Oscillations in the CMB
from Axion Monodromy Inflation

Raphael Flauger,[1] Liam McAllister,[2] Enrico Pajer,[2]
Alexander Westphal,[3] and Gang Xu[2]

[1] *Department of Physics, University of Texas at Austin, Austin, TX 78712*
[2] *Department of Physics, Cornell University, Ithaca, NY 14853*
[3] *Department of Physics, Stanford University, Stanford, CA 94305*

We study the CMB observables in axion monodromy inflation. These well-motivated scenarios for inflation in string theory have monomial potentials over super-Planckian field ranges, with superimposed sinusoidal modulations from instanton effects. Such periodic modulations of the potential can drive resonant enhancements of the correlation functions of cosmological perturbations, with characteristic modulations of the amplitude as a function of wavenumber. We give an analytical result for the scalar power spectrum in this class of models, and we determine the limits that present data places on the amplitude and frequency of modulations. Then, incorporating an improved understanding of the realization of axion monodromy inflation in string theory, we perform a careful study of microphysical constraints in this scenario. We find that detectable modulations of the scalar power spectrum are commonplace in well-controlled examples, while resonant contributions to the bispectrum are undetectable in some classes of examples and detectable in others. We conclude that resonant contributions to the spectrum and bispectrum are a characteristic signature of axion monodromy inflation that, in favorable cases, could be detected in near-future experiments.

Contents

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* [4], in which the potential energy is not periodic under transport around an angular direction in the configuration space. The first examples [4] 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 [5].

The axion monodromy inflation scenario of [5] 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 [6]. 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 [7] 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 §2 by describing the classical evolution of the homogeneous background in axion monodromy inflation with a modulated linear potential. We then solve, in §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, 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.* [8, 9, 10, 11, 12, 13, 14]). Then, in §5 and §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 α' perturbation expansions, successful moduli stabilization, and bounds on higher-derivative terms. In §7 we combine the observational and theoretical constraints, with results presented in figure 7.

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 §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 [15] 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

3

shift symmetries $a \rightarrow a + constant$ 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 [5], 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 [5], an NS5-brane wrapped on a two-cycle Σ 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 ϕ,

$$V = \mu^3 \phi \,, \tag{1.1}$$

with μ 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 [5], 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 [5] that there exist classes of examples in which nonperturbative effects are practically negligible, but we expect – as explained in detail in §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 §2-§4, by studying a phenomenological potential that captures the essential effects; and second, in §5 and §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.

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 Λ^4, and write our potential as in [5],

$$V(\phi) = \mu^3 \phi + \Lambda^4 \cos\left(\frac{\phi}{f}\right) = \mu^3 \left[\phi + bf \cos\left(\frac{\phi}{f}\right)\right] \,, \tag{2.1}$$

4

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

$$\ddot{\phi} + 3H\dot{\phi} + \mu^3 - \mu^3 b \sin\left(\frac{\phi}{f}\right) = 0 \,. \tag{2.2}$$

To solve (2.2), we begin with two approximations. Monotonicity of the potential requires[1] $b < 1$, and as we will see in §4, for the case $b < 1$ observational constraints in fact imply $b \ll 1$. This suggests treating the oscillatory term in 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/ϕ.

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

$$\dot{\phi}_0 = -\sqrt{\frac{\mu^3}{3\phi_0}} \,, \tag{2.3}$$

$$\ddot{\phi}_1 + \sqrt{3\mu^3\phi_0}\,\dot{\phi}_1 - \frac{\mu^3}{2\phi_0}\phi_1 = \mu^3 \sin\left(\frac{\phi_0}{f}\right) \,, \tag{2.4}$$

where we have neglected terms of higher order in M_p/ϕ and we have made use of the slow roll approximation for ϕ_0.[2] Using equation (2.3), we can rewrite equation (2.4) with ϕ_0 as an independent variable instead of t, yielding

$$\phi_1'' - 3\phi_0\phi_1' - \frac{3}{2}\phi_1 = 3\phi_0 \sin\left(\frac{\phi_0}{f}\right) \,. \tag{2.5}$$

where primes denote derivatives with respect to ϕ_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 ϕ_0 everywhere except in the driving term. The inhomogeneous solution is then given by

$$\phi_1(t) = f\frac{6f\phi_*}{(2+3f^2)^2 + 36f^2\phi_*^2}\left[-(2+3f^2)\sin\left(\frac{\phi_0(t)}{f}\right) + 6f\phi_* \cos\left(\frac{\phi_0(t)}{f}\right)\right] \,, \tag{2.6}$$

---

[1] 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 [16, 17, 18] 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.

[2] In approximating $\sin(\phi/f) \simeq \sin(\phi_0/f)$ on the right hand side of (2.4), 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 (2.7), ϕ_1 is of order $f^2\phi_*$. Hence the mild assumption $bf\phi_* \ll 1$ justifies this approximation.

where ϕ_* denotes the value of the field ϕ_0 at the time at which the pivot scale k_* exits the horizon. Assuming 60 e-foldings of inflation, this happens around $\phi_* \simeq 11 M_p$. 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

$$\phi(t) = \phi_0(t) + b\phi_1(t) = \phi_0(t) + bf \frac{3f\phi_*}{1 + (3f\phi_*)^2}\left[-\sin\left(\frac{\phi_0(t)}{f}\right) + 3f\phi_*\cos\left(\frac{\phi_0(t)}{f}\right)\right], \quad (2.7)$$

with $\phi_0(t)$ given by

$$\phi_0(t) = \left[\phi_*^{3/2} - \frac{\sqrt{3}}{2}\mu^{3/2}(t - t_*)\right]^{2/3}. \quad (2.8)$$

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 $\phi_* \simeq 11 M_p$. 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.

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

$$\Delta_{\mathcal{R}}^2(k) = \Delta_{\mathcal{R}}^2(k_*)\left(\frac{k}{k_*}\right)^{n_s - 1}\left[1 + \delta n_s \cos\left(\frac{\phi_k}{f}\right)\right] \approx \Delta_{\mathcal{R}}^2\left(\frac{k}{k_*}\right)^{n_s - 1 + \frac{\delta n_s}{\ln(k/k_*)}\cos\left(\frac{\phi_k}{f}\right)}, \quad (3.1)$$

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 δn_s is given by

$$\delta n_s = \frac{12b}{\sqrt{(1 + (3f\phi_*)^2)}}\sqrt{\frac{\pi}{8}\coth\left(\frac{\pi}{2f\phi_*}\right)}f\phi_*, \tag{3.2}$$

where

$$\phi_k = \sqrt{\phi_*^2 - 2\ln k/k_*} \simeq \phi_* - \frac{\ln k/k_*}{\phi_*} \tag{3.3}$$

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

In §3.1 we will give a derivation of this result that makes no further approximations. In §3.2 and §3.3 we will present two additional derivations of (3.1) 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 (3.1). Let us at this point briefly summarize the scales that will be relevant for our discussion in the next subsections.

Given the potential (2.1), 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 [7]. Using the slow roll equation of motion and the COBE normalization,

$$3H\dot{\phi} \simeq -V'(\phi), \quad \dot{\phi}^2 \simeq \frac{2}{3}\epsilon V, \quad V \simeq 5\cdot 10^{-7}\,\epsilon\,M_p^4, \tag{3.4}$$

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

$$\frac{\omega}{H} \simeq \frac{M_p^2}{\phi f} \simeq \sqrt{2\epsilon}\frac{M_p}{f} > 1, \tag{3.5}$$

$$\frac{\omega}{M_p} \simeq \sqrt{\frac{2\epsilon V}{3}}\frac{1}{fM_{pl}} < 1, \tag{3.6}$$

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 §5 and §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/ϕ expansion. Furthermore, δn_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.

7

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

$$\delta g_{ij}(\mathbf{x}, t) = 2a(t)^2 \mathcal{R}(\mathbf{x}, t)\delta_{ij} \, . \tag{3.7}$$

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

$$\mathcal{R}(\mathbf{x}, t) = \int \frac{d^3 k}{(2\pi)^{3/2}} \left[\mathcal{R}_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} \alpha(\mathbf{k}) + \mathcal{R}_k(t)^* e^{-i\mathbf{k}\cdot\mathbf{x}} \alpha^*(\mathbf{k}) \right] \, , \tag{3.8}$$

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

$$\langle \alpha(\mathbf{k})\alpha^*(\mathbf{k}') \rangle = \delta(\mathbf{k} - \mathbf{k}') \, , \tag{3.9}$$

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[3]

$$\frac{d^2 \mathcal{R}_k}{dx^2} - \frac{2(1 + 2\epsilon + \delta)}{x} \frac{d\mathcal{R}_k}{dx} + \mathcal{R}_k = 0 \, , \tag{3.10}$$

where $x \equiv -k\tau$, with the conformal time τ 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

$$\left| \mathcal{R}_k^{(o)} \right|^2 = 2\pi^2 \frac{\Delta_\mathcal{R}^2(k)}{k^3} \, . \tag{3.11}$$

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,

$$\epsilon = \epsilon_0 + \epsilon_1 + \mathcal{O}(b^2) \, , \tag{3.12}$$

---

[3]We use the same definitions for the slow roll parameters as in [19], *i.e.* $\epsilon \equiv -\frac{\dot{H}}{H^2}$, $\delta \equiv \frac{\ddot{H}}{2H\dot{H}}$. δ 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 $\eta_V = 4\epsilon - \eta$.

$$\delta = \delta_0 + \delta_1 + \mathcal{O}(b^2) \,. \tag{3.13}$$

For the background solution (2.7), the first-order terms are given by

$$\epsilon_1 = -\frac{3bf}{\phi_*[1 + (3f\phi_*)^2]} \left[\cos\left(\frac{\phi_0}{f}\right) + (3f\phi_*)\sin\left(\frac{\phi_0}{f}\right) \right] \,, \tag{3.14}$$

$$\delta_1 = -\frac{3b}{[1 + (3f\phi_*)^2]} \left[\sin\left(\frac{\phi_0}{f}\right) - (3f\phi_*)\cos\left(\frac{\phi_0}{f}\right) \right] \,. \tag{3.15}$$

We now consider an ansatz of the form

$$\mathcal{R}_k = \mathcal{R}_{k,0}^{(o)} \left[i\sqrt{\frac{\pi}{2}} x^{\nu_0} H_{\nu_0}^{(1)}(x) + g(x) \right] \,. \tag{3.16}$$

Here the index ν_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[4]

$$\mathcal{R}_{k,0}^{(o)} = \mp i\sqrt{\frac{\mu^3 \phi_k^3}{6}} \frac{1}{k^{3/2}} \,, \tag{3.17}$$

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

$$\left| \mathcal{R}_k^{(o)} \right|^2 = \left| \mathcal{R}_{k,0}^{(o)} \right|^2 \left[1 + 2\,\mathrm{Re}\,g(0) \right] \approx \left| \mathcal{R}_{k,0}^{(o)} \right|^2 e^{2\,\mathrm{Re}\,g(0)} = \left| \mathcal{R}_{k,0}^{(o)} \right|^2 \left(\frac{k}{k_*}\right)^{\frac{2\,\mathrm{Re}\,g(0)}{\ln(k/k_*)}} \,. \tag{3.18}$$

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

$$\frac{d^2 g}{dx^2} - \frac{2}{x}\frac{dg}{dx} + g = 2e^{ix}(2\epsilon_1 + \delta_1) \,. \tag{3.19}$$

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 ϵ_1 is suppressed relative to δ_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 ϵ_1 on the right hand side of equation (3.19). Furthermore, it turns out to be convenient to rewrite δ_1 using trigonometric identities. Ignoring an unimportant phase, one finds

$$\delta_1 = -\frac{3b}{\sqrt{1 + (3f\phi_*)^2}} \cos\left(\frac{\phi_0}{f}\right) \,. \tag{3.20}$$

---

[4] 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.

