^6 g_s^{3/2}} \frac{\mathcal{V}_E^3}{\pi^3 M_p^4} \left( \frac{1}{2} \dot{c}^8 g_s^4 (g^{88} g^{99})^4 + \frac{1}{2^4} \dot{c}^4 g_s^2 (g^{88} g^{99})^2 \right) \right] \quad (4.130)$$

Now we use  $\phi = cf$  to make the kinetic term canonical, yielding the action in terms of  $\phi$ ,

$$\begin{aligned} S_{4D} &= \int d^4x \left[ -\frac{1}{2} \dot{\phi}^2 + \frac{\zeta(3)}{2^6 g_s^{3/2}} \frac{\mathcal{V}_E^3}{\pi^3} \left( \frac{1}{2} \frac{\dot{\phi}^8}{M_p^{12}} + \frac{1}{2^4} \frac{\ddot{\phi}^4}{M_p^8} \right) \right] \\ &\equiv \int d^4x \left[ -\frac{1}{2} \dot{\phi}^2 + \frac{\dot{\phi}^8}{M_I^{12}} + \frac{\ddot{\phi}^4}{M_{II}^8} \right], \end{aligned} \quad (4.131)$$

where we can now calculate the scale of the higher derivative terms  $M_I$  and  $M_{II}$  to be

$$M_I = M_p \frac{g_s^{1/8}}{\mathcal{V}_E^{1/4}} \left( \frac{\pi^3 2^7}{\zeta(3)} \right)^{1/12} \quad (4.132)$$

and

$$M_{II} = M_p \frac{g_s^{3/16}}{\mathcal{V}_E^{3/8}} \left( \frac{\pi^3 2^{10}}{\zeta(3)} \right)^{1/8} \quad (4.133)$$

To determine whether these higher-derivative terms will become important, we compute the dimensionless quantity  $\frac{\omega}{M_{I,II}}$ , where  $\omega = \frac{\dot{\phi}}{f}$  (4.14) is the physical frequency of oscillations; we obtain

$$\frac{\omega}{M_I} \simeq 5 \cdot 10^{-3} \left( \frac{f}{10^{-3}} \right)^{-1} \left( \frac{g_s}{0.2} \right)^{-1/8} \left( \frac{\mathcal{V}_E}{120} \right)^{1/4}, \quad (4.134)$$

$$\frac{\omega}{M_{II}} \simeq 6 \cdot 10^{-3} \left( \frac{f}{10^{-3}} \right)^{-1} \left( \frac{g_s}{0.2} \right)^{-3/16} \left( \frac{\mathcal{V}_E}{120} \right)^{3/8}. \quad (4.135)$$

For the ranges of  $f$  and  $\mathcal{V}_E$  that will be of interest to us (*cf.* §4.7), the higher-derivative terms are not important and our two-derivative approximation is justified.

#### 4.6.4 Constraints on the axion decay constant

In this section, we discuss direct constraints on the axion decay constant  $f$ . We first recall a rather general (conjectured) upper bound  $f < M_p$  [139], and we then

describe and incorporate a novel lower bound, specific to our setup, that arises from combining the requirements that  $\alpha'$  perturbation theory should be valid and that the inflationary energy should not drive decompactification.

Despite many attempts, at the time of writing there is no known, controllable string theory construction that provides  $f > M_p$ . In particular, the authors of [139] have scanned several classes of string theory models and found sub-Planckian axion decay constants in every case. However, this upper bound on  $f$  is of relatively little importance for the phenomenological signatures we are considering in this paper.

On the other hand, a potential lower bound on  $f$  is of considerable importance for our analysis. Considering oscillations in the CMB spectrum, in the regime  $f \ll M_p$  one can easily find models that range from being observationally excluded to giving undetectably small modifications, depending on the amplitude of the ripples in the inflationary potential. Furthermore, the resonant non-Gaussianity becomes large only for small  $f$  (e.g. we will find that  $f < 3 \cdot 10^{-3}$  is a necessary condition to give a reasonable prospect of detectability). Hence we will move on to consider possible lower bounds on  $f$ .

As discussed in [11] and in the preceding section, a direct lower bound on  $f$  comes from the requirement of small backreaction. In particular, the radius of curvature induced by the energy localized on the wrapped NS5-brane should be smaller than the smallest radius of curvature  $R_\perp$  in a direction transverse to the NS5-brane in the compactification. This requires

$$N_w \ll \frac{R_\perp^4 X}{4\pi g_s} \quad \Rightarrow \quad \frac{f}{M_p} \gg \frac{2\phi g_s}{R_\perp^4 X}, \quad (4.136)$$

where we have defined  $X \equiv \text{Vol}(X_5)/\pi^3$ , with  $X_5$  the base of the cone forming the warped throat. We remark that  $X \leq 1$ , as  $S^5$  is the Sasaki-Einstein manifold with

the largest volume, in the sense defined above. We can estimate  $R_\perp$  as being comparable to the AdS radius  $R$  of the throat containing the NS5-brane. Given that the volume<sup>18</sup>  $\mathcal{V}$  of the Calabi-Yau has to be larger than the volume of any throat it includes, one finds that

$$\mathcal{V} > V_{throat} = \frac{\pi^3}{2} X R^6, \quad (4.137)$$

where for simplicity we have assumed that the UV cutoff of the throat is at  $r \sim R$  where the warp factor becomes of order unity. Putting together (4.136) and (4.137), we find

$$\frac{f}{M_p} > \frac{\pi^2 2^{1/3} \phi g_s}{X^{1/3} \mathcal{V}^{2/3}} \simeq \frac{137 g_s}{X^{1/3} \mathcal{V}^{2/3}} = \frac{0.09}{X^{1/3} \mathcal{V}_E^{2/3}}. \quad (4.138)$$

Although the above constraint substantially restricts our parameter space, an even stronger constraint comes from demanding the validity of  $\alpha'$  perturbation theory: using (4.112) for  $f$  and combining this with the lower bound on two-cycle volumes given in (4.118), we obtain

$$\frac{f^2}{M_p^2} = \frac{\sqrt{g_s}}{(2\pi)^3 \mathcal{V}_E} (c_{\alpha--} v^\alpha \sqrt{g_s} \pi) > \frac{\sqrt{g_s}}{(2\pi)^3 \mathcal{V}_E}, \quad (4.139)$$

where we have assumed that  $c_{\alpha--} \geq 1$ . (4.139) turns out to give the strongest microphysical lower bound on  $f$ . An upper bound is harder to determine from this formula. Assuming again that  $c_{\alpha--} \geq 1$ , assuming that no precise cancellations occur, and using (4.118), we find

$$\frac{f}{M_p} < g_s \frac{\sqrt{3}}{2}. \quad (4.140)$$

---

<sup>18</sup>We always refer to the warped volume, calculated with the whole warped metric.

## 4.6.5 Constraints on the amplitude of the modulations

So far we have seen that with the Kähler potential and superpotential given in (4.99) and (4.100), the axion  $c$  persists as a flat direction after moduli stabilization.<sup>19</sup> As explained in [11], the presence of an NS5-brane wrapping the two-cycle that defines  $c$  introduces a monodromy and results, for large  $c$ , in the linear potential in (4.106). In this section we will consider further nonperturbative corrections that will in general induce small modulations of this linear potential. These are precisely the modulations whose phenomenology we have studied in the first part of this paper.

Nonperturbative corrections could appear both in the Kähler potential and in the superpotential. We focus on the first possibility and comment at the end of this section on the second. Consider the type IIB orientifolds with O3-planes and O7-planes. As we have remarked, the RR two-form  $C_2$  is odd under the orientifold projection and therefore a four-dimensional axion that survives projection comes from integrating  $C_2$  over an odd two-cycle  $v^-$ . Such an odd cycle can be thought of as  $v^- = v^1 - v^2$ , where  $v^1$  and  $v^2$  represent two two-cycles in the parent Calabi-Yau manifold that are mapped into each other by the orientifold action. Now consider a Euclidean D1-brane wrapping the even cycle  $v^+ = v^1 + v^2$ . Such an instanton feels the local  $\mathcal{N} = 1$  supersymmetry of the orientifolded theory, and it breaks this supersymmetry completely.<sup>20</sup> Hence this is a non-BPS instanton with four universal fermionic zero modes, namely the goldstini of the broken  $\mathcal{N} = 1$  supersymmetry. If the Euclidean D1-brane wraps a minimum-

---

<sup>19</sup>As we have remarked, the axion  $b$  has its flat direction lifted by nonperturbative stabilization of the Kähler moduli.

<sup>20</sup>To see this, note that (*cf.* [140]) the instanton action depends on a two-cycle volume, but the proper Kähler coordinates are four-cycle volumes. Therefore, the instanton action cannot be holomorphic, so the instanton cannot contribute to a superpotential, and must instead be non-BPS.

volume cycle in the homology class  $v^+$  then it has the right total number of fermionic zero modes (four) to contribute to a D-term and in particular to the Kähler potential.

More specifically, in [141] it was argued that nonperturbative contributions from worldsheet instantons and their  $SL(2, \mathbb{Z})$  images, Euclidean  $(p, q)$  strings, give rise to corrections to the prepotential of the  $\mathcal{N} = 2$  theory of the parent Calabi-Yau compactification. Such corrections are most naturally expressed inside the logarithm of the Kähler potential,

$$K = -2 \log \left[ \mathcal{V}_E + g(G, \bar{G}) \right], \quad (4.141)$$

where  $g$  is an appropriate function. Invariance under  $SL(2, \mathbb{Z})$ , or more generally under a subgroup  $\Gamma \subset SL(2, \mathbb{Z})$ , is naturally achieved if  $g$  is the sum of some individual correction  $\tilde{g}$  over an orbit of  $\Gamma$ .

At the time of writing, the nonperturbative correction  $g$  is not known explicitly, but a modular-invariant result has been conjectured in [142]. Inspired by the structure of this result (which we will not reproduce here), we will make a simple educated guess based on the following criteria: the non-perturbative correction should go to zero exponentially for large two-cycle volume  $v^+$ ; it should break the continuous shift-symmetry of  $c$  to a discrete shift-symmetry  $c \rightarrow c + 2\pi$ ; and it should be invariant under whatever discrete subgroup  $\Gamma \subset SL(2, \mathbb{Z})$  of the ten-dimensional  $SL(2, \mathbb{Z})$  symmetry is preserved by the compactification. The subgroup  $\Gamma$  may well be trivial, and we will assume this for simplicity; note, however, that one can plausibly obtain a more constrained result when some or all of the symmetry is preserved, as in [142]. Moreover, notice that along the orbits of  $\Gamma$ , the instanton action generally increases compared to that of a single worldsheet instanton or Euclidean D1-brane; thus, when the volume  $v^+$  is not

too small, only a few terms make an important contribution, with the remainder enjoying further exponential suppression.

A reasonable guess satisfying these criteria, for  $\Gamma$  trivial, is

$$K = -2 \log \left[ \mathcal{V}_E + e^{-S_{ED1}} \cos(c) \right] = -2 \log \left[ \mathcal{V}_E + e^{-\frac{2\pi v^+}{\sqrt{g_s}}} \cos(c) \right]. \quad (4.142)$$

In light of this corrected Kähler potential, we should revisit the moduli stabilization before proceeding to calculate the size  $b$  of the periodic contribution to the scalar potential.

We begin by noticing the following implication

$$\begin{cases} D_{T_\alpha} W = \mathcal{O}(e^{-2S_{ED1}}) \\ D_{G^a} W = \mathcal{O}(e^{-S_{ED1}}) \\ W_{G^a} = 0 \end{cases} \Rightarrow \begin{cases} \partial_{T_\alpha} V = 0 + \mathcal{O}(e^{-2S_{ED1}}) \\ \partial_{G^a} V = -2 e^K |W|^2 K_{G^a} + \mathcal{O}(e^{-2S_{ED1}}) \end{cases} \quad (4.143)$$

which can be verified by direct computation. This allows us to use the F-flatness condition to find the minimum in the  $T_\alpha$ -directions even when one of the F-terms, namely  $D_{G^-} W$ , does not vanish. Equipped with this knowledge we repeat, *mutatis mutandis*, the steps of §4.5.

First, the phases of the  $T_\alpha$  are stabilized as in (4.103), with  $k$  being odd as explained below (4.104). The reason is that the sign of  $\partial_{T_\alpha} K$  is not changed by the small nonperturbative correction  $e^{-S_{ED1}}$ . Second,  $\text{Im } G$  is again stabilized at 0. Given (4.143), the equation one needs to solve is  $D_{G^a} W = 0$ , which reduces to

$$0 = W \partial_{G^a} K \propto \partial_{G^a} \mathcal{V}_E - e^{-S_{ED1}} \left[ \pi \sin(c) + \cos(c) \frac{2\pi}{\sqrt{g_s}} \partial_{G^a} v^+ \right] = 0, \quad (4.144)$$

where we made use of (4.96) to perform the derivative on  $c$ . Since  $\mathcal{V}_E$  and  $v^+$  only depend on  $\text{Im } G$  implicitly as in (4.111), we can take the imaginary part of

(4.144),

$$\frac{1}{2}(\partial_{\text{Im} G^-} v^\alpha) c_{\alpha\beta\gamma} v^\beta v^\gamma - e^{-S_{ED1}} \cos(c) \frac{2\pi}{\sqrt{g_s}} \partial_{\text{Im} G^-} v^+ = 0. \quad (4.145)$$

But from (4.111) we know that

$$c_{\alpha\beta\gamma} (\partial_{\text{Im} G^-} v^\beta) v^\gamma = g_s c_{\alpha a} \text{Im} G^a, \quad (4.146)$$

which means that  $\text{Im} G^a = 0$  (for every  $a$ ) is a solution to (4.145). The real part of (4.144) is nonvanishing and of order  $e^{-S_{ED1}}$ . Again because of (4.143), the minimization in the  $\tau_\alpha$  is obtained by imposing  $D_{T_\alpha} W = 0$ . These equations depend on the inflaton  $c$ , appearing explicitly in (4.142), and hence the minimum in the  $T_\alpha$  directions will be a function of  $c$ . Integrating out the  $T_\alpha$  leads to a contribution in the effective potential  $V[T(c), c]$  for  $c$  which is of the same order as the contribution coming from the explicit  $c$ -dependence in the Kähler potential. Therefore this effect cannot be neglected. To take it into account, we solve the  $D_{T_\alpha} W = 0$  equations perturbatively in  $e^{-S_{ED1}}$ .