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 $r(x) \equiv \mathrm{Re}\,(g(x))$, equation (3.19) becomes

$$\frac{d^2 r}{dx^2} - \frac{2}{x}\frac{dr}{dx} + r = -\frac{6b}{\sqrt{1+(3f\phi_*)^2}}\cos(x)\cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right) . \tag{3.21}$$

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

$$r(0) = \frac{6b|\mathcal{I}(f\phi_*)|}{\sqrt{1+(3f\phi_*)^2}}\cos\left(\frac{\phi_k}{f} + \beta(f\phi_*)\right) , \tag{3.22}$$

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

$$\mathcal{I}(f\phi_*) = \frac{\pi}{2}\int_0^\infty dx\, J_{\frac{3}{2}}(x) J_{-\frac{1}{2}}(x)\, x^{\frac{i}{f\phi_*}} . \tag{3.23}$$

Written in this form, the integral can be recognized as a Weber-Schafheitlin integral and can be done analytically (see *e.g.* [20]). One finds

$$|\mathcal{I}| = \sqrt{\frac{\pi}{8}\coth\left(\frac{\pi}{2f\phi_*}\right) f\phi_*} . \tag{3.24}$$

Combining equations (3.18), (3.22) and (3.24), we finally obtain an expression for δn_s,

$$\delta n_s = \frac{2r(0)}{\cos\left(\frac{\phi_k}{f}\right)} = \frac{12b}{\sqrt{1+(3f\phi_*)^2}}\sqrt{\frac{\pi}{8}\coth\left(\frac{\pi}{2f\phi_*}\right) f\phi_*} . \tag{3.25}$$

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 §3.2 and §3.3. A comparison between our analytical result for δn_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 [21] is shown in Figure 1.

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 $f \ll M_p$, 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 (3.23) 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

$$\mathcal{I}_c = \int_0^\infty dx\, \frac{(\sin x - x\cos x)\cos x}{x^2}\cos\left(\frac{\ln x}{f\phi_*}\right) , \tag{3.26}$$

$$\mathcal{I}_s = \int_0^\infty dx\, \frac{(\sin x - x\cos x)\cos x}{x^2}\sin\left(\frac{\ln x}{f\phi_*}\right) . \tag{3.27}$$
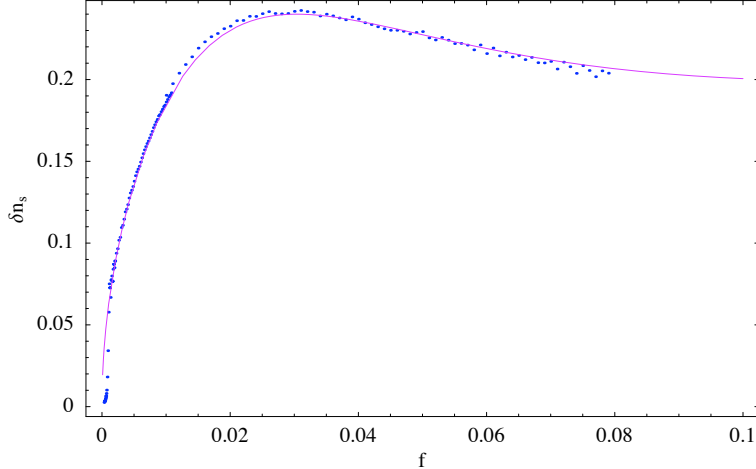
10

Figure 1: The solid line is the analytical result for δn_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 [21].

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 $f\phi_*$

$$\mathcal{I}_c = \sqrt{\frac{\pi}{8} f\phi_*} \sin\left[\frac{1 + \ln(2f\phi_*)}{f\phi_*} - \frac{\pi}{4}\right], \tag{3.28}$$

$$\mathcal{I}_s = \sqrt{\frac{\pi}{8} f\phi_*} \cos\left[\frac{1 + \ln(2f\phi_*)}{f\phi_*} - \frac{\pi}{4}\right], \tag{3.29}$$

which leads to

$$|\mathcal{I}| = \sqrt{\mathcal{I}_c^2 + \mathcal{I}_s^2} = \sqrt{\frac{\pi}{8} f\phi_*}. \tag{3.30}$$

This agrees with our previous result, equation (3.24), 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.[5] 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.

---

[5]This factor of two can be understood from momentum conservation, as will become clear in §3.3.

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 (3.16), where $g(x)$ is the solution of the equation

$$\frac{d^2g}{dx^2} - \frac{2}{x}\frac{dg}{dx} + g = 2e^{ix}\delta_1 \, , \tag{3.31}$$

with δ_1 again given by

$$\delta_1 = -\frac{3b}{\sqrt{1+(3f\phi_*)^2}}\cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right) \, , \tag{3.32}$$

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

$$g(x) = c_k^{(+)}\left(i\sqrt{\frac{\pi}{2}}x^{\frac{3}{2}}H_{3/2}^{(1)}(x)\right) + c_k^{(-)}\left(-i\sqrt{\frac{\pi}{2}}x^{\frac{3}{2}}H_{3/2}^{(2)}(x)\right) \, , \tag{3.33}$$

where $c_k^{(\pm)}$ are momentum dependent coefficients. The solution for equation (3.31) can also be written explicitly as

$$\begin{aligned} g(x) &= (x\cos x - \sin x)\int_x^{\infty}\frac{2e^{iy}(\cos y + y\sin y)}{y^2}\delta_1 \\ &+ (\cos x + x\sin x)\int_x^{\infty}\frac{2e^{iy}(\sin y - y\cos y)}{y^2}\delta_1 \, . \end{aligned} \tag{3.34}$$

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

$$g(x) = \frac{1}{2}(\mathcal{I}^2 + i\mathcal{I}^1)\left(i\sqrt{\frac{\pi}{2}}x^{\frac{3}{2}}H_{3/2}^{(1)}(x)\right) + \frac{1}{2}(\mathcal{I}^2 - i\mathcal{I}^1)\left(-i\sqrt{\frac{\pi}{2}}x^{\frac{3}{2}}H_{3/2}^{(2)}(x)\right) \, , \tag{3.35}$$

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

$$\mathcal{I}^1 = -\frac{6b}{\sqrt{1+(3f\phi_*)^2}}\int_0^{\infty}\frac{e^{iy}(\cos y + y\sin y)}{y^2}\cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right) \, , \tag{3.36}$$

$$\mathcal{I}^2 = -\frac{6b}{\sqrt{1+(3f\phi_*)^2}}\int_0^{\infty}\frac{e^{iy}(\sin y - y\cos y)}{y^2}\cos\left(\frac{\phi_k}{f} + \frac{\ln x}{f\phi_*}\right) \, . \tag{3.37}$$

---

[6]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.

In the saddle point approximation these evaluate to

$$\mathcal{I}^1 = i\mathcal{I}^2 = -\frac{6b}{\sqrt{(1 + (3f\phi_*)^2)}}\sqrt{\frac{\pi}{8}}f\phi_* e^{-i\left(\frac{\phi_k}{f} - \frac{1 + \ln 2f\phi_*}{f\phi_*} + \frac{\pi}{4}\right)}. \tag{3.38}$$

Combining equations (3.16), (3.35), and (3.38), we finally find that the curvature perturbation for $x \ll \frac{1}{2f\phi_*}$ takes the form

$$\mathcal{R}_k = \mathcal{R}_{k,0}^{(o)}\left(i\sqrt{\frac{\pi}{2}}x^{\nu_0}H_{\nu_0}^{(1)}(x) - c_k^{(-)}i\sqrt{\frac{\pi}{2}}x^{\nu_0}H_{\nu_0}^{(2)}(x)\right), \tag{3.39}$$

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

$$c_k^{(-)} = \frac{6b}{\sqrt{(1 + (3f\phi_*)^2)}}\sqrt{\frac{\pi}{8}}f\phi_* e^{-i\left(\frac{\phi_k}{f}\right)}. \tag{3.40}$$

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 (3.39) 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 [22, 23, 8, 9, 10]. The departure from the Bunch-Davies state is of course quantified by $c_k^{(-)}$.