We define the coefficients of the minimum in the  $\tau_\alpha$  directions in a perturbative expansion in  $e^{-S_{ED1}}$  by

$$\tau_{\alpha, \min} \equiv \tau_{\alpha, (0)} + \cos(c) e^{-S} \tau_{\alpha, (1)} + \dots, \quad (4.147)$$

and so on for all other variables. The zeroth-order equations are

$$(D_{T_\alpha} W)_{(0)} = (\partial_{T_\alpha} W)_{(0)} + W_{(0)} (\partial_{T_\alpha} K)_{(0)} = 0, \quad (4.148)$$

which can be solved numerically once the model is specified. The first-order equations are

$$(D_{T_\alpha} W)_{(1)} = (\partial_{T_\alpha} W)_{(1)} + W_{(1)} (\partial_{T_\alpha} K)_{(0)} + W_{(0)} (\partial_{T_\alpha} K)_{(1)} = 0, \quad (4.149)$$

which again can be solved numerically using the solutions of (4.148). We turn now to estimate the parameter  $b$  defined in (4.2). One finds

$$bf \equiv V_{(1)}\mu^{-3}e^{-S_{ED1}} = \frac{\mathcal{U}_{mod}\phi}{\mu^3\phi}e^{-S_{ED1}}\left(K_{(1)} + 2\text{Re}\frac{W_{(1)}}{W_{(0)}}\right), \quad (4.150)$$

where we have defined  $\mathcal{U}_{mod}$  as the moduli stabilization barrier at zeroth order in  $e^{-S_{ED1}}$ , *i.e.*

$$\mathcal{U}_{mod} = \frac{g_s}{2}\left(\frac{3|W|^2}{\mathcal{V}_E^2}\right)_{(0)}. \quad (4.151)$$

More explicitly, using (4.142) and (4.100),

$$\begin{aligned} bf &= \frac{\mathcal{U}_{mod}\phi}{2.4 \cdot 10^{-9}M_p^4}e^{-S_{ED1}}\left[\frac{8\pi}{\sqrt{g_s}}\frac{(\partial_{\tau_\alpha}v^+)_{(0)}}{v_{(0)}^\alpha} - 2a_\alpha\tau_{\alpha,(1)} - \frac{2v_{(1)}^\alpha}{v_{(0)}^\alpha}\right] \\ &= \frac{\mathcal{U}_{mod}\phi}{2.4 \cdot 10^{-9}M_p^4}2e^{-S_{ED1}}\left[\frac{\sum_\beta(\partial_{T_\beta}W)_{(0)}\tau_{\beta,(1)}}{W_{(0)}} - \frac{\mathcal{V}_{E,(1)} + 1}{\mathcal{V}_{E,(0)}}\right], \end{aligned} \quad (4.152)$$

where the first line is valid for any  $\alpha$  and the second line (obtained using (4.149)) shows that the expression for  $b$  is independent of  $\alpha$ . Notice that  $\partial_{T_\alpha}v^+ = \frac{1}{2}\partial_{\tau_\alpha}v^+$  is given implicitly by

$$c_{\alpha\beta\gamma}(\partial_{\tau_\rho}v^\gamma)v^\beta = \partial_{\tau_\rho}\tau_\alpha = \delta_\alpha^\rho. \quad (4.153)$$

Some comments are in order. The size of the ripples in the potential is proportional to the ratio of the moduli stabilization barrier to the scale of inflation, which has to be large for the self-consistency of the estimate. We have used the value of the potential at the would-be AdS minimum to estimate the moduli stabilization barrier once an uplifting term is included. Due to the exponential suppression  $e^{-S_{ED1}}$ , the size of  $bf$  is extremely sensitive to  $g_s$  and  $v^+$ .

An upper bound can be derived from (4.152) using the following considerations. In the KKLT construction, perturbative corrections to the Kähler potential can be neglected as long as  $W_0 \ll 1$ , and generically  $W \sim W_0$ . For larger values of

$W_0$ , perturbative corrections have to be included, as in the large volume scenario [143]. In the present work, we focused on the former setup and we leave an investigation of the latter for the future. The exponential suppression in (4.152) can be bounded by (4.118). Finally, we denote the model-dependent term in square brackets in (4.152) by  $c_0$ . Putting things together leads to the bound

$$bf < 2c_0 \cdot 10^7 \frac{g_s}{\mathcal{V}_E^2} e^{-2/g_s} \left( \frac{W}{0.1} \right)^2. \quad (4.154)$$

Even imposing all the model-independent constraints we have described in the previous sections, one can still have  $bf > 10^{-4}$ , which, as shown in §4.4, is roughly the upper bound imposed by measurements of the scalar power spectrum. Therefore, in certain parameter ranges the primary constraint on modulations of the potential comes from the data, not from microphysics.

## 4.7 Combined Theoretical and Observational Constraints

We now summarize our results, combining the observational constraints from §4.4 with the theoretical constraints from §4.6. As an aid to the reader, we will now briefly recall the qualitative properties of those results.

Axion monodromy can produce characteristic signatures in the CMB: the oscillations in the axion potential generated by nonperturbative effects source resonant contributions to the scalar power spectrum and bispectrum. The amplitude and frequency of the oscillations in the potential can therefore be bounded by comparison to observations. We recall from §4.4 that the observational constraints take the form of exclusion contours in the space of the phenomenological parameters, after marginalization over additional model parameters that

have important degeneracies with those displayed. For convenience, we have chosen to display constraints in terms of the parameters  $f$  and  $bf$  defined in (4.2), marginalizing over the phase  $\Delta\phi$  and over  $\Omega_B h^2$ .

The first new step is to combine the exclusion contours based on the temperature two-point function with estimates of constraints from the three-point function. Based on the rough estimates described in §4.3.4, we present, in figure 4.6, three contours at  $f_{res} = 200, 20, 2$ , with the expectation that the gray region ( $f_{res} > 200$ ) might plausibly be excluded, while the colored, lighter regions ( $20 < f_{res} < 200$  and  $2 < f_{res} < 20$ , respectively) are possibly within detectability. A careful study of the constraints on resonant non-Gaussianity would be a worthwhile topic for future research.

Next, we recall that in §4.6, we found that the requirements of consistency and computability in the string compactifications giving rise to axion monodromy models led to constraints on the parameters  $f$  and  $bf$ . Let us remark that as these constraints are not rooted in deep principles of string theory or of quantum field theory, but rather originate in practical limitations in our present ability to construct computable models, they may well be loosened in further work. As such, the theoretical constraints we present here should be understood as designating *included* rather than excluded regions: in contrast to experimental contours, theoretical contours of this sort may expand rather than contract given improved understanding.

Because the parameter  $b$  measures the amplitude of a nonperturbative effect, it is exponential in the natural input parameters, and can therefore be made small without substantial fine-tuning. We therefore do not present a lower

bound on  $b$ . However, we found the theoretical upper bound (4.154)<sup>21</sup>

$$bf < 2c_0 \cdot 10^9 M_p^4 \frac{g_s}{\mathcal{V}_E^2} e^{-2/g_s}, \quad (4.155)$$

with a model-dependent constant  $c_0$  that can be estimated in explicit examples, and which we find to be typically of order  $10^{-2}$ .

Next, we obtained a lower bound for  $f$  in (4.139).<sup>22</sup> A precise upper bound, however, is highly model-dependent. We estimate an upper limit by assuming that the intersection numbers are of order one<sup>23</sup> and that no precise cancellations occur. From (4.139) and (4.140), the complement of the theoretically excluded range for  $f$  is then

$$\frac{g_s^{1/4}}{(2\pi)^{3/2} \sqrt{\mathcal{V}_E}} < f < g_s \frac{\sqrt{3}}{2}. \quad (4.156)$$

Notice that the theoretical constraints depend mainly on two quantities: the string coupling  $g_s$  and the volume  $\mathcal{V}_E$  of the compactification. The former appears in the exponential suppression of the nonperturbative effect generating the modulations of the potential. Hence,  $g_s \lesssim 0.1$  suppresses any possible signature of the modulations. For  $g_s \gtrsim 0.1$ , there is always a theoretically allowed region in which the oscillations in the inflaton potential lead to observable ripples in the two-point function of the CMB. On the other hand, the size of the non-Gaussianity depends critically on  $\mathcal{V}_E$  as well. Assuming  $g_s \gtrsim 0.1$ , larger  $\mathcal{V}_E$  allows for a larger range of  $f$  and therefore larger non-Gaussianity (see (4.156)). A way to quantify this is to use the estimate obtained in §4.3.4,

$$f_{res} \simeq \frac{9}{4} \frac{b}{(\phi f)^{3/2}}, \quad (4.157)$$

---

<sup>21</sup>We stress that this ‘bound’ is *not* universal and depends on the assumptions enumerated in §4.6. We include it here as a representative example of the constraints that arise in particular scenarios.

<sup>22</sup>The constraint from the backreaction described in §4.6.2 is weaker than (4.139).

<sup>23</sup>Larger intersection numbers are an interesting possibility that we will not investigate here.

and the lower bound in (4.156). The result is

$$\mathcal{V}_E > 170 \left( \frac{g_s}{0.2} \right)^{1/2} \left( \frac{f_{res}}{10} \right)^{4/5} \left( \frac{10^{-4}}{bf} \right)^{4/5}. \quad (4.158)$$

We now combine the theoretical and observational constraints, presenting them in the plane  $\{\log(f), \log(b)\}$ . We choose as boundaries  $10^{-4} < f \ll M_p$  and  $10^{-4} < b \ll 1$  based on the following considerations. The number of oscillations per e-folding is roughly  $10^{-2}M_p/f$ . Hence for  $f \gtrsim 0.1M_p$  there is less than one oscillation in the whole range of scales probed by the CMB, and the signal from modulations becomes degenerate with the overall amplitude. Furthermore, in §4.3, we systematically used the expansion  $b \ll 1$ , where  $b = 1$  divides monotonic from non-monotonic potentials. Finally, the lower boundaries  $10^{-4} < f$  and  $10^{-4} < b$  exclude regions that are relatively uninteresting in the present context: smaller values of  $b$  lead to an unobservably small signal, while smaller values of  $f$  are rather difficult to obtain in the class of string theory constructions we considered. In the  $\{\log(f), \log(b)\}$  plane, the theoretically allowed region looks like an interval in  $f$ , whose size is determined by  $\mathcal{V}_E$ , with an upper cut effectively determined by  $g_s$  as in (4.155).

Finally, in figure 4.8 we show where a particular numerical toy example, with specific choices of the intersection numbers, lies in the  $\{\log(f), \log(b)\}$  plane.

## 4.8 Conclusions

The goal of this investigation was to characterize the predictions of axion monodromy inflation for the CMB temperature anisotropies. Nonperturbative effects in these models generically introduce sinusoidal modulations of the in-

flaton potential, which in turn lead to resonantly-enhanced modulations of the scalar spectrum and bispectrum.

We have provided a simple analytic result for the modulated scalar power spectrum in this class of models. We also presented an alternative derivation in terms of episodes of particle production driven by resonance between a mode inside the horizon and the driving force of the oscillatory background evolution. We then determined in detail the constraints that the five-year WMAP data places on models with modulations of this sort.

Next, after reviewing the realization of axion monodromy inflation in string theory, we performed a comprehensive study of the parameter constraints implied by the requirements of microphysical consistency and computability. The resulting allowed parameter regions are very plausibly realizable in sensible string theory constructions.

We also identified a new contribution to the inflaton potential in axion monodromy inflation: the backreaction of the inflationary energy on the compact space can source an important correction to the potential by correcting the volumes of four-cycles and hence affecting the scale of nonperturbative moduli-stabilizing effects. We then presented a model-building solution to this problem, in which the NS5-brane and anti-NS5-brane driving inflation are in the same warped region, or more generally are distant from the four-cycles of interest.

Finally, we combined the observational and theoretical constraints, in order to ascertain whether detectable modulations of the scalar spectrum and/or bispectrum are possible, consistent with current observational bounds and known

theoretical restrictions. Our conclusion is that both sorts of modulations are possible, and in fact in many cases the strongest bound on the amplitude of the modulations comes from data, not from microphysics. Moreover, even though observational limits on the amplitude and frequency of modulations in the scalar power spectrum provide a strong constraint on the parameter space of axion monodromy models, and even though microphysical constraints sharply restrict the allowed frequencies, detectably-large non-Gaussianity can indeed be produced in a class of controllable models. Such models enjoy three non-trivial signatures: detectable tensors with  $r \approx 0.07$ , a modulated scalar power spectrum, and resonant non-Gaussianity.

Let us remark that even in the absence of non-Gaussianity, this class of models is eminently testable: axion monodromy inflation unambiguously predicts a large tensor signal, and the parameters of the models are already strongly constrained by limits on modulations in the scalar power spectrum.

There are several interesting directions for future work. First, we have not analyzed the constraints on the model from the three-point function; more generally, understanding the prospects for constraining or detecting resonant non-Gaussianity is an important task. Moreover, it would be instructive to construct an explicit model in which the many theoretical constraints we have checked can be combined in a coordinated way. In addition, it would be interesting to determine whether chain inflation can be realized in this context.

It is intriguing that the modulated power spectrum we have found is very similar in form to that proposed in the context of modifications of the initial state, as in *e.g.* [118, 119, 104, 105, 106]. In light of our calculation of the power spectrum in terms of particle production (§4.3.3), this coincidence is not entirely

surprising: the driving force of the oscillating background eventually generates an excited state, even if one begins in the Bunch-Davies vacuum. We leave for the future a more systematic exploration of this connection.

Finally, it would be most valuable to develop a broader understanding of the connection, if any, between symmetries and signatures in models of large-field inflation.