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

$$\Delta H^{(2)} = \frac{1}{2}V''(\phi_0(t)) \cdot \delta\phi^2, \tag{3.41}$$

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

$$H[\delta\phi_{\vec{p}}] = \frac{1}{2}\delta\dot{\phi}_{\vec{p}}^2 + \frac{1}{2}\vec{p}^2\delta\phi_{\vec{p}}^2 + \frac{1}{2}V''(\phi_0(t)) \cdot \delta\phi_{\vec{p}}^2, \tag{3.42}$$

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,

$$H[\delta\phi_p] = \frac{1}{2}\delta\dot{\phi}_p^2 + \frac{1}{2}\omega_p^2\delta\phi_p^2 + \underbrace{\frac{1}{2}V''(\phi_0(t)) \cdot \delta\phi_p^2}_{\Delta H^{(2)}[\delta\phi_p]}. \tag{3.43}$$

Now we compare this to the perturbed harmonic oscillator in one-dimensional quantum mechanics,

$$H = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega_0^2 x^2 + \frac{1}{2}\delta\omega(t)^2 x^2. \tag{3.44}$$

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

$$H = \frac{1}{2}\omega_0\left(p^2 + q^2\right) + \underbrace{\frac{\delta\omega(t)^2}{2\omega_0}q^2}_{\Delta H^{(2)}}, \tag{3.45}$$

where for our case of a periodic perturbation periodic with frequency ω we have

$$\delta\omega(t)^2 = \delta\omega^2\cos(\omega t). \tag{3.46}$$

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

$$\Delta H^{(2)}(t) = \frac{\delta\omega(t)^2}{2\omega_0}q^2 = \frac{\delta\omega^2}{4\omega_0}q^2(e^{i\omega t} + e^{-i\omega t}) \equiv F\left(e^{i\omega t} + e^{-i\omega t}\right), \tag{3.47}$$

in the notation of equations (40.1) through (40.9) of [24]. 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

14

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

$$\Delta H^{(2)}(t) = \frac{\delta\omega(t)^2}{2\omega_p}\delta\varphi_{\vec{p}}^2 = \frac{\delta\omega^2}{4\omega_p}\delta\varphi_{\vec{p}}^2(e^{i\omega t} + e^{-i\omega t}) \equiv F\left(e^{i\omega t} + e^{-i\omega t}\right). \tag{3.48}$$

For the transition matrix element one then finds

$$\langle p, -p|\Delta H^{(2)}|0\rangle = F_{20}(e^{i\omega t} + e^{-i\omega t}) \quad \text{with} \quad F_{20} = \frac{\delta\omega^2}{4\omega_p}\langle 0|\frac{a^2}{\sqrt{2}}\left(\frac{a + a^\dagger}{\sqrt{2}}\right)^2|0\rangle = \frac{\delta\omega^2}{4\sqrt{2}\omega_p} \tag{3.49}$$

Here we have used that

$$|p, -p\rangle = \frac{(a^\dagger)^2}{\sqrt{2}}|0\rangle, \tag{3.50}$$

and

$$\langle 0|a^2(a + a^\dagger)^2|0\rangle = \langle 0|a^2(a^\dagger)^2|0\rangle = 2. \tag{3.51}$$

If the energy of the two-particle state $E_{2p} = 2k/a$ were not too close to the perturbation frequency ω, we could use time-dependent perturbation theory with the above matrix element and obtain the first order transition probability $P_{0\rightarrow 2k}$,

$$
\begin{aligned}
P_{0\rightarrow 2k} &= \left| -i \int^t dt' \langle k, -k|\Delta H^{(2)}(t')|0\rangle e^{i\omega_{20}t'} \right|^2 \\
&= 2|F_{20}|^2 \frac{\left(\frac{2k}{a}\right)^2 + \omega^2 + \left[\left(\frac{2k}{a}\right)^2 - \omega^2\right]\cos(2\omega t)}{\left[\left(\frac{2k}{a}\right)^2 - \omega^2\right]^2} \\
&= \frac{\delta\omega^4}{16(k/a)^2} \frac{\left(\frac{2k}{a}\right)^2 + \omega^2 + \left[\left(\frac{2k}{a}\right)^2 - \omega^2\right]\cos(2\omega t)}{\left[\left(\frac{2k}{a}\right)^2 - \omega^2\right]^2},
\end{aligned} \tag{3.52}
$$

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 Δt_{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 [24]. One finds that on resonance the transition probability is

$$P_{0\rightarrow 2k} = \frac{1}{2}\left[1 - \cos\left(2\Omega t\right)\right] = \frac{1}{2}\left[1 - \cos\left(\frac{\delta\omega^2}{2\sqrt{2}k/a}t\right)\right], \quad \text{where} \quad \Omega \equiv F_{20}. \tag{3.53}$$

15

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 Δt_{res} during which a two-particle state of comoving momentum k stays in near-resonance. We will follow the analysis in [7] 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 (3.52). 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 ω only for a time roughly estimated to be (for the relative phase shifting from $-\pi$ to π)

$$\Delta t_1 \sim \frac{2\pi}{\Delta\omega} . \tag{3.54}$$

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

$$\Delta t_2 \sim \frac{2\Delta\omega}{\omega H} \tag{3.55}$$

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,

$$\Delta t_{res} \equiv \Delta t_1 = \Delta t_2 \sim 2\sqrt{\pi\frac{H}{\omega}}H^{-1} . \tag{3.56}$$

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

$$P_{0\to 2k} \simeq \frac{1}{2}\left[1 - \cos\left(\sqrt{2}\frac{\delta\omega^2}{\omega}\sqrt{\pi\frac{H}{\omega}}H^{-1}\right)\right] . \tag{3.57}$$

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

$$\delta\omega(t)^2 = V''(\phi_0(t)) = \frac{\Lambda^4}{f^2}\cos\left(\frac{\phi_0(t)}{f}\right) = \delta\omega^2\cos(\omega t) \quad , \quad \omega = \frac{H}{f\phi_0} . \tag{3.58}$$

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

$$P_{0\to 2k} \simeq \frac{\pi}{2}\frac{\Lambda^8}{f^4\omega^2 H^2}\cdot\frac{H}{\omega} = \frac{\pi}{2}\frac{\Lambda^8 f\phi_*^3}{f^2 H^4} = \frac{\pi}{2}\frac{9\Lambda^8 f\phi_*^3}{f^2\mu^6\phi_*^2} = \frac{36\pi}{8}b^2 f\phi_* , \tag{3.59}$$

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

Next, because $P_{0\to 2k}$ 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

$$u_k(out) = c^{(-)}u_{-k} + c^{(+)}u_k , \tag{3.60}$$

16

whereas the original Bunch-Davies in-vacuum had modes

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

(3.61)

We therefore find that

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

(3.62)

In comparing these results with the general treatment of the Mukhanov-Sasaki equation above, we see by looking at (3.39) and (3.18) that we can identify

$$u_k = i \sqrt{\frac{\pi}{2}} x^{\nu_0} H_{\nu_0}^{(1)}(x)$$

(3.63)

$$u_{-k} = -i \sqrt{\frac{\pi}{2}} x^{\nu_0} H_{\nu_0}^{(2)}(x)$$

(3.64)

and thus from (3.18) we conclude that

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

(3.65)

which agrees with the general result (3.25) in the appropriate limit $\omega > H$ and $\delta\omega \ll H$, corresponding to $f\phi_* < 1$, where $\coth(\pi/2f\phi_*) \to 1$.

Note that in calculating the transition probability we lose information about the phase of the transition matrix element as given in (3.49). Therefore, if we estimate the population coefficient $c^{(-)}$ from $\sqrt{P_{0 \to 2k}}$, 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 δn_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.

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 [7]. We then present an estimate for the size of the non-Gaussianity for the model (2.1).

The three-point function can be calculated as [25]

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

(3.66)

17

where H_I is the interacting part of the Hamiltonian. H_I was calculated for a generic potential (see e.g. [25, 7]) at cubic order in the perturbations; it takes the form

$$
\begin{aligned}
H_I \;=\; & -\int d^3x\Big[a\epsilon^2\mathcal{R}\mathcal{R}'^2 + a\epsilon^2\mathcal{R}(\partial\mathcal{R})^2 - 2\epsilon\mathcal{R}'(\partial\mathcal{R})(\partial\chi) \\
& + \frac{a}{2}\epsilon\eta'\mathcal{R}^2\mathcal{R}' + \frac{\epsilon}{2a}(\partial\mathcal{R})(\partial\chi)(\partial^2\chi) + \frac{\epsilon}{4a}(\partial^2\mathcal{R})(\partial\chi)^2\Big]\,,
\end{aligned}
\tag{3.67}
$$

where ∂ denote space derivatives,

$$
\chi \equiv a^2\epsilon\partial^{-2}\dot{\mathcal{R}}\,,
\tag{3.68}
$$

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 η than in terms of δ.

We would like to stress that (3.67) is exact for arbitrary values of the slow roll parameters ϵ and η. Substituting H_I into (3.66) produces six terms, plus an additional term coming from a field redefinition. For the modulated linear potential (2.1), ϵ is small, as in standard slow roll inflation. On the other hand, contrary to the standard slow-roll approximation, $\dot\eta$ can be much larger than ϵ^2. This suggests that the leading term comes from the $\epsilon\dot\eta$ term in the Hamiltonian.[7] Hence we have [11, 7]

$$
\langle\mathcal{R}(t,\mathbf{k}_1)\mathcal{R}(t,\mathbf{k}_2)\mathcal{R}(t,\mathbf{k}_3)\rangle \simeq i\left(\prod_i u_i(\tau_{end})\right)\times
$$

$$
\int_{-\infty}^{\tau_{end}} d\tau\,\epsilon\eta'a^2\left(u_1^*(\tau)u_2^*(\tau)\frac{d}{d\tau}u_3^*(\tau) + \text{sym}\right)\delta^3(\mathbf{K})(2\pi)^3 + c.c.\,.
\tag{3.70}
$$

As in [7], we parameterize the non-Gaussianity as

$$
\langle\mathcal{R}(\tau,\mathbf{k}_1)\mathcal{R}(\tau,\mathbf{k}_2)\mathcal{R}(\tau,\mathbf{k}_3)\rangle \equiv \frac{G(k_1,k_2,k_3)}{(k_1k_2k_3)^3}\delta^3(\mathbf{K})\,\Delta_{\mathcal{R}}^4(2\pi)^7\,,
\tag{3.71}
$$

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

$$
\frac{G(k_1,k_2,k_3)}{k_1k_2k_3} = f_{res}\sin\left(\frac{2}{\phi f}\ln K + \text{phase}\right)
\tag{3.72}
$$

Following [7] and comparing (3.70), (3.71) and (3.72), we obtain the estimate

$$
f_{res} \simeq \frac{3\,\dot\eta_1}{8H\sqrt{\phi f}}\,,
\tag{3.73}
$$

---

[7]In (3.9) of [25] this term was written as

$$
\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho}\dot{\mathcal{R}}\mathcal{R}^2\frac{d}{dt}\left(\frac{\ddot\phi}{2\dot\phi\dot\rho} + \frac{\dot\phi^2}{4\dot\rho^2}\right)\,,
\tag{3.69}
$$

which can be reduced to the term in (3.67) using $H' = -\dot\phi^2/2$.

where we have again used the notation $\eta = \eta_0 + b\eta_1 + \ldots$. Using the background solution obtained in §2, it is straightforward to find

$$\dot{\eta}_1 \simeq 2\dot{\delta}_1 \simeq -\sqrt{\frac{\mu^3}{3\phi_*}} \frac{6b}{f[1 + (3f\phi_*)^2]} \left[\cos\left(\frac{\phi_0}{f}\right) + (3f\phi_*)\sin\left(\frac{\phi_0}{f}\right) \right] . \tag{3.74}$$