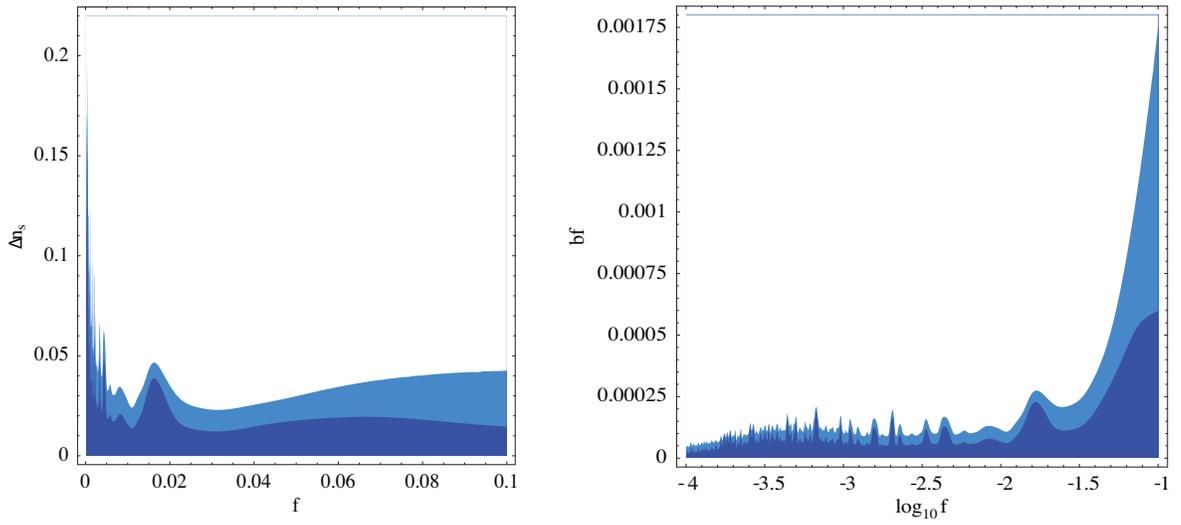


Figure 4.2: This plot shows the 68% and 95% likelihood contours in the  $\delta n_s$ - $f$ , and  $bf$ - $\log_{10} f$  plane, respectively, from the five-year WMAP data on the temperature angular power spectrum.

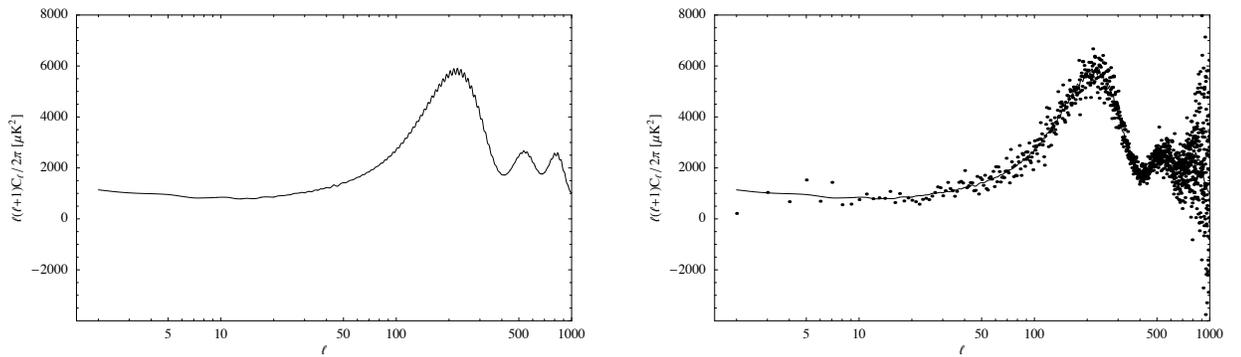


Figure 4.3: The left plot shows the angular power spectrum for the best fit point  $f = 6.67 \times 10^{-4}$ , and  $\delta n_s = 0.17$ . The right plot shows the angular power spectrum for the best fit point together with the unbinned WMAP five-year data.

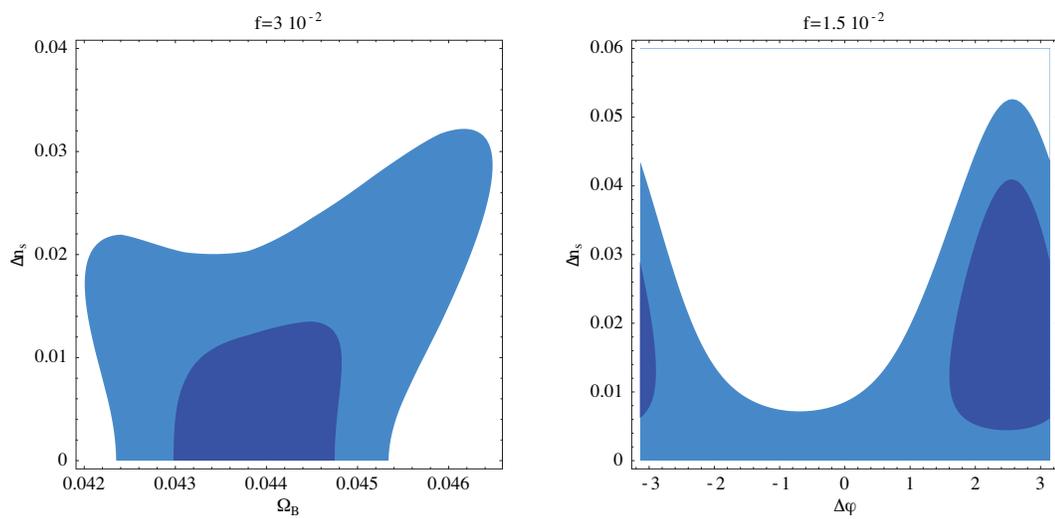


Figure 4.4: These plots show the 68% and 95% likelihood contours for the five-year WMAP data on the temperature angular power spectrum in the  $\delta n_s - \Omega_B h^2$  plane and  $\delta n_s - \Delta\phi$  plane for an axion decay constant of  $f = 3 \times 10^{-2}$  and  $f = 1.5 \times 10^{-2}$ , respectively.

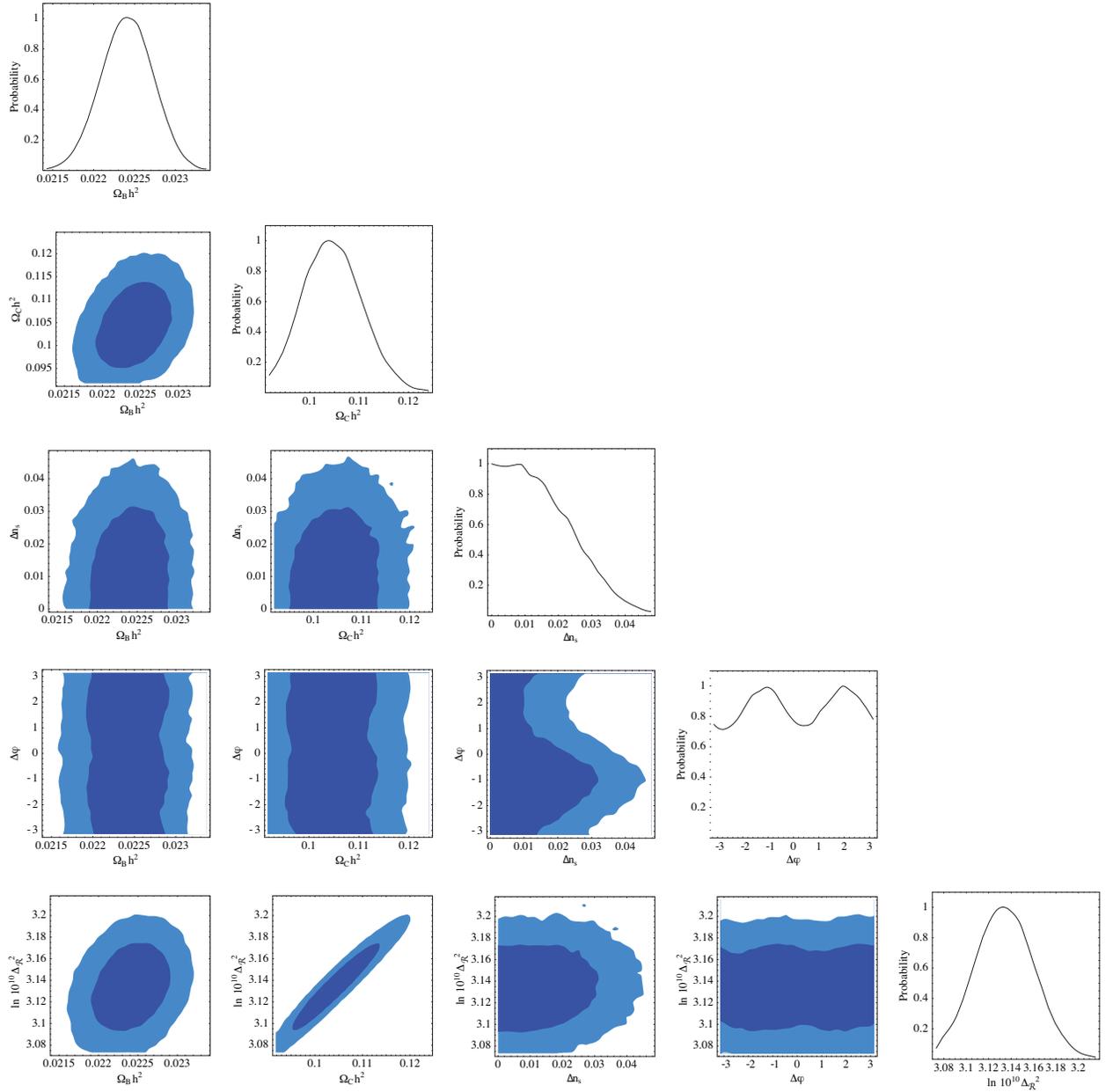


Figure 4.5: This figure shows a triangle plot for some of the parameters that were sampled in a Markov chain Monte Carlo for an axion decay constant of  $f = 10^{-2}$ . The contours again represent 68% and 95% confidence levels.

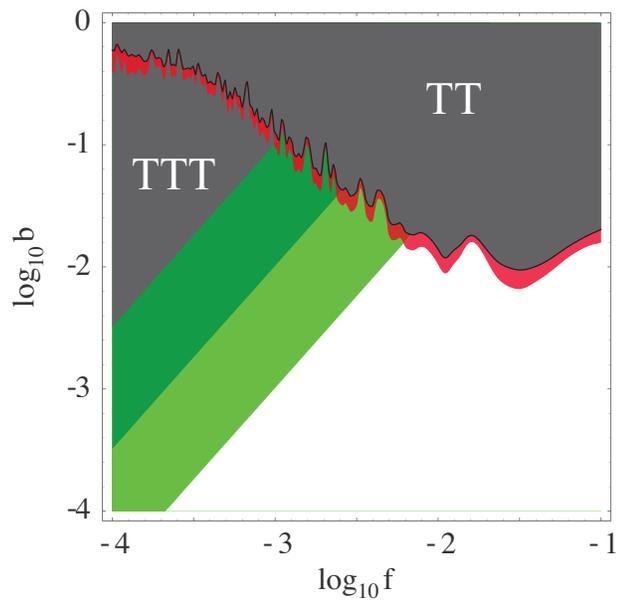


Figure 4.6: We show the (one- and two-sigma) likelihood contours for the temperature two-point function together with three contours that characterize the amplitude of the three point function, for  $f_{res} = 200, 20, 2$ .

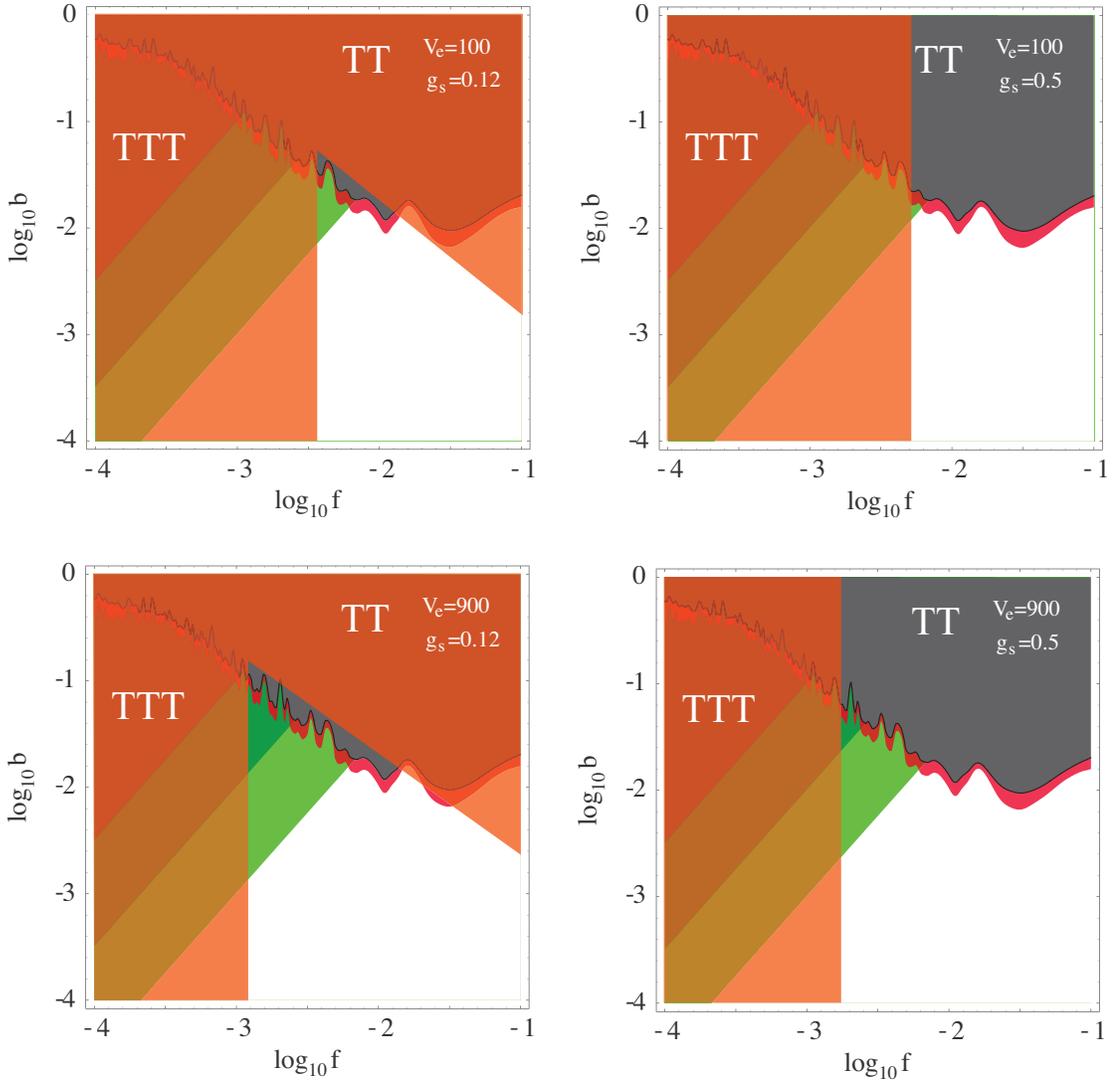


Figure 4.7: We superimpose the theoretical constraints, summarized in (4.155) and (4.156), on the constraints imposed by observations, which are shown in figure 4.6. The orange overlay indicates regions of the parameter space that are difficult to reach in the class of models considered in the present work. The theoretical constraints are shown for  $g_s = 0.12, 0.5$  and  $\mathcal{V}_E = 100, 900$ .

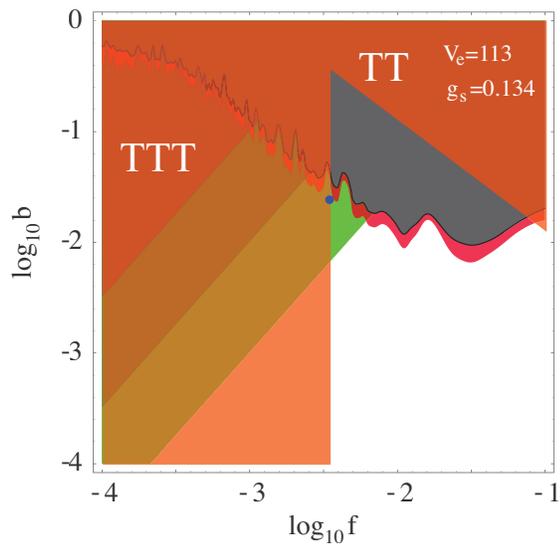


Figure 4.8: The blue dot represents the explicit numerical example presented in appendix B.4. It represents a case in which upcoming experiments could detect the signatures of modulations in both the two-point function and the three-point function.

## APPENDIX A

### CHAPTER 1 OF APPENDIX

#### A.1 Structure of the Scalar Potential

In this appendix we summarize the computation of bulk contributions to the inflaton potential in warped D-brane inflation, following [6], to which we refer for notation and for further details.

D3-branes experience the potential

$$V_{D3} = T_3 (e^{4A} - \alpha) \equiv T_3 \Phi_- . \quad (\text{A.1})$$

The classical equations of motion imply

$$\nabla^2 \Phi_- = \frac{g_s}{96} |\Lambda|^2 + \mathcal{R}_4 + \mathcal{S}_{\text{local}} , \quad (\text{A.2})$$

where  $\nabla^2$  is constructed using the unwarped metric on the compact space,  $\Lambda$  is proportional to the imaginary anti-self-dual three-form flux, and  $\mathcal{S}_{\text{local}}$  is a localized source due to anti-D3-branes.

The homogeneous solutions to (A.2) are harmonic functions on the conifold, i.e. solutions to the Laplace equation  $\nabla^2 h = 0$ . Expanding  $h$  in angular harmonics  $Y_{LM}(\Psi)$  on  $T^{1,1}$ , we have

$$h(x, \Psi) = \sum_{L,M} h_{LM} x^{\delta(L)} Y_{LM}(\Psi) + c.c. \quad (\text{A.3})$$

where  $h_{LM}$  are coefficients,  $L \equiv (j_1, j_2, R_f)$  and  $M \equiv (m_1, m_2)$  label the  $SU(2) \times SU(2) \times U(1)_R$  quantum numbers under the isometries of  $T^{1,1}$ , and the radial scaling dimensions  $\delta(L)$  are related to the eigenvalues of the angular Laplacian,

$$\delta(L) \equiv -2 + \sqrt{H(j_1, j_2, R_f) + 4} , \quad (\text{A.4})$$

where [27]

$$H(j_1, j_2, R_f) \equiv 6(j_1(j_1 + 1) + j_2(j_2 + 1) - R_f^2/8). \quad (\text{A.5})$$

Next, let us consider inhomogeneous contributions to the D3-brane potential sourced by imaginary anti-self-dual flux  $\Lambda$ . Each mode of flux is a three-form with specified quantum numbers  $L = (j_1, j_2, R_f)$  and  $M = (m_1, m_2)$  under the angular isometries. Thus, we can consider a general imaginary anti-self-dual flux  $\Lambda$ , compute<sup>1</sup>  $|\Lambda|^2$ , and expand the result in terms of irreducible representations of the angular isometries. (That is, we expand products of two harmonic functions on  $T^{1,1}$  in terms of individual harmonic functions, ensuring that each term still obeys the selection rules.) Using the Green's function given in [144], one readily obtains the corresponding  $\Phi_-$  profile.

The general potential for a D3-brane therefore involves a sum of homogeneous terms  $h$ , inhomogeneous contributions sourced by flux, the Coulomb potential sourced by  $\mathcal{S}_{\text{local}}$ , and curvature contributions. (For the computation of higher-order contributions from curvature, see [6].)

---

<sup>1</sup>Some care is needed because certain contractions vanish identically.

## APPENDIX B

### CHAPTER 2 OF APPENDIX

#### B.1 Notation and Conventions

In this appendix we review our conventions, emphasizing differences with the existing literature.

A good starting point is the ten-dimensional string-frame action<sup>1</sup> [129]

$$S_{10} = \frac{1}{(2\pi)^7 \alpha'^4} \int d^{10}x \sqrt{g_{string}} \left( e^{-2\Phi} R_{string} - \frac{1}{2} |dC_2|^2 \right) \quad (\text{B.1})$$

which after the rescaling to the ten-dimensional Einstein-frame metric  $e^{-\Phi/2} g_{string, MN} = g_{E, MN}$  becomes

$$S_{10} = \frac{1}{(2\pi)^7 \alpha'^4} \int d^{10}x \sqrt{g_E} \left( R_E - \frac{1}{2} g_s |dC_2|^2 \right), \quad (\text{B.2})$$

where we assumed that the axio-dilaton is  $\tau = i/g_s$ . Upon compactifying on a six-dimensional manifold  $Y$ , the resulting four-dimensional reduced Planck mass is

$$M_p^2 = \frac{\int_Y \sqrt{g_E}}{(2\pi)^6 \alpha'^3} \frac{1}{\alpha' \pi} \equiv \frac{\mathcal{V}_E}{\alpha' \pi}, \quad (\text{B.3})$$

where  $\mathcal{V}_E$  is the (dimensionless) Einstein volume of  $Y$  measured in units of  $2\pi \sqrt{\alpha'}$ . When  $Y$  is (conformally equal to) a Calabi-Yau space, we have

$$\mathcal{V}_E = \frac{1}{6} \frac{\int J \wedge J \wedge J}{(2\pi)^6 \alpha'^3} = \frac{1}{6} v^J v^J v^K \frac{\int \omega_I \wedge \omega_J \wedge \omega_K}{(2\pi)^6 \alpha'^3} \equiv \frac{1}{6} v^J v^J v^K c_{IJK}, \quad (\text{B.4})$$

where  $\omega_I$  for  $I = 1, \dots, h^{1,1}$  are a basis of the cohomology  $H^2(Y, \mathbb{Z})$  normalized such that

$$\int_{\Sigma_I} \omega_J = (2\pi)^2 \alpha' \delta_J^I \quad (\text{B.5})$$

---

<sup>1</sup>Remember that  $2\kappa_{10}^2 = (2\pi)^7 \alpha'^4$ .

for a basis  $\Sigma_I$  of the dual homology  $H_2(Y, \mathbb{Z})$ . With the ansatz for the ten-dimensional RR two-form

$$C_2 = \frac{1}{2\pi} c(x) \omega, \quad (\text{B.6})$$

for some base two-cycle  $\omega$ , we get a four-dimensional axion  $c(x)$  with periodicity<sup>2</sup>  $2\pi$ , as can be seen *e.g.* via S-duality starting from the world-sheet coupling  $\int B/(2\pi\alpha')$ . The axion decay constant of  $c$  is

$$\frac{f^2}{M_p^2} = \frac{g_s}{12\mathcal{V}_E(2\pi)^2} \left[ \frac{\int \omega \wedge *\omega}{(2\pi)^6 \alpha'^3} \right]. \quad (\text{B.7})$$

The four-dimensional  $\mathcal{N} = 1$  Kähler potential for the Kähler moduli is

$$K = -2 \log \mathcal{V}_E. \quad (\text{B.8})$$

## B.2 Induced Shift of the Four-Cycle Volume

In this appendix we address the issue raised in §4.6.2: the inflationary energy can correct the warped volumes of four-cycles in the compact space, leading to corrections to the moduli potential, and hence inducing new terms in the inflaton potential itself.

More specifically, if an NS5-brane wraps some cycle  $\Sigma$ , then a nonvanishing integral  $\int_{\Sigma} C_2 \neq 0$  leads to the presence of energy that is localized near  $\Sigma$  in the compact space; this energy corresponds to the increased tension of the NS5-brane. Moreover, there is a corresponding induced D3-brane charge via the coupling  $\int C_2 \wedge C_4$ . The increased tension creates a backreaction on the metric (and in particular, on the warp factor) of the compact space, while the induced charge

---

<sup>2</sup>Note that this choice differs from that in [11], where the axion periodicity was  $(2\pi)^2$ .

sources five-form flux. We must determine whether these effects substantially correct the nonperturbative effects that are responsible, in our KKLT-like scenario, for stabilization of the Kähler moduli.

Whether the nonperturbative superpotential arises from gaugino condensation on D7-branes or from Euclidean D3-branes, it is exponentially sensitive to the warped volume of the four-cycle wrapped by these D-branes. Therefore, we will carefully consider the possibility of an inflaton-dependent shift of the warped volume of various four-cycles.

Concretely, we will consider a fivebrane/anti-fivebrane pair wrapping two homologous cycles, and will compute the leading correction to the volume of a particular four-cycle in the same throat region as the fivebranes. This will serve as a conservative upper bound on the effect of the worldvolume flux, as more distant four-cycles would be more weakly affected.

It would be very interesting to perform a systematic study of this backreaction in the four-dimensional effective theory and in ten-dimensional supergravity/string theory. We leave this task for future investigation. In what follows, we simply show that the effect described above can be ameliorated by choosing an appropriate configuration.

There are two mechanisms to suppress the backreaction on a given four-cycle volume. A first improvement comes from choosing a setup in which the leading backreaction is due to a dipole as opposed to a monopole potential. This allows for a parametric suppression. The second improvement can be achieved by a carefully chosen geometry of the four-cycle under consideration. In general, this latter mechanism requires fine tuning.

The problem of estimating the backreaction from two-form flux on an NS5-brane pair may be simplified by a series of approximations. First, the inflaton-dependent backreaction is generated by the increased tension and the induced D3-brane charge of the NS5-branes, which may be understood as corresponding to some number of D3-branes (or anti-D3-branes) dissolved in the NS5-branes. In practice, it is much simpler to study the effect of the D3-branes themselves; this captures the leading inflaton-dependent contributions.

The configuration of interest involves an NS5-brane wrapping  $\Sigma$ , with

$$\frac{1}{(2\pi)^2 \alpha'} \int_{\Sigma} C_2 \equiv N_w, \quad (\text{B.9})$$

as well as a distant NS5-brane wrapping a homologous cycle  $\Sigma'$ , but with opposite orientation. (We will refer to the latter object as the anti-NS5-brane.) Next, we recall that the COBE normalization requires each fivebrane to be in a warped region. Let us denote by  $N/2$  the amount of D3-brane charge that creates the background warping for each of the fivebranes.<sup>3</sup>

In light of the above discussion, we may approximate the fivebrane by a stack of  $N_w$  D3-branes and the anti-fivebrane by a stack of  $N_w$  anti-D3-branes. Combining this with the background D3-brane charge, we conclude that a convenient proxy for our system consists of two stacks of D3-branes, which we call A and B respectively. The first consists of  $N/2 + N_w$  D3-branes and the second of  $N/2$  D3-branes and  $N_w$  anti-D3-branes, which we may more conveniently represent as  $N/2 - N_w$  D3-branes and  $N_w$  brane-antibrane pairs.

Next, using the results of [136], we recognize that the leading backreaction effect comes from the total D3-brane charge on each stack, while the brane-

---

<sup>3</sup>More generally, one could consider different degrees of warping for each fivebrane; extending our considerations to this case is straightforward.

antibrane pairs lead to subleading effects that are suppressed by powers of the warp factor. Thus, we can simplify even further, so that at last we are considering a supersymmetric system involving two stacks that contain  $N/2 + N_w$  and  $N/2 - N_w$  D3-branes, respectively.

Equipped with this much simpler system, we may now estimate the inflaton-dependent backreaction, by computing how the presence of the stacks A and B leads to a  $N_w$ -dependent change in the warped volume of some four-cycle.

We choose the usual D3-brane ansatz

$$ds_6^2 = \sqrt{H^{-1}(y)} ds_4^2 + \sqrt{H(y)} ds_2^2 \quad (\text{B.10})$$

$$\tilde{F}_5 = (1 + *)dH^{-1} \wedge \text{Vol}_4, \quad \Phi = \text{const}. \quad (\text{B.11})$$

The resulting equation of motion is linear in  $H(y)$ . Therefore we may simply add the solutions obtained in the presence of either of the two individual stacks. Once the resulting warp factor is used to compute the volume of a four-cycle, the  $N_w$  dependence of the result gives us an estimate of the inflaton-dependence of the nonperturbative superpotential.