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 (3.74) is always negligible, i.e. $3f\phi \ll 1$. Hence our estimate for the non-Gaussianity is

$$f_{res} \simeq \frac{9b}{4(f\phi)^{3/2}} = \frac{9}{4}b\left(\frac{\omega}{H}\right)^{3/2} . \tag{3.75}$$

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 §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 [26]), 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 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 [6]. 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

$$T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}) \tag{4.1}$$

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

$$\langle a_{\ell m}^{th} a_{\ell' m'}^{th\,*} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'} \,, \tag{4.2}$$

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_ℓ, encoding their correlations are of interest. These multipole coefficients C_ℓ can be estimated from the measured expansion coefficients $a_{\ell m}$ via

$$C_\ell^{\text{sky}} = \frac{1}{2\ell+1} \sum_m |a_{\ell m}|^2 \,. \tag{4.3}$$

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

$$\langle C_\ell^{\text{sky,th}} \rangle \equiv \frac{1}{2\ell+1} \left\langle \sum_m |a_{\ell m}^{th}|^2 \right\rangle = C_\ell \,. \tag{4.4}$$

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

$$\left\langle \left(\frac{C_\ell^{\text{sky,th}} - C_\ell}{C_\ell} \right)^2 \right\rangle = \frac{2}{2\ell+1} \,. \tag{4.5}$$

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 ℓ (specifically, for $\ell \leq 32$), and a pseudo-C_ℓ estimator for higher ℓ. For details we refer the reader to [27] 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 [28, 29][8], using the primordial power spectrum derived in the previous section in the form

$$\Delta_{\mathcal{R}}^2(k) = \Delta_{\mathcal{R}}^2(k_*) \left(\frac{k}{k_*} \right)^{n_s - 1 + \frac{\delta n_s}{\ln(k/k_*)} \cos\left(\frac{\phi_k}{f} + \Delta\varphi \right)} \,. \tag{4.6}$$

---

[8]Of course, we modify the CAMB code to calculate all the multipole coefficients rather than calculating some and interpolating.

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 [30]. The power spectrum in our model contains additional parameters beyond those of the WMAP five-year ΛCDM 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 δn_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 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 ΛCDM 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 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 δn_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 2. To convert the resulting observational constraints on δn_s as a function of f into constraints on the microscopic parameter bf as a function of f, we make use of equation (3.25). The resulting 68% and 95% contours in the bf-f plane are shown in the right plot of Figure 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 3.
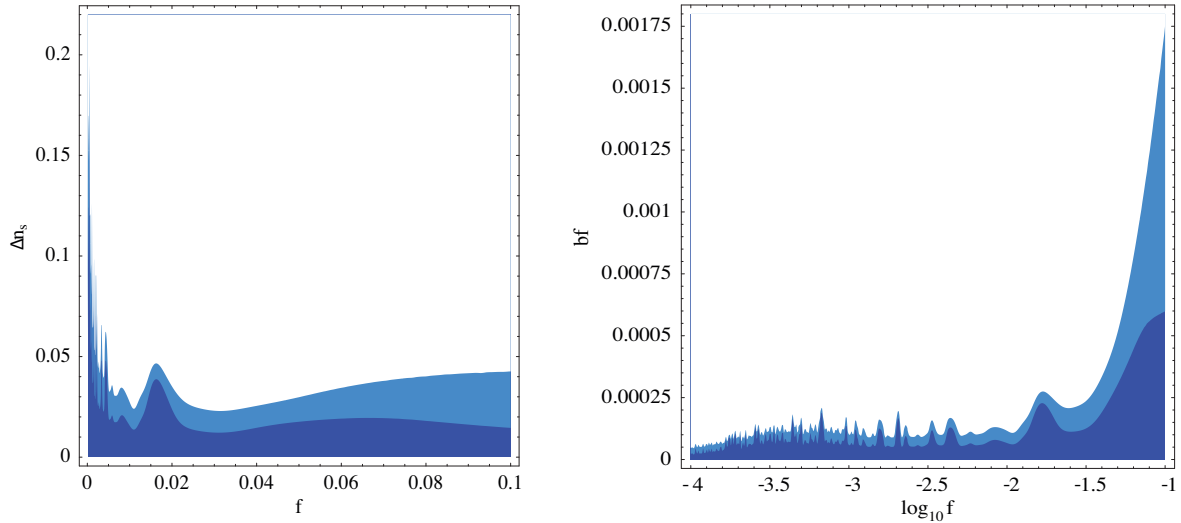
Figure 2: This plot shows the 68% and 95% likelihood contours in the δn_s-f, and bf-$\log_{10} f$ plane, respectively, from the five-year WMAP data on the temperature angular power spectrum.
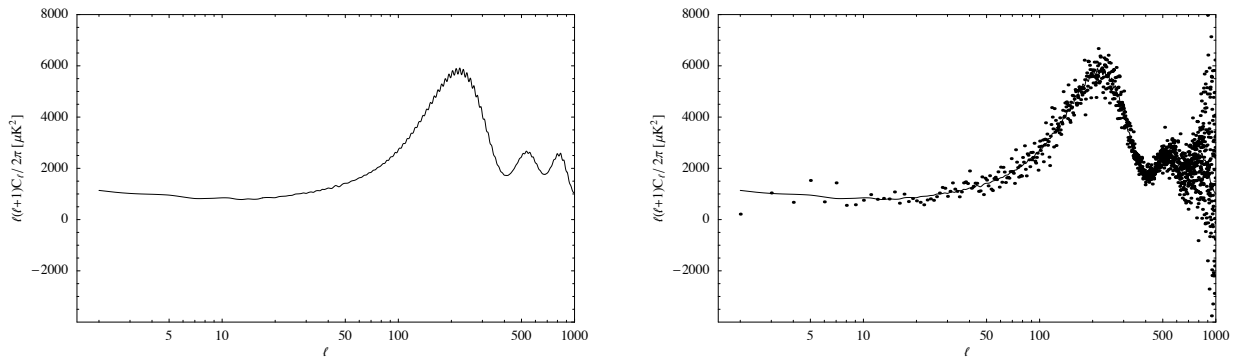

Figure 3: The left plot shows the angular power spectrum for the best fit point $f = 6.67 \times 10^{-4}$, and $\delta n_s = 0.17$. The right plot shows the angular power spectrum for the best fit point together with the unbinned WMAP five-year data.

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 δn_s at least for a certain range of f.

The most straightforward way to demonstrate this degeneracy between $\Omega_B h^2$ and δn_s arising for certain 'resonant' values of f is to present a likelihood plot in the Ω_B-δn_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. It shows that marginalizing over $\Omega_B h^2$ is necessary to obtain correct exclusion contours on δn_s and f. That marginalization over the phase is necessary can easily
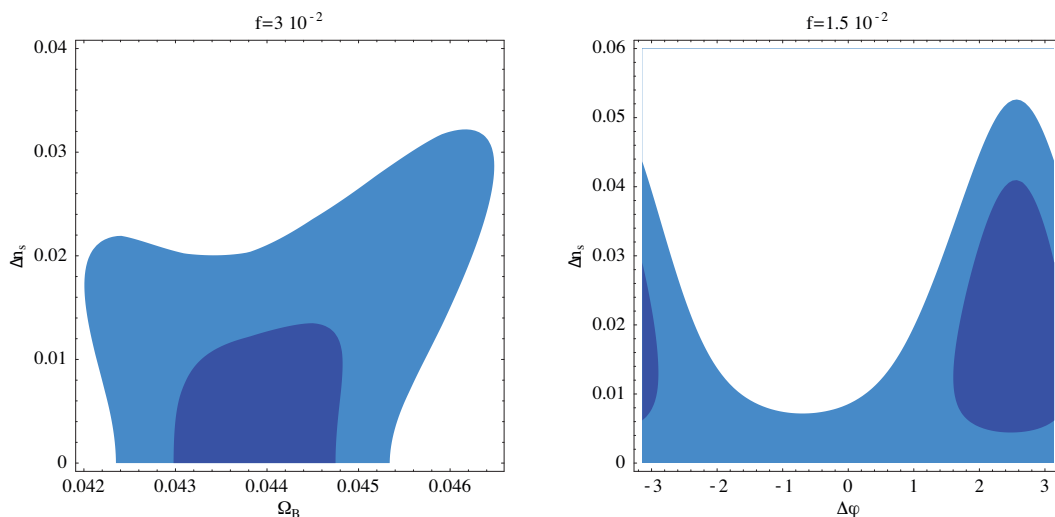

Figure 4: These plots show the 68% and 95% likelihood contours for the five-year WMAP data on the temperature angular power spectrum in the δn_s - $\Omega_B h^2$ plane and δn_s - $\Delta\phi$ plane for an axion decay constant of $f = 3 \times 10^{-2}$ and $f = 1.5 \times 10^{-2}$, respectively.

be seen from a likelihood plot in the δn_s-$\Delta\varphi$ plane. This is shown in the plot on the right side of Figure 4.

We have also performed a Markov chain Monte Carlo analysis for the model using the publicly available CosmoMC code [31], [32]. 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.
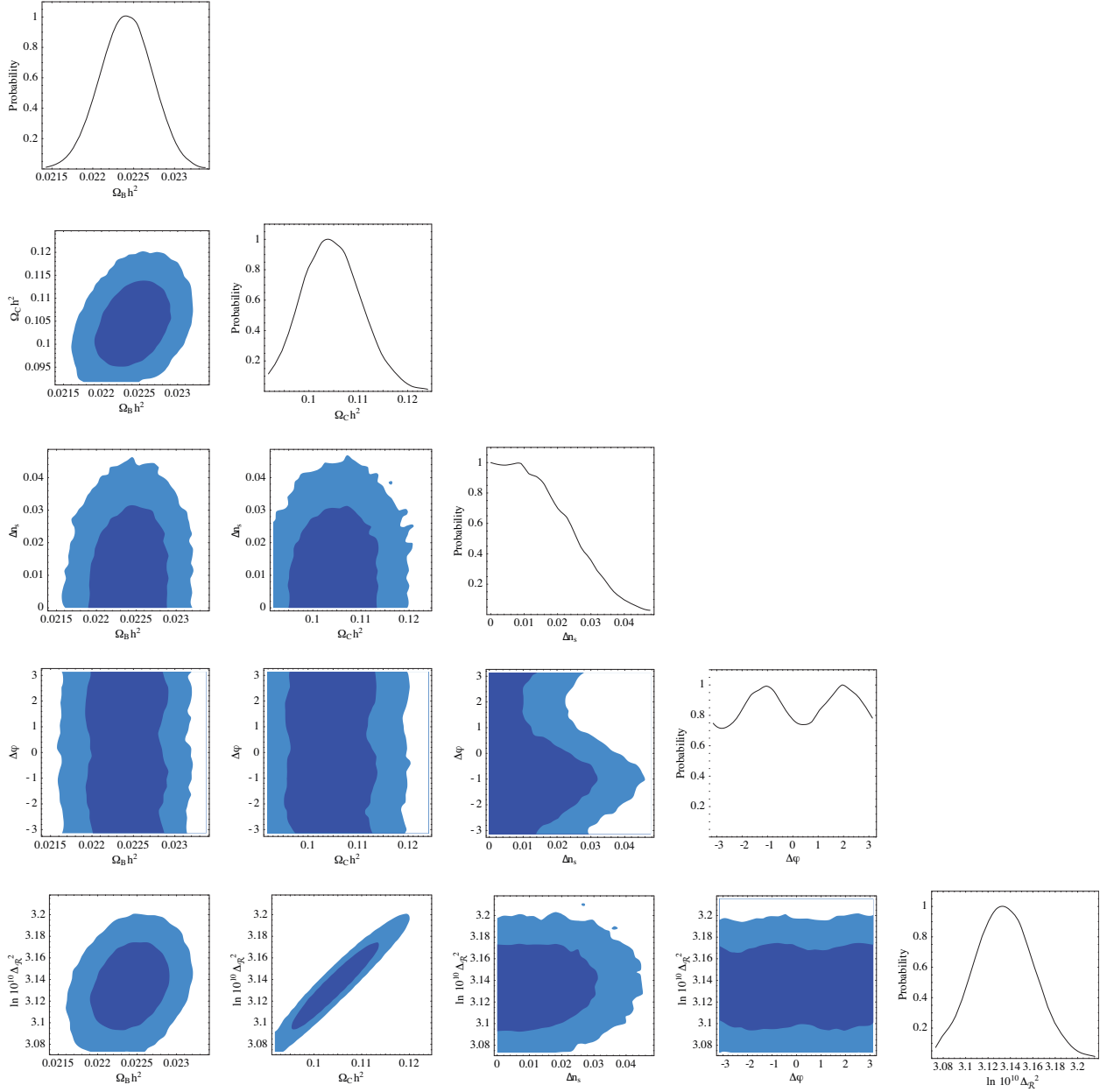

Figure 5: This figure shows a triangle plot for some of the parameters that were sampled in a Markov chain Monte Carlo for an axion decay constant of $f = 10^{-2}$. The contours again represent 68% and 95% confidence levels.

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 5. The plot shows marginalized one-dimensional distributions and two-dimensional 68% and 95% confidence level limits for the

most important ΛCDM parameters as well as δn_s and $\Delta \phi$. In the Monte Carlo, we sampled the parameters δn_s, $\Delta \phi$, all parameters of the ΛCDM except the scalar spectral index, as well as the Sunyaev-Zel'dovich amplitude.

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, μ, Λ are in principle derivable from the data of a string compactification, and as such they obey nontrivial 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, μ, Λ. 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.

5.1 Axions in string theory

Let us first review the origin of the relevant axions. Our conventions and notation are summarized in appendix A. Consider type IIB string theory compactified on an orientifold of a Calabi-Yau threefold X. Let the forms ω^I 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 Σ_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[9]

$$C_2 = \frac{1}{2\pi} c_I(x) \omega^I \,, \tag{5.1}$$

where x is a four-dimensional spacetime coordinate. The ten-dimensional Einstein-frame action [33] that follows is

$$\int d^{10}x \frac{g_s \sqrt{-g_E}}{2(2\pi)^7 \alpha'^4} |dC_2|^2 = \int d^{10}x \frac{g_s \sqrt{-g_E}}{12(2\pi)^9 \alpha'^4} g_E^{\mu\nu} \partial_\mu c_I \partial_\nu c_J \omega_{ij}^I \omega_{i'j'}^J g_E^{ii'} g_E^{jj'} \,. \tag{5.2}$$

Notice that the axions only have derivative couplings, and hence enjoy a continuous shift symmetry at the level of the classical action. In §5.2.3 we will recall the origin of this

---

[9]The factor of 2π is introduced so that the four-dimensional axions c_I have periodicity 2π, 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 ω^I have the dimensions of length-squared, while c_I are dimensionless.

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 α',

$$\alpha' M_p^2 = \frac{\mathcal{V}_E}{\pi}\,, \tag{5.3}$$

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

$$\frac{f^2}{M_p^2} = \frac{g_s}{48\pi^2 \mathcal{V}_E} \left[\frac{\int \omega \wedge *\omega}{(2\pi)^6 \alpha'^3} \right]\,. \tag{5.4}$$

The present definition of the axion decay constant differs by a factor of 2π from that in [5], i.e. $f^{here} = 2\pi f^{there}$. As a consequence our canonically normalized axion has periodicity $2\pi f$, consistent with (2.1).

5.2 Dimensional reduction and moduli stabilization

5.2.1 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 [34]. 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 [35]. We are interested in orientifold actions under which the holomorphic three-form Ω 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,

$$H^{(r,s)} = H_+^{r,s} \oplus H_-^{r,s}\,. \tag{5.5}$$

We therefore divide the basis ω_A, $A = 1, \ldots, h^{1,1}$ into ω_α, $\alpha = 1, \ldots h_+^{1,1}$ and ω_a, $a = 1, \ldots h_-^{1,1}$. 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 ω_a. Grimm and Louis [35] 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:[10]

$$G^a \equiv \frac{1}{2\pi} \left(c^a - i\frac{b^a}{g_s} \right) \tag{5.6}$$

$$T_\alpha \equiv i\rho_\alpha + \frac{1}{2}c_{\alpha\beta\gamma}v^\beta v^\gamma + \frac{g_s}{4}c_{\alpha bc}G^b(G - \bar{G})^c \tag{5.7}$$

---

[10]We use the same notation as [35] with two exceptions: we rescale T_α 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 2π. See appendix A for more details on our conventions.

where ρ_α comes from the RR four-form C_4 integrated over some orientifold-even four-cycle Σ_α, with $\alpha = 1, \ldots, h_+^{1,1}$; 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 Σ_a with $a = 1, \ldots, h_-^{1,1}$. The tree-level Kähler potential is given by[11]

$$K = \log\left(\frac{g_s}{2}\right) - 2\log \mathcal{V}_E \tag{5.8}$$

where the (dimensionless) Einstein-frame volume $\mathcal{V}_E$ of the Calabi-Yau manifold is defined in (A.4). The dependence of this Kähler potential on the multiplets (5.6) and (5.7) 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^α

$$K = \log\left(\frac{g_s}{2}\right) - 2\log\left[\frac{1}{6}c_{\alpha\beta\gamma}v^\alpha(T,G)v^\beta(T,G)v^\gamma(T,G)\right]. \tag{5.9}$$

Then one has to solve (5.7) for v^α 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^α, and in turn of $\tau_\alpha \equiv \operatorname{Re} T_\alpha$ and $\operatorname{Im} G$, but does not depend on $\operatorname{Re} G$ and $\operatorname{Im} T_\alpha$ (as can be seen by taking the real part of (5.7)). One might be tempted to conclude that c enjoys a shift symmetry but that b does not, but, as we will explain in §5.2.3, both fields have shift symmetries.

The tree-level superpotential W_0 does not depend on the multiplets (5.6) and (5.7). 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.

5.2.2 Nonperturbative stabilization of the Kähler moduli

Let us now proceed to consider nonperturbative effects. We follow the KKLT strategy [34] for the construction of a de Sitter vacuum. We assume that each four-cycle T_α 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.[12] This results in the following four-dimensional superpotential:

$$W = W_0 + \sum_{\alpha=1}^{h_+^{1,1}} A_\alpha e^{-a_\alpha T_\alpha}, \tag{5.10}$$

where A_α 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_α the number of D7-branes in

---

[11]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.

[12]In general, Euclidean D3-branes or D7-branes will wrap some linear combinations $\tilde{T}_\alpha$ of the cycles appearing in (5.7), rather than the basis cycles T_α themselves, but for simplicity we will suppress this issue.

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

$$
\begin{aligned}
0 &= D_\alpha W \equiv \partial_{T_\alpha} W + W \partial_{T_\alpha} K = -A_\alpha a_\alpha e^{-a_\alpha T_\alpha} - 2W \frac{\partial_{T_\alpha} \mathcal{V}_E}{\mathcal{V}_E} \\
&= -A_\alpha a_\alpha e^{-a_\alpha T_\alpha} - W \frac{v^\alpha}{2\mathcal{V}_E} \,,
\end{aligned}
\tag{5.11}
$$

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

$$
0 = D_a W \equiv \partial_{G^a} W + W \partial_{G^a} K = -iW \frac{c_{\alpha ac} v^\alpha b^c}{4\pi \mathcal{V}_E}
\tag{5.12}
$$

where in both cases in the last step we used the chain rule and the definitions of G^a and T_α in terms of two-cycle volumes v^α. The condition (5.11) is simplified if we first solve for $\operatorname{Im} T_\alpha$, which gives

$$
a_\alpha \operatorname{Im} T_\alpha = \theta_{A_\alpha} - \theta_{W_0} + k_\alpha \pi \,, \qquad k_\alpha \in \mathbb{Z} \,.
\tag{5.13}
$$

Then we are left with the set of real equations for each α,

$$
(\pm 1)_\alpha |A_\alpha| a_\alpha e^{-a_\alpha \tau_\alpha} = \partial_{T_\alpha} K \left(|W_0| + \sum_\beta (\pm 1)_\alpha |A_\beta| e^{-a_\beta \tau_\beta} \right) \,,
\tag{5.14}
$$

where (± 1) depends on the value of k in (5.13). As long as the orientifold-even four-cycle Kähler moduli are defined as in (5.7), then $\partial_{T_\alpha} K = -v^\alpha/(2\mathcal{V}_E) < 0$ for every α. Now we prove that in (5.14) the minus sign has to be chosen for every α in order to have a supersymmetric solution. First we notice that the sign of the right hand side does not depend on α, so k_α and hence $(\pm 1)_\alpha$ have to be the same for every α. If we choose the positive sign in (5.14), 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 $\operatorname{Im} T_\alpha$ boils down to taking all A_α real and negative and W_0 real and positive or the other way around.[13]

Concerning (5.12), an obvious solution is given by $b^a = 0$ for every a. As argued in [5], 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.