We tackle the problem in two steps of increasing complexity. First, in §B.2.1 we give a very simple, (conformally) flat toy example in which the calculations are easy. This already shows the relevant features of the more complicated solution. The inflaton-dependent shift of the volume can be suppressed by having the distance between A and B much smaller than the distance between the four-cycle and either of A and B. This corresponds to a configuration in which the leading interaction is via a dipole. In addition, one can fine-tune the four-cycle embedding so that the  $N_w$ -dependent correction to its volume actually cancels.<sup>4</sup>

---

<sup>4</sup>Although suppression from symmetry of the embedding appears unappealing because of

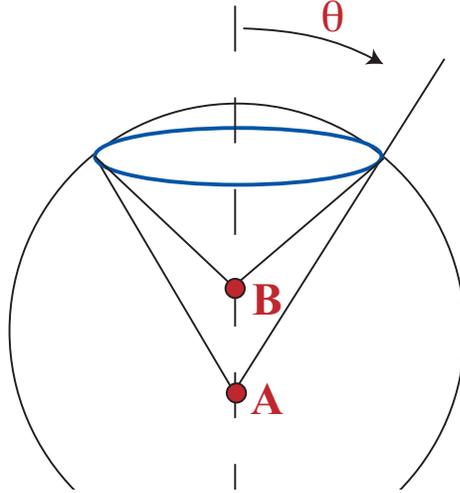


Figure B.1: This diagram illustrates the positions of the A and B stacks of D3-branes in  $\mathbb{R}^6$ , the choice of the angular coordinate  $\theta$ , and, in blue, a (topologically trivial) four-cycle.

Then, in §B.2.2, §B.2.2 and §B.2.3, we describe the case of a resolved conifold using the solution of [144, 145]. We consider a particular holomorphic embedding of a four-cycle and compute numerically the inflaton-dependence of its warped volume.

### B.2.1 A simple illustration of the suppression mechanism

Consider two stacks of  $N/2 \pm N_w$  D3-branes, called A and B, respectively, in conformally flat space  $M_4 \times \mathbb{R}^6$ . The A stack is located at the origin of  $\mathbb{R}^6$  and the fine-tuning required, one should keep in mind that it could conceivably be enforced by a discrete symmetry of the compactification.

the B stack is located at some position  $(u, 0, 0, 0, 0, 0)$  for<sup>5</sup>  $u \in \mathbb{R}^+$ , where we have chosen spherical coordinates (see figure B.2.1) with the metric

$$ds_6^2 = dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\Omega_4^2), \quad (\text{B.12})$$

where  $d\Omega_4$  is the volume form of  $S^4$ . With the usual D3-brane ansatz (B.10) one finds the solution

$$H = 1 + \frac{R_A^4}{r^4} + \frac{R_B^4}{(r^2 + u^2 - 2ru \cos \theta)^2}, \quad (\text{B.13})$$

$$R_{A,B} = 4\pi g_s \alpha'^2 \left( \frac{N}{2} \pm N_w \right). \quad (\text{B.14})$$

Let us consider a (topologically trivial) four-cycle  $\Sigma_4$  defined by<sup>6</sup>  $r = \mu$  and  $\theta = \bar{\theta}$ , whose unwarped volume is  $V_4 = \frac{8\pi^2}{3} \bar{r}^4 \sin^4 \bar{\theta}$ . The warped volume is

$$\begin{aligned} \int_{\Sigma_4} H(r, \theta) \sin^4 \theta d\Omega_4 &= V_4 H(\mu, \bar{\theta}) \\ &= V_4 \left[ 1 + \frac{R_A^4}{\mu^4} \left( 1 + \mathcal{O}\left(\frac{u}{\mu}\right) \right) - N_w \frac{u}{\mu} \left( 4 \cos \bar{\theta} - 2 \frac{u}{\mu} + \mathcal{O}\left(\frac{u^2}{\mu^2}\right) \right) \right]. \end{aligned} \quad (\text{B.15})$$

From this result, one can see that a fine-tuning of the embedding can suppress the backreaction, *i.e.* if  $\cos \bar{\theta} \simeq u/(2\mu)$ . On the other hand, a parametric suppression is also clearly visible. Both the factors  $N_w/N$  and  $u/\mu$  can be made small by construction. The former must be small in order for the background geometry to be at all trustworthy. The latter can be made small by arranging for all the four-cycles bearing nonperturbative effects to be far away in units of the separation of the fivebranes. Physically, this means that the four-cycle is sensitive only to the dipole field generated by the A and B stacks.

---

<sup>5</sup>We choose the letter  $u$  in analogy with the setup of the next subsections, where the distance between the A and the B stack is given by the resolution parameter of the resolved conifold.

<sup>6</sup>We choose the letter  $\mu$  in analogy with the (usually complex) parameter appearing in other known embeddings of four-cycles in the conifold [146, 147].

## B.2.2 The conifold and its resolution

In the following we review some relevant definitions and conventions regarding the resolved conifold. The treatment is based on [93, 145]. The (singular) conifold is a cone over  $T^{1,1}$  (the coset space  $SU(2) \times SU(2)/U(1)$ , which is topologically  $S^2 \times S^3$ ). It is defined as the hyperspace in  $\mathbb{C}^4$  that is a solution of the complex constraint

$$\det W \equiv \det \begin{pmatrix} X & U \\ V & Y \end{pmatrix} = XY - VU = 0 \quad (\text{B.16})$$

where  $(X, U, V, Y)$  are coordinates on  $\mathbb{C}^4$ . The resolved conifold can be defined as the zero locus in  $\mathbb{C}^4 \times \mathbb{CP}_1$  of the two linear complex equations

$$\begin{pmatrix} X & U \\ V & Y \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0 \quad (\text{B.17})$$

where  $(\lambda_1, \lambda_2)$  are complex coordinates on  $\mathbb{CP}_1$ , *i.e.* they are identified by  $(\lambda_1, \lambda_2) \simeq (\alpha\lambda_1, \alpha\lambda_2)$ , for every  $\alpha \in \mathbb{C}_*$ . For every  $W \neq 0$  (B.16) and (B.17) are equivalent, but when  $W = 0$ , *i.e.* at the tip,  $(\lambda_1, \lambda_2)$  are arbitrary and (B.17) defines a  $\mathbb{CP}_1 \simeq S^2$ . The radial direction is defined by

$$\text{Tr} W^\dagger W = r^2. \quad (\text{B.18})$$

One can check that the resolved conifold is an  $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$  bundle over  $\mathbb{CP}_1$  with fiber  $\mathbb{C}_2$ . If we define  $\lambda \equiv \lambda_2/\lambda_1$ , then we can choose coordinates on a patch  $H_+ \equiv \{\lambda \neq 0\}$  of the resolved conifold using the following solution of (B.17):

$$W = \begin{pmatrix} -\lambda U & U \\ -\lambda Y & Y \end{pmatrix}. \quad (\text{B.19})$$

Defining  $\tilde{\lambda} \equiv \lambda_1/\lambda_2$ , one can find coordinates on a complementary patch  $H_- \equiv \{\tilde{\lambda} \neq 0\}$  using the following solution:

$$W = \begin{pmatrix} X & -\tilde{\lambda}X \\ V & -\tilde{\lambda}V \end{pmatrix}. \quad (\text{B.20})$$

The complex structure is given by

$$\Omega = dU \wedge dY \wedge d\lambda = dV \wedge dX \wedge d\tilde{\lambda}. \quad (\text{B.21})$$

For later use, we introduce a parametrization of the resolved conifold in terms of real coordinates and give the explicit Kähler metric. We start by noting that a particular solution of (B.17) is given by

$$W_0 = \begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \lambda_0 \\ 0 \end{pmatrix} \Rightarrow \lambda = 0. \quad (\text{B.22})$$

The base of the resolved conifold with respect to  $r$  can be obtained by acting on this solution with two  $SU(2)$  transformations,  $L_1$  and  $L_2$ ,

$$L_i = \begin{pmatrix} \cos \frac{\theta_i}{2} e^{\frac{i}{2}(\psi_i + \phi_i)} & -\sin \frac{\theta_i}{2} e^{-\frac{i}{2}(\psi_i - \phi_i)} \\ \sin \frac{\theta_i}{2} e^{\frac{i}{2}(\psi_i - \phi_i)} & \cos \frac{\theta_i}{2} e^{-\frac{i}{2}(\psi_i + \phi_i)} \end{pmatrix}, \quad i = 1, 2 \quad (\text{B.23})$$

written in terms of Euler angles. This gives

$$W = L_1 W_0 L_2^\dagger, \quad \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = L_2 \begin{pmatrix} \lambda_0 \\ 0 \end{pmatrix} \quad (\text{B.24})$$

which depends only on the combination  $\psi \equiv \psi_1 + \psi_2$ . A Kähler metric on the resolved conifold with resolution parameter  $u$  is given by [145]

$$ds_6^2 = \kappa^{-1}(\rho) d\rho^2 + \frac{1}{9} \kappa(\rho) \rho^2 e_\psi^2 + \frac{1}{6} \rho^2 (e_{\theta_1}^2 + e_{\phi_1}^2) + \frac{1}{6} (\rho^2 + 6u^2) (e_{\theta_2}^2 + e_{\phi_2}^2), \quad (\text{B.25})$$

where

$$\kappa(\rho) = \frac{\rho^2 + 9u^2}{\rho^2 + 6u^2}. \quad (\text{B.26})$$

Here, following [145], we have defined a new radial coordinate  $\rho$  by

$$r^4 = \frac{4}{9}\rho^4 \left( \frac{2}{3}\rho^2 + 6u^2 \right), \quad (\text{B.27})$$

The explicit expression for the  $e$ 's is

$$e_\psi = d\psi + \sum_{i=1}^2 \cos \theta_i d\phi_i, \quad e_{\theta_i} = d\theta_i, \quad e_{\phi_i} = \sin \theta_i d\phi_i. \quad (\text{B.28})$$

### The $\lambda UY$ embedding

In this subsection, we consider a particular holomorphic embedding of a four-cycle. A simple embedding would be  $\lambda = \mu$  because this is trivial to solve for in real coordinates,  $\tan \theta_2 = \mu$  and  $\phi_2 = 0$  for  $\mu \in \mathbb{R}$ . The trouble is that this embedding reaches the tip, and in fact  $r$  is unconstrained. This can also be seen from

$$r^2 = (1 + |\lambda|^2)(|U|^2 + |Y|^2). \quad (\text{B.29})$$

As a result, this embedding does not give us the dipole suppression factor analogous to the  $(u/\mu)$  of appendix B.2.1. The next-simplest embedding (whose defining equation depends on  $r$ ) is

$$\lambda UY = \mu^3, \quad \mu \in \mathcal{R}, \quad (\text{B.30})$$

which in real coordinates gives

$$\psi = 0, \quad \sin(\theta_2) \sin(\theta_1) = 4 \frac{\mu^3}{r^2} \sim \frac{\mu^3}{\rho^3} \quad \text{for large } r. \quad (\text{B.31})$$

After some algebra (in particular, expressing  $d\theta_2$  as a function of  $d\theta_1$  and  $dr$ ) one finds the metric in terms of  $d\rho$ ,  $d\theta_2$ ,  $d\phi_1$  and  $d\phi_2$ . Its determinant  $g_4^{\text{ind}}$  is

independent of  $\phi_{1,2}$  and reads

$$\begin{aligned}
g_4^{ind} &= \frac{\rho^2 \csc^4(\theta_2)}{20736 (9u^2 + \rho^2)^3 (\rho^4(9u^2 + \rho^2) \sin^2(\theta_2) - 54\mu^6)} \times \\
&\left( (6u^2 + \rho^2)(9u^2\rho + \rho^3)^2 \cos(4\theta_2) \right. \\
&- 4 \cos(2\theta_2) (486u^6\rho^2 + 189u^4\rho^4 + u^2(24\rho^6 - 324\mu^6) - 27\mu^6\rho^2 + \rho^8) \\
&\left. + 3(54\rho^2(9u^6 + 2\mu^6) + 189u^4\rho^4 + 864u^2\mu^6 + 24u^2\rho^6 + \rho^8) \right)^2 . \quad (\text{B.32})
\end{aligned}$$

We see that there is a boundary beyond which the sign of the determinant becomes negative, which thus defines the integration boundary in  $\rho, \theta_2$ -space:

$$\begin{aligned}
\rho_{min}(\theta_2) &= \sqrt{3} \mu \sqrt{A - \frac{u^2}{\mu^2} \cdot \left(1 - \frac{u^2}{\mu^2} \frac{1}{A}\right)} \rightarrow \sqrt{3 \cdot 2^{1/3}} \mu \csc^{1/3}(\theta_2) \quad \text{for } \frac{u}{\mu} \ll 1 \\
\text{with : } A &= \sqrt[3]{\sqrt{-\csc^2(\theta_2) \left(2\frac{u^6}{\mu^6} - \csc^2(\theta_2)\right) - \frac{u^6}{\mu^6} + \csc^2(\theta_2)}} . \quad (\text{B.33})
\end{aligned}$$

### B.2.3 The shift of the four-cycle volume

The solution with the branes smeared over the  $S^2$  was obtained in [145]. Later, the solutions with pointlike sources were given in [144]. If the D3-brane stacks are at the north and south pole of the resolution  $S^2$ , respectively, *i.e.*  $\theta_2^A = \pi - \theta_2^B = 0$ , then one finds

$$H = \sum_l (2l + 1) H_l(\rho) \left[ L_A^4 P_l(\cos(\theta_2)) + L_B^4 P_l(\cos(\theta_2)) (-1)^l \right], \quad (\text{B.34})$$

$$H_l = \frac{2}{9u^2} \frac{C_\beta}{\rho^{2+2\beta}} {}_2F_1\left(\beta, 1 + \beta, 1 + 2\beta, -\frac{9u^2}{\rho^2}\right), \quad (\text{B.35})$$

$$C_\beta = \frac{(3u)^{2\beta} \Gamma(1 + \beta)^2}{\Gamma(1 + 2\beta)} \quad \beta = \sqrt{1 + (3/2)l(l + 1)}, \quad (\text{B.36})$$

$$L_{A,B} = \frac{27}{16} 4\pi g_s (\alpha')^2 (N \mp N_w), \quad (\text{B.37})$$

where  ${}_2F_1$  is a hypergeometric function. We want to integrate this warp factor on some supersymmetric four-cycle  $\Sigma_4$ . This gives us an estimate of the inflaton dependence of the gauge kinetic function of a stack of D7-branes wrapping  $\Sigma_4$ .

Using this information and (B.34) and (B.35) we can now calculate the integral

$$\mathcal{V}_{warped} = \int_{\Sigma_4} d\rho d\theta_2 d\phi_1 d\phi_2 \sqrt{-g_4^{ind}} H(\rho, \theta_2) = 4\pi^2 \int_{\Sigma_4} d\rho d\theta_2 \sqrt{-g_4^{ind}} H(\rho, \theta_2) \quad (\text{B.38})$$

numerically, as a function of  $\mu$ . To facilitate this we will expand (B.34) up to  $\ell = 1$ , the dipole term, and take the large  $\rho$  limit

$$H(\rho, \theta_2) = \frac{L^4}{2\rho^4} \left[ 1 + 3(2\ell + 1) \frac{N_w u^2}{N \rho^2} P_\ell(\cos \theta_2) \Big|_{\ell=1} \right] = H_{\ell=0}(\rho) + \delta H_{\ell=1}(\rho, \theta_2) \quad (\text{B.39})$$

We can now calculate

$$\mathcal{V}_{warped}^{(0)} = 4\pi^2 \int_0^\pi d\theta_2 \int_{\rho_{min}(\theta_2)}^{\rho_R} d\rho \sqrt{-g_4^{ind}} H_{\ell=0}(\rho) \quad (\text{B.40})$$

$$\delta \mathcal{V}_{warped}^{\ell=1}(\delta\theta_2) = 4\pi^2 \int_0^\pi d\theta_2 \int_{\rho_{min}(\theta_2)}^{\rho_R} d\rho \sqrt{-g_4^{ind}} \delta H_{\ell=1}(\rho, \theta_2 + \delta\theta_2) \quad (\text{B.41})$$

where  $\rho_R \gg 1$  denotes a UV cutoff to compactify the resolved conifold geometry for the purpose of integration, and  $\delta\theta_2$  denotes the angular misalignment of the D3-brane dipole configuration with respect to the four-cycle symmetry axis at  $\theta_2 = \pi/2$ .

As  $\sqrt{g_4^{ind}}$  is a symmetric function with respect to  $\theta_2 = \pi/2$  and  $\delta H_{\ell=1}(\rho, \theta_2)$  is anti-symmetric with respect to  $\theta_2 = \pi/2$ , we immediately find  $\delta \mathcal{V}_{warped}(\delta\theta_2 = 0) = 0$ . So by fine-tuning a  $\mathbb{Z}_2$ -symmetric configuration we can forbid the  $\ell = 1$  term in the warped volume, whose corrections in this case start with the  $\ell = 2$  quadrupole terms.

We will now display the numerical results for the case  $\delta\theta_2 = -\pi/4$  in which the  $\ell = 1$  term will not vanish under the integral, and compare the scaling with

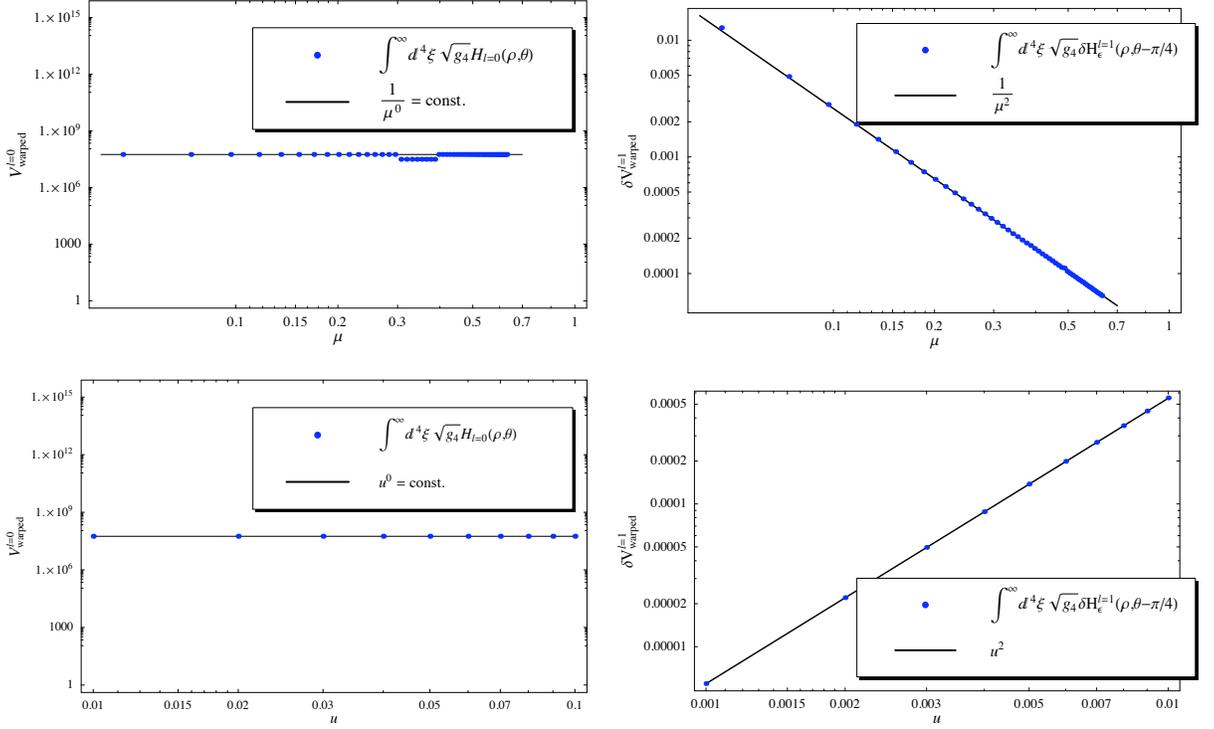


Figure B.2: 1st row: Plot of  $\mathcal{V}_{warped}^{(0)}$  and  $\delta\mathcal{V}_{warped}(\delta\theta_2)$  as functions of  $\mu$  at constant  $u = 0.01$ . 2nd row: Plot of  $\mathcal{V}_{warped}^{(0)}$  and  $\delta\mathcal{V}_{warped}(\delta\theta_2)$  as functions of  $u$  at constant  $\mu = 0.1$ . The leading  $\ell = 0$  term scales as  $u^0\mu^0 = \text{const.}$  while the  $\ell = 1$  dipole term scales as  $(u/\mu)^2$ . Note that the  $\ell = 0$  scaling ensues only in the strictly noncompact limit (*i.e.* when the integration goes all the way  $\rho \rightarrow \infty$ ), while for a finite cutoff, resembling a crude approximation to a compact setting, there remains a weak dependence of the  $\ell = 0$  term on  $\mu$ , of the form  $(u/\mu)^\delta$ , where  $\delta \rightarrow 0$  for  $\rho_{bulk} \rightarrow \infty$ . For the example we have chosen  $\epsilon \equiv N_w/N = 0.1$ .

$\mu$  between  $\delta\mathcal{V}_{warped}(\delta\theta_2)$  and  $\mathcal{V}_{warped}^{(0)}$ . This is displayed in Fig. B.2. We see clearly that the leading  $\ell = 0$  term scales as  $u^0\mu^0 = \text{const.}$  while the  $\ell = 1$  dipole term scales as  $(u/\mu)^2$ . Therefore, the  $\ell = 1$  dipole term has a parametric suppression  $(u/\mu)^2$  relative to the leading  $\ell = 0$  term, and can therefore be made parametrically small (even in the non- $\mathbb{Z}_2$ -symmetric general situation) in the limit where the four-cycle recedes far from the resolution  $S^2$  (*i.e.* in the limit of large  $\mu/u$ ).

Let us finally note that this relative suppression of the  $\ell = 1$  term with  $(u/\mu)^2$

might have been guessed without any integration, as the integration boundary tells us that  $\rho_{min}(\theta_2) \geq \rho_{min}(\pi/2)$ , which corresponds to  $r > 2\mu^{3/2}$  or  $\rho \gtrsim \mu$ , and thus the relative scaling  $u^2/\rho^2$  should be replaced by the scaling  $u^2/\mu^2$ .

### B.3 The Kaluza-Klein Spectrum

In this appendix we obtain the (5+1)-dimensional effective action for a D5-brane wrapped on a two-cycle with  $\int B \neq 0$ . We show how a Kaluza-Klein reduction to four dimensions leads to masses that are suppressed with respect to the fluxless case. We then comment on the consequences of these light KK modes for axion monodromy inflation.

#### B.3.1 The effective theory

The DBI action for a D5-brane is

$$S = T_5 \int d^4x dy dz \sqrt{-\det(G_{ab}^{ind} + \mathcal{F}_{ab})}, \quad (\text{B.42})$$

where  $y, z$  are two coordinates in the internal space, which we take to be toroidal for the purpose of this derivation. The indices are defined as follows: worldvolume indices are  $a, b = 0, \dots, 5$ ; spacetime indices are  $\mu, \nu = 0, \dots, 3$  as usual; ten-dimensional indices are  $M, N = 0, \dots, 9$ ; six-dimensional compact indices are  $m, n = 4, \dots, 9$ ; and indices transverse to the D5-brane are  $i, j = 6, \dots, 9$ . We first expand the square root using

$$\sqrt{\det(M_0 + \delta M)} = \sqrt{\det M_0} \left\{ 1 + \frac{1}{2} \text{Tr}(M_0^{-1} \delta M) + \frac{1}{8} [\text{Tr}(M_0^{-1} \delta M)]^2 \right. \quad (\text{B.43})$$

$$\left. - \frac{1}{4} \text{Tr}(M_0^{-1} \delta M M_0^{-1} \delta M) + \dots \right\}. \quad (\text{B.44})$$

We will consider a background with two-form flux on the two-cycle

$$\int \mathcal{F} = \int B = \int dy \wedge dz B_{yz}(x, y, z) = b(x) = b, \quad (\text{B.45})$$

*i.e.* the four-dimensional axion field  $b(x)$  has a homogeneous expectation value that is approximately constant, up to terms suppressed by the slow-roll parameters. So the background is given by

$$B_{MN} = b\delta_{My}\delta_{Nz} - b\delta_{Mz}\delta_{Ny}, \quad F_{ab} = 0, \quad (\text{B.46})$$

$$ds_{10}^2 = g_{\mu\nu}dx^\mu dx^\nu + g_{yy}dy^2 + g_{zz}dz^2 + 2g_{yz}dydz + g_{ij}dy^i dy^j. \quad (\text{B.47})$$

Hence

$$(M_0)_{ab} = \begin{pmatrix} g_{\mu\nu} & & & \\ & g_{yy} & g_{yz} + b & \\ & g_{zy} - b & g_{zz} & \\ & & & \end{pmatrix}. \quad (\text{B.48})$$

The perturbations are

$$(\delta M)_{ab} = \partial_a X^i \partial_b X^j (g_{ij} + B_{ij}) + F_{ab} + \delta B_{ab}. \quad (\text{B.49})$$

The calculation is simplified by the block-diagonal form of the background  $M_0$ . The  $2 \times 2$  block is the sum of a symmetric and an antisymmetric piece that we call  $S$  and  $A$  respectively. We have that

$$\det(A + S) = \det(A) + \det(S), \quad (\text{B.50})$$

$$(S + A)^{-1} = S^{-1} \frac{\det(S)}{\det(A) + \det(S)} + A^{-1} \frac{\det(A)}{\det(A) + \det(S)}, \quad (\text{B.51})$$

which substantially simplifies the calculation. Using (B.43) we get at leading order

$$S = T_5 \int d^4x dy dz \sqrt{-g_4} \sqrt{g_2 + b^2} \left[ 1 + \frac{1}{2} \partial_\mu X^i \partial^\mu X_i \right. \quad (\text{B.52})$$

$$\left. + \frac{1}{2} \frac{g_2}{g_2 + b^2} (\partial_y X^i \partial^y X_i + \partial_z X^i \partial^z X_i) \right. \quad (\text{B.53})$$

$$\left. + \frac{1}{2} \frac{2b}{g_2 + b^2} (\partial_y X^i \partial_z X^j \delta B_{ij} + F_{yz} + \delta B_{yz}) \right] + \dots, \quad (\text{B.54})$$

where  $g_2 \equiv g_{yy}g_{zz} - g_{yz}^2$ . After a KK reduction one finds the four-dimensional kinetic and potential terms, in the first line, as well as the Kaluza-Klein mass terms, in the second line. The Kaluza-Klein masses in the presence of fluxes are

$$m_{bKK}^2 = \frac{g_2}{g_2 + b^2} m_{KK}^2, \quad (\text{B.55})$$

where  $m_{KK}$  are the Kaluza-Klein masses in the absence of fluxes. This leads to the central point of this appendix: for  $b \gg 1$ , the Kaluza-Klein masses are *suppressed* by a factor of  $\sqrt{g_2}/b \simeq L^2/b \ll 1$ .<sup>7</sup> This phenomenon is intuitively understood in the T-dual picture in which flux becomes the angle of the D-brane. A large flux means that the T-dual brane winds around the torus many times, and thus becomes quite long. The Kaluza-Klein reduction of the fields living on the worldvolume of the T-dual brane therefore produces b-suppressed Kaluza-Klein masses.

### B.3.2 Effects of the light Kaluza-Klein modes

Throughout this paper we have been careful to work in parameter ranges for which the typical Kaluza-Klein mass scale  $m_{KK}$  obeys  $m_{KK} \gg H$ , as required for a consistent four-dimensional analysis of inflation. However, from (B.55) we learn that a subclass of Kaluza-Klein modes, namely those associated with transverse excitations of the fivebrane, have considerably smaller masses,  $m_{bKK} \ll m_{KK}$ . For the numerical examples we have considered, we find that, very roughly,  $m_{bKK} \sim (f_c/f)H$ , where  $f_c$  is a fiducial value of the decay constant,  $f_c \sim 10^{-2}M_p$ . Therefore, for constructions with small values of  $f$ , the transverse excitations of the fivebrane can be lighter than  $H$ .

---

<sup>7</sup>We have assumed for simplicity that the internal space is isotropic, with typical size  $L\sqrt{\alpha'}$ .

We leave a comprehensive study of this constraint for future work, as a proper implementation plausibly requires a more explicit compact model that we have been able to present in this work. In particular, one should carefully compute the Kaluza-Klein mass, incorporating anisotropy in the geometry, warping, and, as we have explained above, the effect of worldvolume two-forms. To accomplish this, one needs a reasonably explicit construction of the warped throat region, of the two-cycle within the throat, and of the gluing of the throat into the compact space, which are beyond the scope of this work.