5.2.3 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

---

[13]We notice that if one chooses as Kähler variable a linear combination of the T_α defined in (5.7), 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 α. In this case, the minimization of $\operatorname{Im} T_\alpha$ boils down to taking W_0 real and positive and A_α real with $\operatorname{Sign}(A_\alpha) = \operatorname{Sign}(\partial_{T_\alpha} K)$, up to multiplying W by an overall phase.

result [36, 37] establishing the shift symmetry to 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 Σ is a two-cycle in the Calabi-Yau manifold. The vertex operator representing the coupling of b to the string worldsheet is [36]

$$V(k) = \frac{1}{2\pi\alpha'} \int_\Sigma d^2\xi \exp\Big(ik \cdot X(\xi)\Big)\epsilon^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu B_{\mu\nu}(X)\,. \qquad (5.15)$$

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 Σ_a creates a monodromy for the axion c_a, spoiling its shift symmetry and inducing an asymptotically linear potential [5]. 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

$$V(c_a) = \frac{\epsilon}{g_s(2\pi)^5\alpha'^2} \sqrt{\ell^4 + (2\pi g_s c_a)^2}\,, \qquad (5.16)$$

where $\ell\sqrt{\alpha'}$ is the size of Σ_a and ϵ 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 ϕ 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 ϕ.

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 (5.10) with $h^{1,1}_+ = 1$. The Kähler potential is then [35]

$$K = -3M_p^2 \log \left(T + \bar{T} - d\, b^2 \right), \qquad (5.17)$$

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 [5] choosing b as an axion: for a generic uplifting,[14] $V''(b) \sim V(b)$, 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 [39] and can be intuitively understood in the same way. Here we will take $b = 0$ as the stabilized value [15] 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 $\int F_2 \wedge C_2$. As observed in [5], 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 [5], is to stabilize the Kähler moduli via gaugino condensation on D7-branes, which leads to an exponentially smaller (and hence negligible) mass for c.

---

[14]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 [38].

[15]It is easy to check that $b = 0$ is still the stabilized value after the inclusion of nonperturbative corrections to the Kähler potential, *cf.* §6.5.

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 §5.1, the decay constant of an axion $C_2 = c(x)\omega/(2\pi)$ is given by

$$\frac{f^2}{M_p^2} = \frac{g_s}{48\pi^2 \mathcal{V}_E} \left[\frac{\int \omega \wedge *\omega}{(2\pi)^6 \alpha'^3} \right], \tag{5.18}$$

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 [40].) We will first recall, in §5.3.1, 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 §5.3.2, we will present a concrete example that illustrates the geometry of a configuration that leads to small f.

5.3.1 Decay constants in terms of $\mathcal{N} = 1$ data

In §5.2.1 we have reviewed, following [35], 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_α. The tree-level Kähler potential given in (5.9) 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_α 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 [35]. We are interested in one particular mode from among the G^a, which we will denote by G^-; Σ_- 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 1×1 block $K_{G^- \bar{G}^-}$. The kinetic term for c^- is then given by

$$-\frac{1}{2}f^2 \left(\partial c^-\right)^2 = M_p^2 K_{G^- \bar{G}^-} \frac{1}{(2\pi)^2} \left(\partial c^-\right)^2 \subset M_p^2 K_{G^- \bar{G}^-} \left|\partial G^-\right|^2, \tag{5.19}$$

where

$$K_{G^- \bar{G}^-} = \frac{\partial^2 K(G, T)}{\partial G^- \partial \bar{G}^-} = -g_s \frac{c_{\alpha--} v^\alpha}{4\mathcal{V}_E}, \tag{5.20}$$

and we used

$$c_{\alpha\beta\gamma} v^\beta v^\gamma = 2\tau_\alpha + g_s c_{\alpha bc} \operatorname{Im} G^b \operatorname{Im} G^c. \tag{5.21}$$

31

Hence we can express the decay constant of the axion c^- as

$$\frac{f^2}{M_p^2} = \frac{g_s}{8\pi^2} \frac{c_{\alpha--}v^\alpha}{\mathcal{V}_E} .$$ (5.22)

As promised, we have expressed the norm $\int \omega \wedge *\omega$ in terms of the intersection numbers

$$\frac{\int \omega \wedge *\omega}{(2\pi)^6 \alpha'^3} = \frac{2}{3} c_{\alpha--} v^\alpha .$$ (5.23)

In §6.4 we will discuss the constraints that follow from the result (5.22). First, in the following subsection we provide some geometrical intuition for (5.22).

5.3.2   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 Σ of complex dimension one. Let ω be the two-form dual to the blowup cycle of the orbifold, and let Σ be the two-manifold of fixed points, *i.e.* the base over which the Eguchi-Hanson space is fibered.[16] 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 Σ 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 [41], 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 ω corresponding to this cycle may be written

$$\omega = 2\pi\alpha' \frac{a^2}{r^2} \left(\frac{dr}{r} \wedge d\psi + \cos\theta \frac{dr}{r} \wedge d\phi + \frac{1}{2}\sin\theta \, d\theta \wedge d\phi \right)$$ (5.24)

in terms of r and the angular coordinates ψ, θ, ϕ. By observing that $*_4\omega = -\omega$ and that $\int \omega \wedge \omega = -(2\pi)^4 \alpha'^2/2$, one finds

$$\frac{\int_{EH} \omega \wedge *_4\omega}{(2\pi)^4 \alpha'^2} = \frac{1}{2} .$$ (5.25)

Clearly, given the form of ω, this integral has its support in a region $a \leq r \lesssim few \times a$. This justifies the local approximation as long as the compact space has a radius that is large

---

[16]Concretely, we are imagining that Σ 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 [5] for further details, and for an example of a suitable orbifold action in a Klebanov-Strassler throat.

compared to a. Next, we observe that

$$\frac{\int \omega \wedge *_6 \omega}{(2\pi)^6 \alpha'^3} = \frac{\int_{EH} d^4 x \, \omega \wedge *_4 \omega}{(2\pi)^4 \alpha'^2} \times \frac{\int_\Sigma \sqrt{g}}{(2\pi)^2 \alpha'} = \frac{1}{2} \text{Vol}(\Sigma) \tag{5.26}$$

Substituting this in (5.18), we recover the parametric scaling of §5.3.1.

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 §6.1 by considering the constraints from computability, then give, in §6.2, a qualitative description of the constraints from backreaction, deferring details to appendix B. Next, in §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 §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 §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.

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 α' 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 α' expansion. To do this including a reasonable *estimate* of numerical factors such as 2π, it is convenient to use worldsheet instantons as a proxy for

perturbative α' 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 worldsheet instanton action obeys $e^{-S_{WS}} \lesssim e^{-2}$, or

$$\frac{1}{2\pi\alpha'} \int \sqrt{g_{string}} \gtrsim 2 \,, \tag{6.1}$$

which using $g_{string} = g_{Einstein}\sqrt{g_s}$ is converted to Einstein frame

$$2 < \frac{\sqrt{g_s}}{2\pi\alpha'} \int_{\Sigma^\alpha} J = \sqrt{g_s} v^\alpha 2\pi \quad \Rightarrow \quad v^\alpha > \frac{1}{\pi\sqrt{g_s}} \,, \tag{6.2}$$

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

$$e^{-a_\alpha T_\alpha} < e^{-2} \ll 1 \quad \Rightarrow \quad \tau_\alpha > \frac{N_\alpha}{\pi} \,. \tag{6.3}$$

Additional constraints come from the moduli stabilization process. To use the single-field inflationary analysis we have developed in §2 and §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

$$V(\phi, \mathcal{V}_E) \approx \left(\frac{\langle \mathcal{V}_E \rangle}{\mathcal{V}_E}\right)^2 \mu^3 \phi \tag{6.4}$$

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

$$V_{AdS} = -\frac{g_s}{2} \frac{3|W|^2}{\mathcal{V}_E^2} \,. \tag{6.5}$$

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

$$V(\phi_{CMB}) = \epsilon \, (\, 0.027 M_p)^4 \simeq 2.4 \cdot 10^{-9} M_p^4 \,. \tag{6.6}$$

Hence we obtain the constraint

$$\frac{g_s}{2}\frac{3|W|^2}{\mathcal{V}_E^2} = |V_{AdS}| \simeq \mathcal{U}_{mod} \gg 2.4 \cdot 10^{-9} M_p^4 \,. \tag{6.7}$$

To extract a useful form of the above constraints, let us substitute for W the solution of any of the equations (5.11)

$$W = +|A_\alpha| a_\alpha \, e^{-a_\alpha \tau_\alpha} \frac{2\mathcal{V}_E}{v^\alpha} \,, \tag{6.8}$$

with no sum over α. We will also assume $|A_\alpha| \sim 1$ (see [5] for a discussion of this point). After some manipulations we find

$$\tau_\alpha \ll -\frac{N_\alpha}{2\pi} \log\left(N_\alpha 10^{-5} \frac{v^\alpha}{\pi\sqrt{g_s}}\right) \,, \tag{6.9}$$

again with no summation over α. Finally, we should limit the number of D7-branes in each stack; although there plausibly exist examples with N_α quite large, we will impose $N_\alpha \le 50$. This gives us

$$\tau_\alpha \ll 73 - 8\log\left(\frac{v^\alpha \pi\sqrt{g_s}}{2g_s}\right) \,. \tag{6.10}$$

We notice that $v^\alpha(\pi\sqrt{g_s}) > 1$ was the condition in (6.2) that enabled us to neglect α' corrections, so that as long as $g_s \le 0.5$ the second term on the right hand side of (6.10) is negative.

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 Σ_- of the form

$$\frac{1}{(2\pi)^2\alpha'}\int_{\Sigma_-} C_2 \equiv N_w = \frac{\phi}{2\pi f} \gg 1 \tag{6.11}$$

In the absence of an NS5-brane wrapping Σ, 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 Σ'_- homologous to Σ_-, 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 Σ_- to Σ'_-.

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 [42, 43].

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 [44] 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.[17]

In appendix B we give more details about the above setup. We show, through two explicit models of increasing complexity, the robustness of the above suppression mechanism.

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.

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 α'^3 corrections to the effective action due to Gross and Sloan [45] (at the four-point level) and Kehagias and Partouche [46] (up to the eight-point level). These corrections are of the same lineage as the

---

[17]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.

famous Riemann[4] term, but involve NS-NS three-form flux. This yields corrections to the axion kinetic terms. Following [46], the ten-dimensional Einstein-frame action including the leading (α'^3) corrections is

$$S_{10D,E} = \frac{1}{(2\pi)^7 \alpha'^4} \int d^{10}x \sqrt{g_E} \left(R_E - \frac{1}{12g_s} H_{KMN} H^{KMN} + \frac{\zeta(3)}{3 \times 2^6} g_s^{-3/2} \alpha'^3 \bar{R}^4 + \ldots \right) . \quad (6.12)$$

where

$$\bar{R}_{MN}{}^{PQ} = R_{MN}{}^{PQ} + \frac{1}{2} g_s^{-1/2} \nabla_{[M} H_{N]}{}^{PQ} - \frac{1}{4} g_s^{-1} H_{[M}{}^{C[P} H_{N]C}{}^{Q]} + \ldots \quad (6.13)$$

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 α'^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 [45], 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

$$S_{4D} = \int d^4x \left[-M_p^2 \frac{g_s}{2} \dot{c}^2 g^{88} g^{99} + \frac{\zeta(3)}{2^6 g_s^{3/2}} \frac{\mathcal{V}_E^3}{\pi^3 M_p^4} \left(\frac{1}{2} \dot{c}^8 g_s^4 (g^{88} g^{99})^4 + \frac{1}{2^4} \ddot{c}^4 g_s^2 (g^{88} g^{99})^2 \right) \right] \quad (6.14)$$

Now we use $\phi = cf$ to make the kinetic term canonical, yielding the action in terms of ϕ,

$$\begin{aligned} S_{4D} &= \int d^4x \left[-\frac{1}{2} \dot{\phi}^2 + \frac{\zeta(3)}{2^6 g_s^{3/2}} \frac{\mathcal{V}_E^3}{\pi^3} \left(\frac{1}{2} \frac{\dot{\phi}^8}{M_p^{12}} + \frac{1}{2^4} \frac{\ddot{\phi}^4}{M_p^8} \right) \right] \\ &\equiv \int d^4x \left[-\frac{1}{2} \dot{\phi}^2 + \frac{\dot{\phi}^8}{M_I^{12}} + \frac{\ddot{\phi}^4}{M_{II}^8} \right] , \end{aligned} \quad (6.15)$$

where we can now calculate the scale of the higher derivative terms M_I and M_{II} to be

$$M_I = M_p \frac{g_s^{1/8}}{\mathcal{V}_E^{1/4}} \left(\frac{\pi^3 2^7}{\zeta(3)} \right)^{1/12} \quad (6.16)$$

and

$$M_{II} = M_p \frac{g_s^{3/16}}{\mathcal{V}_E^{3/8}} \left(\frac{\pi^3 2^{10}}{\zeta(3)} \right)^{1/8} \quad (6.17)$$

To determine whether these higher-derivative terms will become important, we compute the dimensionless quantity $\frac{\omega}{M_{I,II}}$, where $\omega = \frac{\dot{\phi}}{f}$ (3.5) is the physical frequency of oscillations; we obtain

$$\frac{\omega}{M_I} \simeq 5 \cdot 10^{-3} \left(\frac{f}{10^{-3}}\right)^{-1} \left(\frac{g_s}{0.2}\right)^{-1/8} \left(\frac{\mathcal{V}_E}{120}\right)^{1/4}, \tag{6.18}$$

$$\frac{\omega}{M_{II}} \simeq 6 \cdot 10^{-3} \left(\frac{f}{10^{-3}}\right)^{-1} \left(\frac{g_s}{0.2}\right)^{-3/16} \left(\frac{\mathcal{V}_E}{120}\right)^{3/8}. \tag{6.19}$$

For the ranges of f and $\mathcal{V}_E$ that will be of interest to us (*cf.* §7), the higher-derivative terms are not important and our two-derivative approximation is justified.

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 $f < M_p$ [47], and we then describe and incorporate a novel lower bound, specific to our setup, that arises from combining the requirements that α' 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 [47] 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 [5] 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

$$N_w \ll \frac{R_\perp^4 X}{4\pi g_s} \quad \Rightarrow \quad \frac{f}{M_p} \gg \frac{2\phi g_s}{R_\perp^4 X}, \tag{6.20}$$

where we have defined $X \equiv \mathrm{Vol}(X_5)/\pi^3$, 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[18] $\mathcal{V}$ of the Calabi-Yau has to be

---

[18]We always refer to the warped volume, calculated with the whole warped metric.

larger than the volume of any throat it includes, one finds that

$$\mathcal{V} > V_{throat} = \frac{\pi^3}{2} X R^6 \,, \tag{6.21}$$

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 (6.20) and (6.21), we find

$$\frac{f}{M_p} > \frac{\pi^2 2^{1/3} \phi g_s}{X^{1/3} \mathcal{V}^{2/3}} \simeq \frac{137 g_s}{X^{1/3} \mathcal{V}^{2/3}} = \frac{0.09}{X^{1/3} \mathcal{V}_E^{2/3}} \,. \tag{6.22}$$

Although the above constraint substantially restricts our parameter space, an even stronger constraint comes from demanding the validity of α' perturbation theory: using (5.22) for f and combining this with the lower bound on two-cycle volumes given in (6.2), we obtain

$$\frac{f^2}{M_p^2} = \frac{\sqrt{g_s}}{(2\pi)^3 \mathcal{V}_E} \left(c_{\alpha--} v^\alpha \sqrt{g_s} \pi \right) > \frac{\sqrt{g_s}}{(2\pi)^3 \mathcal{V}_E} \,, \tag{6.23}$$

where we have assumed that $c_{\alpha--} \geq 1$. (6.23) 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 (6.2), we find

$$\frac{f}{M_p} < g_s \frac{\sqrt{3}}{2} \,. \tag{6.24}$$

6.5   Constraints on the amplitude of the modulations

So far we have seen that with the Kähler potential and superpotential given in (5.9) and (5.10), the axion c persists as a flat direction after moduli stabilization.[19] As explained in [5], 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 (5.16). 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.

---

[19]As we have remarked, the axion b has its flat direction lifted by nonperturbative stabilization of the Kähler moduli.

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.[20] 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-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 [49] 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,

$$K = -2 \log \left[\mathcal{V}_E + g(G, \bar{G}) \right] , \qquad (6.25)$$

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 Γ.

At the time of writing, the nonperturbative correction g is not known explicitly, but a modular-invariant result has been conjectured in [50]. 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 $c \to c + 2\pi$; 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 Γ 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 [50]. Moreover, notice that along the orbits of Γ, 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 Γ trivial, is

$$K = -2 \log \left[\mathcal{V}_E + e^{-S_{ED1}} \cos(c) \right] = -2 \log \left[\mathcal{V}_E + e^{-\frac{2\pi v^+}{\sqrt{g_s}}} \cos(c) \right] . \qquad (6.26)$$

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.

---

[20]To see this, note that (*cf.* [48]) 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.

We begin by noticing the following implication

$$
\begin{cases}
D_{T_\alpha} W = \mathcal{O}\left(e^{-2S_{ED1}}\right) \\
D_{G^a} W = \mathcal{O}\left(e^{-S_{ED1}}\right) \\
W_{G^a} = 0
\end{cases}
\Rightarrow
\begin{cases}
\partial_{T_\alpha} V = 0 + \mathcal{O}\left(e^{-2S_{ED1}}\right) \\
\partial_{G^a} V = -2\, e^K \, |W|^2 K_{G^a} + \mathcal{O}\left(e^{-2S_{ED1}}\right)
\end{cases}
\tag{6.27}
$$

which can be verified by direct computation. This allows us to use the F-flatness condition to find the minimum in the T_α-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 §5.

First, the phases of the T_α are stabilized as in (5.13), with k being odd as explained below (5.14). The reason is that the sign of $\partial_{T_\alpha} K$ is not changed by the small nonperturbative correction $e^{-S_{ED1}}$. Second, $\mathrm{Im}\, G$ is again stabilized at 0. Given (6.27), the equation one needs to solve is $D_{G^a} W = 0$, which reduces to

$$
0 = W \partial_{G^a} K \propto \partial_{G^a} \mathcal{V}_E - e^{-S_{ED1}} \left[\pi \sin(c) + \cos(c)\frac{2\pi}{\sqrt{g_s}} \partial_{G^a} v^+ \right] = 0 \,,
\tag{6.28}
$$

where we made use of (5.6) to perform the derivative on c. Since $\mathcal{V}_E$ and v^+ only depend on $\mathrm{Im}\, G$ implicitly as in (5.21), we can take the imaginary part of (6.28),

$$
\frac{1}{2}(\partial_{\mathrm{Im}\, G^-} v^\alpha) c_{\alpha\beta\gamma} v^\beta v^\gamma - e^{-S_{ED1}} \cos(c)\frac{2\pi}{\sqrt{g_s}} \partial_{\mathrm{Im}\, G^-} v^+ = 0 \,.
\tag{6.29}
$$

But from (5.21) we know that

$$
c_{\alpha\beta\gamma}(\partial_{\mathrm{Im}\, G^-} v^\beta) v^\gamma = g_s c_{\alpha a -} \mathrm{Im}\, G^a \,,
\tag{6.30}
$$

which means that $\mathrm{Im}\, G^a = 0$ (for every a) is a solution to (6.29). The real part of (6.28) is nonvanishing and of order $e^{-S_{ED1}}$. Again because of (6.27), the minimization in the τ_α is obtained by imposing $D_{T_\alpha} W = 0$. These equations depend on the inflaton c, appearing explicitly in (6.26), and hence the minimum in the T_α directions will be a function of c. Integrating out the T_α 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 τ_α directions in a perturbative expansion in $e^{-S_{ED1}}$ by

$$
\tau_{\alpha,min} \equiv \tau_{\alpha,(0)} + \cos(c)\, e^{-S} \tau_{\alpha,(1)} + \dots \,,
\tag{6.31}
$$

and so on for all other variables. The zeroth-order equations are

$$
(D_{T_\alpha} W)_{(0)} = (\partial_{T_\alpha} W)_{(0)} + W_{(0)}(\partial_{T_\alpha} K)_{(0)} = 0 \,,
\tag{6.32}
$$

42

which can be solved numerically once the model is specified. The first-order equations are

$$(D_{T_\alpha}W)_{(1)} = (\partial_{T_\alpha}W)_{(1)} + W_{(1)}(\partial_{T_\alpha}K)_{(0)} + W_{(0)}(\partial_{T_\alpha}K)_{(1)} = 0\,, \qquad (6.33)$$

which again can be solved numerically using the solutions of (6.32). We turn now to estimate the parameter b defined in (2.1). One finds

$$bf \;\equiv\; V_{(1)}\mu^{-3}e^{-S_{ED1}} = \frac{\mathcal{U}_{mod}\,\phi}{\mu^3\phi}e^{-S_{ED1}}\left(K_{(1)} + 2\mathrm{Re}\,\frac{W_{(1)}}{W_{(0)}}\right)\,, \qquad (6.34)$$

where we have defined $\mathcal{U}_{mod}$ as the moduli stabilization barrier at zeroth order in $e^{-S_{ED1}}$, *i.e.*

$$\mathcal{U}_{mod} = \frac{g_s}{2}\left(\frac{3|W|^2}{\mathcal{V}_E^2}\right)_{(0)}\,. \qquad (6.35)$$