In this appendix, we will restrict ourselves to some qualitative statements that explain how our inflationary analysis can be consistent even in parameter regimes for which  $m_{bKK}$  is slightly smaller than  $H$ . Broadly speaking, one might worry about corrections to the inflationary Lagrangian, and about new contributions to the cosmological perturbations. Concerning the first point, we remark that the excitations of the fivebrane depend on the inflaton expectation value only through their masses. Therefore, the primary correction to the background evolution from these light modes would come if large numbers of Kaluza-Klein particles were produced by the time-dependent background. In practice, the particle production is negligible, as can be seen by computing the adiabatic parameter  $\dot{m}_{bKK}/m_{bKK}^2$  and substituting the constraints on the volume, and hence on the Kaluza-Klein mass, from §4.6.

More generally, let us stress that only a small subclass of the Kaluza-Klein modes (a small portion of the tower of excitations of the fivebrane) have masses smaller than  $H$ . From the viewpoint of the inflationary analysis, these fields constitute a small number of harmless spectators. These light fields will fluctuate, absorbing energy, but this yields a very small correction unless the number of

fields approaches  $(M_p/H)^2$ . Moreover, any entropy perturbations produced by these fields can turn into visible isocurvature perturbations only if their decays are distinct from that of the inflaton. Although we have not specified a concrete reheating mechanism, one can argue that the most straightforward scenario involves visible sector degrees of freedom that are well-separated in the compact space from the inflationary fivebranes. Thus, we expect that excitations of the fivebranes will not give visible isocurvature perturbations, because they must first decay [148, 149] to degrees of freedom localized in the inflationary throat, just as the inflaton does, and will plausibly do so with rather similar couplings, as the modes correspond to small excitations of the NS5-brane that drives inflation.

## B.4 Numerical Examples

In this appendix, we specify two different sets of intersection numbers and show the relevant formulas for the volumes. For these two toy models, we explicitly performed the moduli stabilization outlined in §4.6.5, finding numerical values leading to the dot in figure 4.8.

### B.4.1 Intersection numbers: set I

We consider as a toy-model Calabi-Yau manifold one with  $H_+^{1,1} = \text{span}(\omega^L, \omega^+)$  for the orientifold-even homology two-cycles and  $H_-^{1,1} = \text{span}(\omega^-)$  for the orientifold-odd homology two-cycles. We assume the following simple set of

intersection numbers

$$c_{LLL} = c_{LL+} = c_{+--} = 1, \quad (\text{B.56})$$

with all the others vanishing. We believe that, although very simplistic, the above toy model captures the relevant features of more realistic constructions. Notice that the intersection numbers in a basis for the homology of the covering space of the orientifold, *i.e.* without a definite parity with respect to the orientifold projection, are just linear combinations of those given above.

Using the standard relations

$$\mathcal{V}_E = \frac{1}{6} c_{\alpha\beta\gamma} v^\alpha v^\beta v^\gamma, \quad \tau_\alpha = \partial_{v^\alpha} \mathcal{V}_E = \frac{1}{2} c_{\alpha\beta\gamma} v^\beta v^\gamma, \quad (\text{B.57})$$

one finds

$$v_L = \sqrt{2\tau_+}, \quad v_+ = \frac{\tau_L - \tau_+}{\sqrt{2\tau_+}}, \quad (\text{B.58})$$

and

$$\mathcal{V}_E = \frac{\sqrt{2\tau_+}}{2} \tau_L - \frac{\sqrt{2}}{6} \tau_+^{3/2}. \quad (\text{B.59})$$

## B.4.2 Intersection numbers: set II

Again assuming  $H_+^{1,1} = \text{span}(\omega^L, \omega^+)$  and  $H_-^{1,1} = \text{span}(\omega^-)$ , we consider the intersection numbers

$$c_{LLL} = c_{L++} = c_{+--} = 1, \quad (\text{B.60})$$

with all the others vanishing. We find

$$v_L = \frac{1}{\sqrt{2}} \left( (\tau_L + \tau_+)^{1/2} + (\tau_L - \tau_+)^{1/2} \right), \quad v_+ = \frac{1}{\sqrt{2}} \left( (\tau_L + \tau_+)^{1/2} - (\tau_L - \tau_+)^{1/2} \right), \quad (\text{B.61})$$

and

$$\mathcal{V}_E = \frac{1}{3\sqrt{2}} \left( (\tau_L + \tau_+)^{3/2} + (\tau_L - \tau_+)^{3/2} \right). \quad (\text{B.62})$$

## BIBLIOGRAPHY

- [1] A. H. Guth, "The Inflationary Universe: A Possible Solution To The Horizon And Flatness Problems," *Phys. Rev. D* **23**, 347 (1981).
- [2] A. 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 (1982).
- [3] A. Albrecht and P. J. Steinhardt, "Cosmology For Grand Unified Theories With Radiatively Induced Symmetry Breaking," *Phys. Rev. Lett.* **48**, 1220 (1982).
- [4] D. H. Lyth and A. Riotto, *Phys. Rept.* **314**, 1 (1999) [hep-ph/9807278].
- [5] S. Kachru, R. Kallosh, A. Linde, J. M. Maldacena, L. P. McAllister and S. P. Trivedi, "Towards inflation in string theory," *JCAP* **0310**, 013 (2003) [arXiv:hep-th/0308055].
- [6] D. Baumann, A. Dymarsky, S. Kachru, I. R. Klebanov and L. McAllister, *JHEP* **1006**, 072 (2010) [arXiv:1001.5028 [hep-th]].
- [7] D. Baumann, A. Dymarsky, I. R. Klebanov *et al.*, *Phys. Rev. Lett.* **99**, 141601 (2007). [arXiv:0705.3837 [hep-th]].
- [8] A. Krause, E. Pajer, *JCAP* **0807**, 023 (2008). [arXiv:0705.4682 [hep-th]].
- [9] D. Baumann, A. Dymarsky, I. R. Klebanov and L. McAllister, *JCAP* **0801**, 024 (2008) [arXiv:0706.0360 [hep-th]].
- [10] E. Silverstein and A. Westphal, "Monodromy in the CMB: Gravity Waves and String Inflation," arXiv:0803.3085 [hep-th].
- [11] L. McAllister, E. Silverstein and A. Westphal, "Gravity Waves and Linear Inflation from Axion Monodromy," arXiv:0808.0706 [hep-th].
- [12] E. Komatsu *et al.* [WMAP Collaboration], "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations:Cosmological Interpretation," *Astrophys. J. Suppl.* **180**, 330 (2009) [arXiv:0803.0547 [astro-ph]].
- [13] D. Larson *et al.*, arXiv:1001.4635 [astro-ph.CO].

- [14] G. R. Dvali and S. H. H. Tye, Phys. Lett. B **450**, 72 (1999) [arXiv:hep-ph/9812483].
- [15] G. R. Dvali, Q. Shafi, S. Solganik, [hep-th/0105203].
- [16] C. P. Burgess, M. Majumdar, D. Nolte *et al.*, JHEP **0107**, 047 (2001) [hep-th/0105204].
- [17] D. Baumann, A. Dymarsky, I. R. Klebanov, J. M. Maldacena, L. P. McAllister and A. Murugan, "On D3-brane potentials in compactifications with fluxes and wrapped D-branes," JHEP **0611**, 031 (2006) [arXiv:hep-th/0607050].
- [18] C. P. Burgess, J. M. Cline, K. Dasgupta *et al.*, JHEP **0703**, 027 (2007) [hep-th/0610320].
- [19] S. Panda, M. Sami and S. Tsujikawa, Phys. Rev. D **76**, 103512 (2007) [arXiv:0707.2848 [hep-th]].
- [20] A. Ali, R. Chingangbam, S. Panda and M. Sami, Phys. Lett. B **674**, 131 (2009) [arXiv:0809.4941 [hep-th]].
- [21] H. Y. Chen and J. O. Gong, Phys. Rev. D **80**, 063507 (2009) [arXiv:0812.4649 [hep-th]].
- [22] A. Ali, A. Deshamukhya, S. Panda *et al.*, [arXiv:1010.1407 [hep-th]].
- [23] H. Y. Chen, J. O. Gong and G. Shiu, JHEP **0809**, 011 (2008) [arXiv:0807.1927 [hep-th]].
- [24] H. Y. Chen, J. O. Gong, K. Koyama and G. Tasinato, JCAP **1011**, 034 (2010) [arXiv:1007.2068 [hep-th]].
- [25] L. Hoi and J. M. Cline, Phys. Rev. D **79**, 083537 (2009) [arXiv:0810.1303 [hep-th]].
- [26] D. Baumann, A. Dymarsky, S. Kachru, I. R. Klebanov and L. McAllister, JHEP **0903**, 093 (2009) [arXiv:0808.2811 [hep-th]].
- [27] A. Ceresole, G. Dall'Agata, R. D'Auria and S. Ferrara, Phys. Rev. D **61**, 066001 (2000) [arXiv:hep-th/9905226].

- [28] D. Baumann and L. McAllister, *Phys. Rev. D* **75**, 123508 (2007) [arXiv:hep-th/0610285].
- [29] Z. D. Bai, *Statistica Sinica* **9**, 61 (1999).
- [30] B. Freivogel, M. Kleban, M. Rodriguez Martinez and L. Susskind, *JHEP* **0603**, 039 (2006) [arXiv:hep-th/0505232].
- [31] E. Silverstein and D. Tong, *Phys. Rev. D* **70**, 103505 (2004) [arXiv:hep-th/0310221].
- [32] M. Alishahiha, E. Silverstein and D. Tong, *Phys. Rev. D* **70**, 123505 (2004) [arXiv:hep-th/0404084].
- [33] R. Bean, S.E. Shandera, S.H. Tye, and J. Xu, *JCAP* **0705**, 004 (2007) [arXiv:hep-th/0702107].
- [34] S. Kecskemeti, J. Maiden, G. Shiu and B. Underwood, *JHEP* **0609**, 076 (2006) [arXiv:hep-th/0605189].
- [35] E. Pajer, *JCAP* **0804**, 031 (2008) [arXiv:0802.2916 [hep-th]].
- [36] S. Kachru, R. Kallosh, A. Linde and S. P. Trivedi, “de Sitter vacua in string theory,” *Phys. Rev. D* **68**, 046005 (2003) [arXiv:hep-th/0301240].
- [37] C. Gordon, D. Wands, B. A. Bassett and R. Maartens, *Phys. Rev. D* **63**, 023506 (2001) [arXiv:astro-ph/0009131].
- [38] S. H. Tye and J. Xu, *Phys. Lett. B* **683**, 326 (2010) [arXiv:0910.0849 [hep-th]].
- [39] S. Groot Nibbelink and B. J. W. van Tent, *Class. Quant. Grav.* **19**, 613 (2002) [arXiv:hep-ph/0107272].
- [40] B. J. W. van Tent and S. Groot Nibbelink, arXiv:hep-ph/0111370.
- [41] N. Agarwal et al., work in progress.
- [42] R. Easther and L. McAllister, *JCAP* **0605**, 018 (2006). [hep-th/0512102].
- [43] S. Dimopoulos, S. Kachru, J. McGreevy and J. Wacker, *JCAP* **0808**, 003 (2008). [hep-th/0507205].

- [44] D. Baumann, arXiv:0907.5424 [hep-th].
- [45] J. Frazer and A. R. Liddle, arXiv:1101.1619 [astro-ph.CO].
- [46] D. Baumann and D. Green, arXiv:1109.0292 [hep-th].
- [47] M. Dias, J. Frazer and A. R. Liddle, arXiv:1203.3792 [astro-ph.CO].
- [48] N. Agarwal, R. Bean, L. McAllister, G. Xu, JCAP **1109**, 002 (2011) [arXiv:1103.2775 [astro-ph.CO]].
- [49] D. A. Easson, R. Gregory, D. F. Mota, G. Tasinato and I. Zavala, JCAP **0802**, 010 (2008) [arXiv:0709.2666 [hep-th]].
- [50] D. Langlois, S. Renaux-Petel, D. A. Steer and T. Tanaka, Phys. Rev. Lett. **101** (2008) 061301 [arXiv:0804.3139 [hep-th]].
- [51] D. Langlois, S. Renaux-Petel, D. A. Steer and T. Tanaka, Phys. Rev. D **78** (2008) 063523 [arXiv:0806.0336 [hep-th]].
- [52] X. Gao and B. Hu, JCAP **0908** (2009) 012 [arXiv:0903.1920 [astro-ph.CO]].
- [53] D. Langlois, S. Renaux-Petel and D. A. Steer, JCAP **0904** (2009) 021 [arXiv:0902.2941 [hep-th]].
- [54] S. Mizuno, F. Arroja, K. Koyama and T. Tanaka, Phys. Rev. D **80** (2009) 023530 [arXiv:0905.4557 [hep-th]].
- [55] S. Renaux-Petel, JCAP **0910** (2009) 012 [arXiv:0907.2476 [hep-th]].
- [56] S. Mizuno, F. Arroja and K. Koyama, Phys. Rev. D **80** (2009) 083517 [arXiv:0907.2439 [hep-th]].
- [57] R. Gregory and D. Kaviani, JHEP **1201** (2012) 037 [arXiv:1107.5522 [hep-th]].
- [58] S. Renaux-Petel, S. Mizuno and K. Koyama, JCAP **1111** (2011) 042 [arXiv:1108.0305 [astro-ph.CO]].
- [59] D. Langlois and S. Renaux-Petel, JCAP **0804**, 017 (2008) [arXiv:0801.1085 [hep-th]].

- [60] M. Sasaki and E. D. Stewart, *Prog. Theor. Phys.* **95** (1996) 71 [[astro-ph/9507001](#)].
- [61] A. A. Starobinsky, *JETP Lett.* **42** (1985) 152 [*Pisma Zh. Eksp. Teor. Fiz.* **42** (1985) 124].
- [62] M. Sasaki and T. Tanaka, *Prog. Theor. Phys.* **99** (1998) 763 [[gr-qc/9801017](#)].
- [63] D. H. Lyth and Y. Rodriguez, *Phys. Rev. Lett.* **95** (2005) 121302 [[astro-ph/0504045](#)].
- [64] D. Wands, N. Bartolo, S. Matarrese and A. Riotto, *Phys. Rev. D* **66** (2002) 043520 [[astro-ph/0205253](#)].
- [65] G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, *Phys. Rev. D* **72** (2005) 083507 [[astro-ph/0410486](#)].
- [66] G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, *Phys. Rev. D* **73** (2006) 083521 [[astro-ph/0504508](#)].
- [67] G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, *Phys. Rev. D* **73** (2006) 083522 [[astro-ph/0506704](#)].
- [68] G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, *Phys. Rev. D* **76** (2007) 083512 [[astro-ph/0511041](#)].
- [69] D. Langlois and F. Vernizzi, *JCAP* **0702** (2007) 017 [[astro-ph/0610064](#)].
- [70] S. Renaux-Petel and G. Tasinato, *JCAP* **0901** (2009) 012 [[arXiv:0810.2405](#) [[hep-th](#)]].
- [71] J.-L. Lehners and S. Renaux-Petel, *Phys. Rev. D* **80** (2009) 063503 [[arXiv:0906.0530](#) [[hep-th](#)]].
- [72] D. J. Mulryne, D. Seery and D. Wesley, *JCAP* **1001** (2010) 024 [[arXiv:0909.2256](#) [[astro-ph.CO](#)]].
- [73] D. J. Mulryne, D. Seery and D. Wesley, *JCAP* **1104** (2011) 030 [[arXiv:1008.3159](#) [[astro-ph.CO](#)]].
- [74] M. Dias and D. Seery, *Phys. Rev. D* **85** (2012) 043519 [[arXiv:1111.6544](#) [[astro-ph.CO](#)]].

- [75] G. J. Anderson, D. J. Mulryne and D. Seery, arXiv:1205.0024 [astro-ph.CO].
- [76] H.-C. Lee, M. Sasaki, E. D. Stewart, T. Tanaka and S. Yokoyama, JCAP **0510** (2005) 004 [astro-ph/0506262].
- [77] C. T. Byrnes and D. Wands, Phys. Rev. D **74** (2006) 043529 [astro-ph/0605679].
- [78] Z. Lalak, D. Langlois, S. Pokorski and K. Turzyski, JCAP **0707** (2007) 014 [arXiv:0704.0212 [hep-th]].
- [79] K.-Y. Choi, L. M. H. Hall and C. van de Bruck, JCAP **0702** (2007) 029 [astro-ph/0701247].
- [80] A. Avgoustidis, S. Cremonini, A.-C. Davis, R. H. Ribeiro, K. Turzyski and S. Watson, JCAP **1202** (2012) 038 [arXiv:1110.4081 [astro-ph.CO]].
- [81] S. Cremonini, Z. Lalak and K. Turzyski, JCAP **1103** (2011) 016 [arXiv:1010.3021 [hep-th]].
- [82] A. Achucarro, J.-O. Gong, S. Hardeman, G. A. Palma and S. P. Patil, JCAP **1101** (2011) 030 [arXiv:1010.3693 [hep-ph]].
- [83] D. Baumann and D. Green, JCAP **1109** (2011) 014 [arXiv:1102.5343 [hep-th]].
- [84] G. Shiu and J. Xu, Phys. Rev. D **84** (2011) 103509 [arXiv:1108.0981 [hep-th]].
- [85] S. Cespedes, V. Atal and G. A. Palma, JCAP **1205** (2012) 008 [arXiv:1201.4848 [hep-th]].
- [86] A. Avgoustidis, S. Cremonini, A. -C. Davis, R. H. Ribeiro, K. Turzyski and S. Watson, arXiv:1203.0016 [hep-th].
- [87] A. Achucarro, V. Atal, S. Cespedes, J. -O. Gong, G. A. Palma and S. P. Patil, arXiv:1205.0710 [hep-th].
- [88] X. Gao, D. Langlois and S. Mizuno, arXiv:1205.5275 [hep-th].
- [89] S. Tsujikawa, D. Parkinson and B. A. Bassett, Phys. Rev. D **67** (2003) 083516 [astro-ph/0210322].

- [90] C. M. Peterson and M. Tegmark, Phys. Rev. D **83** (2011) 023522 [arXiv:1005.4056 [astro-ph.CO]].
- [91] B. J. W. van Tent, Ph.D. thesis, Utrecht University, 2002.
- [92] S. Weinberg, Oxford, UK: Oxford Univ. Pr. (2008).
- [93] P. Candelas and X. C. de la Ossa, "Comments on Conifolds," Nucl. Phys. B **342**, 246 (1990).
- [94] D. Marsh, L. McAllister and T. Wrase, JHEP **1203**, 102 (2012) [arXiv:1112.3034 [hep-th]].
- [95] J. Meyers and N. Sivanandam, Phys. Rev. D **83** (2011) 103517 [arXiv:1011.4934 [astro-ph.CO]].
- [96] J. Elliston, D. J. Mulryne, D. Seery and R. Tavakol, JCAP **1111** (2011) 005 [arXiv:1106.2153 [astro-ph.CO]].
- [97] J. Meyers and N. Sivanandam, Phys. Rev. D **84** (2011) 063522 [arXiv:1104.5238 [astro-ph.CO]].
- [98] D. Seery, D. J. Mulryne, J. Frazer and R. H. Ribeiro, arXiv:1203.2635 [astro-ph.CO].
- [99] C. M. Peterson and M. Tegmark, arXiv:1111.0927 [astro-ph.CO].
- [100] D. Wands, K. A. Malik, D. H. Lyth and A. R. Liddle, Phys. Rev. D **62** (2000) 043527 [astro-ph/0003278].
- [101] N. Bartolo, S. Matarrese and A. Riotto, Phys. Rev. D **64** (2001) 123504 [astro-ph/0107502].
- [102] X. Chen and Y. Wang, Phys. Rev. D **81** (2010) 063511 [arXiv:0909.0496 [astro-ph.CO]];  
X. Chen and Y. Wang, JCAP **1004** (2010) 027 [arXiv:0911.3380 [hep-th]];  
X. Chen and Y. Wang, arXiv:1205.0160 [hep-th].
- [103] X. Chen, R. Easther and E. A. Lim, "Generation and Characterization of Large Non-Gaussianities in Single Field Inflation," JCAP **0804**, 010 (2008) [arXiv:0801.3295 [astro-ph]].

- [104] J. Martin and C. Ringeval, "Superimposed Oscillations in the WMAP Data?," *Phys. Rev. D* **69**, 083515 (2004) [arXiv:astro-ph/0310382].
- [105] J. Martin and C. Ringeval, "Exploring the superimposed oscillations parameter space," *JCAP* **0501**, 007 (2005) [arXiv:hep-ph/0405249].
- [106] R. Easther, W. H. Kinney and H. Peiris, "Observing trans-Planckian signatures in the cosmic microwave background," *JCAP* **0505**, 009 (2005) [arXiv:astro-ph/0412613].
- [107] X. Chen, R. Easther and E. A. Lim, "Large non-Gaussianities in single field inflation," *JCAP* **0706** (2007) 023 [arXiv:astro-ph/0611645].
- [108] J. Hamann, L. Covi, A. Melchiorri and A. Slosar, "New constraints on oscillations in the primordial spectrum of inflationary perturbations," *Phys. Rev. D* **76** (2007) 023503 [arXiv:astro-ph/0701380].
- [109] R. Bean, X. Chen, G. Hailu, S. H. Tye and J. Xu, "Duality Cascade in Brane Inflation," *JCAP* **0803**, 026 (2008) [arXiv:0802.0491 [hep-th]].
- [110] C. Pahud, M. Kamionkowski and A. R. Liddle, "Oscillations in the inflaton potential?," arXiv:0807.0322 [astro-ph].
- [111] D. Baumann and L. McAllister, "Advances in Inflation in String Theory," arXiv:0901.0265 [hep-th].
- [112] K. Freese and D. Spolyar, "Chain inflation: 'Bubble bubble toil and trouble'," *JCAP* **0507**, 007 (2005) [arXiv:hep-ph/0412145].
- [113] K. Freese, J. T. Liu and D. Spolyar, "Chain inflation via rapid tunneling in the landscape," arXiv:hep-th/0612056.
- [114] D. Chialva and U. H. Danielsson, "Chain inflation revisited," *JCAP* **0810**, 012 (2008) [arXiv:0804.2846 [hep-th]].
- [115] S. Weinberg, *Cosmology*, Oxford University Press, Oxford, UK (2008).
- [116] Bateman Manuscript Project, *Higher Transcendental Functions, Volume 2*, McGraw-Hill, New York, USA, (1953-55).
- [117] J. Lesgourgues, A. A. Starobinsky and W. Valkenburg, "What do

- WMAP and SDSS really tell about inflation?," JCAP **0801**, 010 (2008) [arXiv:0710.1630 [astro-ph]].
- [118] R. Easther, B. R. Greene, W. H. Kinney and G. Shiu, "Inflation as a probe of short distance physics," Phys. Rev. D **64**, 103502 (2001) [arXiv:hep-th/0104102].
- [119] N. Kaloper, M. Kleban, A. E. Lawrence and S. Shenker, "Signatures of short distance physics in the cosmic microwave background," Phys. Rev. D **66**, 123510 (2002) [arXiv:hep-th/0201158].
- [120] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Butterworth-Heinemann, Oxford, UK (2002). "Course of Theoretical Physics. Vol. 3: Quantum Mechanics,"
- [121] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," JHEP **0305** (2003) 013 [arXiv:astro-ph/0210603].
- [122] P. D. Meerburg, J. P. van der Schaar and P. S. Corasaniti, "Signatures of Initial State Modifications on Bispectrum Statistics," JCAP **0905**, 018 (2009) [arXiv:0901.4044 [hep-th]].
- [123] G. Hinshaw *et al.* [WMAP Collaboration], "Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Temperature analysis," Astrophys. J. Suppl. **170**, 288 (2007) [arXiv:astro-ph/0603451].
- [124] A. Lewis, A. Challinor and A. Lasenby, "Efficient Computation of CMB anisotropies in closed FRW models", Astrophys. J. **538**, 473 (2000) [arXiv:astro-ph/9911177].
- [125] The code is available at <http://camb.info/>
- [126] <http://lambda.gsfc.nasa.gov/>
- [127] A. Lewis and S. Bridle, "Cosmological parameters from CMB and other data: a Monte-Carlo approach", Phys. Rev. D **66**, 103511 (2002) [arXiv:astro-ph/0205436].
- [128] The code is available at <http://cosmologist.info/cosmomc/>

- [129] J. Polchinski, *String Theory. Vol. 2: Superstring Theory and Beyond*, Cambridge University Press, Cambridge, UK (1998).
- [130] T. W. Grimm and J. Louis, “The effective action of  $N = 1$  Calabi-Yau orientifolds,” *Nucl. Phys. B* **699**, 387 (2004) [arXiv:hep-th/0403067].
- [131] X. G. Wen and E. Witten, “World Sheet Instantons And The Peccei-Quinn Symmetry,” *Phys. Lett. B* **166**, 397 (1986).
- [132] M. Dine and N. Seiberg, “Nonrenormalization Theorems in Superstring Theory,” *Phys. Rev. Lett.* **57**, 2625 (1986).
- [133] P. Svrcek and E. Witten, “Axions in string theory,” *JHEP* **0606** (2006) 051 [arXiv:hep-th/0605206].
- [134] M. Bertolini, V. L. Campos, G. Ferretti, P. Fre', P. Salomonson and M. Trigiante, “Supersymmetric 3-branes on smooth ALE manifolds with flux,” *Nucl. Phys. B* **617**, 3 (2001) [arXiv:hep-th/0106186].
- [135] D. Baumann, A. Dymarsky, I. R. Klebanov and L. McAllister, “Towards an Explicit Model of D-brane Inflation,” *JCAP* **0801**, 024 (2008) [arXiv:0706.0360 [hep-th]].
- [136] O. DeWolfe, S. Kachru and M. Mulligan, “A Gravity Dual of Metastable Dynamical Supersymmetry Breaking,” *Phys. Rev. D* **77**, 065011 (2008) [arXiv:0801.1520 [hep-th]].
- [137] D. J. Gross and J. H. Sloan, “The Quartic Effective Action for the Heterotic String,” *Nucl. Phys. B* **291**, 41 (1987).
- [138] A. Kehagias and H. Partouche, “On the exact quartic effective action for the type iib superstring,” *Phys. Lett. B* **422**, 109 (1998) [arXiv:hep-th/9710023].
- [139] T. Banks, M. Dine, P. J. Fox and E. Gorbatov, “On the possibility of large axion decay constants,” *JCAP* **0306**, 001 (2003) [arXiv:hep-th/0303252].
- [140] E. Witten, “Non-Perturbative Superpotentials In String Theory,” *Nucl. Phys. B* **474**, 343 (1996) [arXiv:hep-th/9604030].
- [141] T. W. Grimm, “Non-Perturbative Corrections and Modularity in  $N=1$

- Type IIB Compactifications," JHEP **0710**, 004 (2007) [arXiv:0705.3253 [hep-th]].
- [142] D. Robles-Llana, M. Rocek, F. Saueressig, U. Theis and S. Vandoren, "Non-perturbative corrections to 4D string theory effective actions from  $SL(2, Z)$  duality and supersymmetry," Phys. Rev. Lett. **98**, 211602 (2007) [arXiv:hep-th/0612027].
- [143] V. Balasubramanian, P. Berglund, J. P. Conlon and F. Quevedo, JHEP **0503**, 007 (2005) [arXiv:hep-th/0502058].
- [144] I. R. Klebanov and A. Murugan, "Gauge/Gravity Duality and Warped Resolved Conifold," JHEP **0703**, 042 (2007) [arXiv:hep-th/0701064].
- [145] L. A. Pando Zayas and A. A. Tseytlin, "3-branes on resolved conifold," JHEP **0011**, 028 (2000) [arXiv:hep-th/0010088].
- [146] S. Kuperstein, "Meson spectroscopy from holomorphic probes on the warped deformed conifold," JHEP **0503**, 014 (2005) [arXiv:hep-th/0411097].
- [147] P. Ouyang, "Holomorphic D7-branes and flavored  $N = 1$  gauge theories," Nucl. Phys. B **699**, 207 (2004) [arXiv:hep-th/0311084].
- [148] N. Barnaby, C. P. Burgess and J. M. Cline, "Warped reheating in brane-antibrane inflation," JCAP **0504**, 007 (2005) [arXiv:hep-th/0412040].
- [149] L. Kofman and P. Yi, "Reheating the universe after string theory inflation," Phys. Rev. D **72**, 106001 (2005) [arXiv:hep-th/0507257].
- [150] L. P. McAllister, Thesis SLAC-R-750.
- [151] D. Baumann, Thesis