More explicitly, using (6.26) and (5.10),

$$
\begin{aligned}
bf &= \frac{\mathcal{U}_{mod}\,\phi}{2.4\cdot 10^{-9}M_p^4}e^{-S_{ED1}}\left[\frac{8\pi}{\sqrt{g_s}}\frac{(\partial_{T_\alpha}v^+)_{(0)}}{v_{(0)}^\alpha} - 2a_\alpha\tau_{\alpha,(1)} - \frac{2v_{(1)}^\alpha}{v_{(0)}^\alpha}\right] \\
&= \frac{\mathcal{U}_{mod}\,\phi}{2.4\cdot 10^{-9}M_p^4}2e^{-S_{ED1}}\left[\frac{\sum_\beta(\partial_{T_\beta}W)_{(0)}\tau_{\beta,(1)}}{W_{(0)}} - \frac{\mathcal{V}_{E,(1)}+1}{\mathcal{V}_{E,(0)}}\right]\,,
\end{aligned}
\qquad (6.36)
$$

where the first line is valid for any α and the second line (obtained using (6.33)) shows that the expression for b is independent of α. Notice that $\partial_{T_\alpha}v^+ = \frac{1}{2}\partial_{\tau_\alpha}v^+$ is given implicitly by

$$c_{\alpha\beta\gamma}(\partial_{\tau_\rho}v^\gamma)v^\beta = \partial_{\tau_\rho}\tau_\alpha = \delta_\alpha^{\ \rho}\,. \qquad (6.37)$$

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 (6.36) 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 [51]. 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 (6.36) can be bounded by (6.2). Finally, we denote the model-dependent term in square brackets in (6.36) by c_0. Putting things together leads to the bound

$$bf < 2c_0\cdot 10^7\frac{g_s}{\mathcal{V}_E^2}e^{-2/g_s}\left(\frac{W}{0.1}\right)^2\,. \qquad (6.38)$$

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, 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.

7 Combined Theoretical and Observational Constraints

We now summarize our results, combining the observational constraints from §4 with the theoretical constraints from §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 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 (2.1), 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 §3.4, we present, in figure 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 §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 (6.38)[21]

$$bf < 2c_0 \cdot 10^9 M_p^4 \frac{g_s}{\mathcal{V}_E^2} e^{-2/g_s} \,, \tag{7.1}$$

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}.

---

[21]We stress that this 'bound' is *not* universal and depends on the assumptions enumerated in §6. We include it here as a representative example of the constraints that arise in particular scenarios.
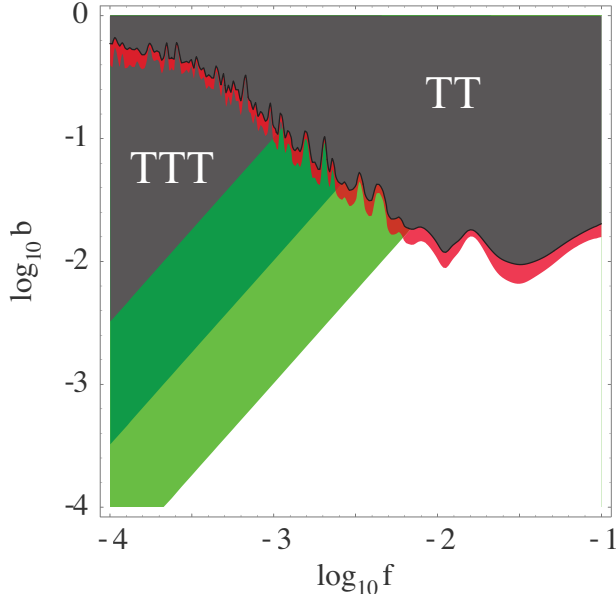
Figure 6: We show the (one- and two-sigma) likelihood contours for the temperature two-point function together with three contours that characterize the amplitude of the three point function, for $f_{res} = 200, 20, 2$.

Next, we obtained a lower bound for f in (6.23).[22] A precise upper bound, however, is highly model-dependent. We estimate an upper limit by assuming that the intersection numbers are of order one[23] and that no precise cancellations occur. From (6.23) and (6.24), the complement of the theoretically excluded range for f is then

$$\frac{g_s^{1/4}}{(2\pi)^{3/2}\sqrt{\mathcal{V}_E}} < f < g_s \frac{\sqrt{3}}{2} \ . \tag{7.2}$$

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 (7.2)). A way to quantify this is to use the estimate obtained in §3.4,

$$f_{res} \simeq \frac{9}{4} \frac{b}{(\phi f)^{3/2}} \ , \tag{7.3}$$

---

[22]The constraint from the backreaction described in §6.2 is weaker than (6.23).

[23]Larger intersection numbers are an interesting possibility that we will not investigate here.
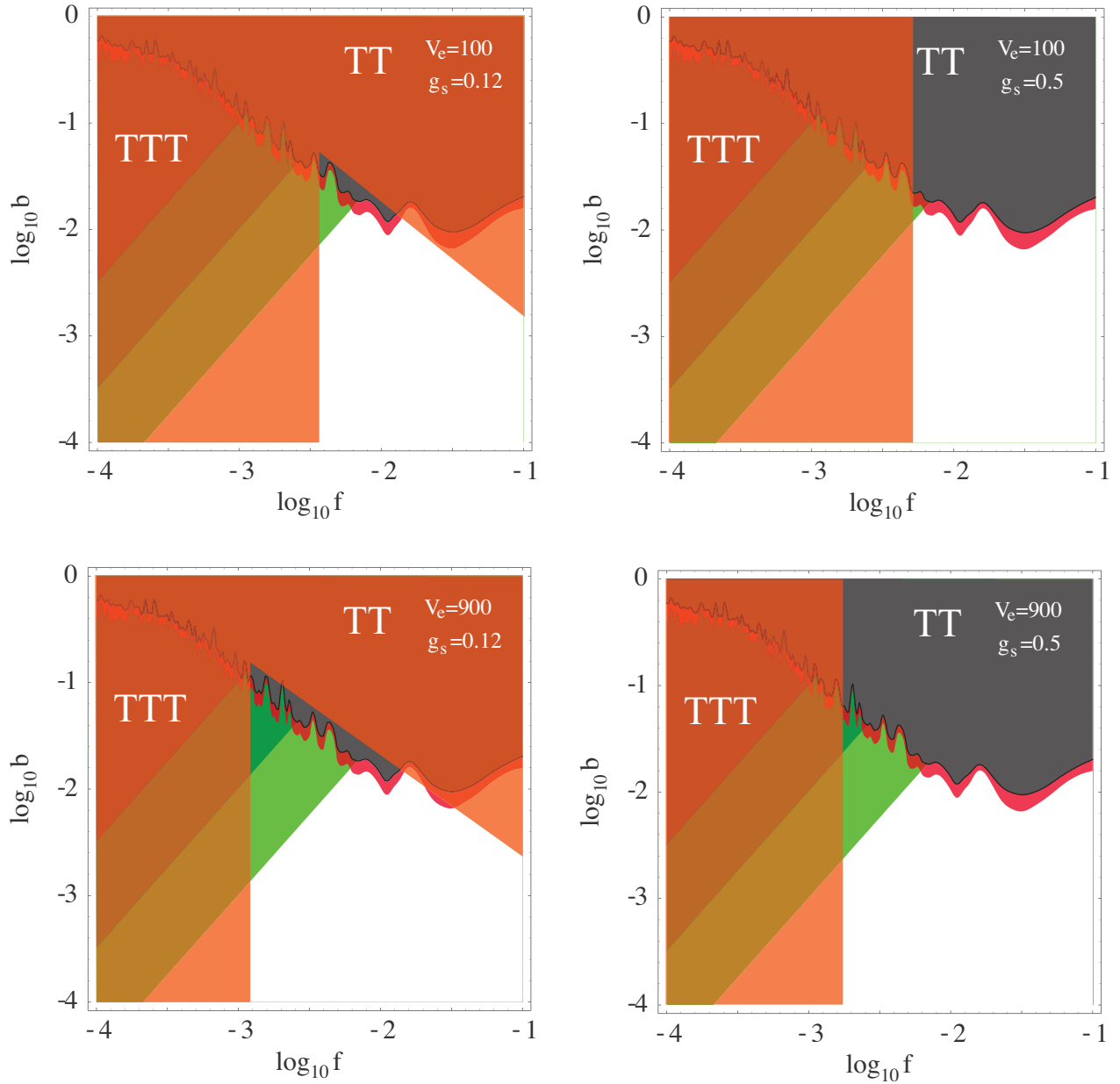
45

Figure 7: We superimpose the theoretical constraints, summarized in (7.1) and (7.2), on the constraints imposed by observations, which are shown in figure 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 $\mathcal{V}_E = 100, 900$.
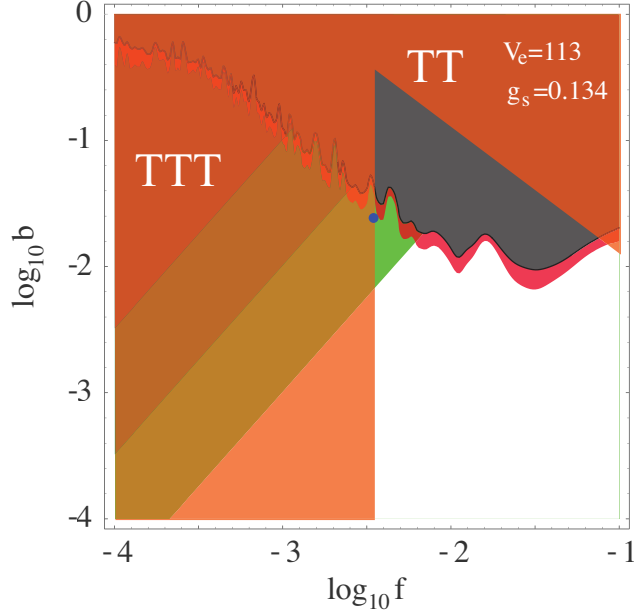
Figure 8: The blue dot represents the explicit numerical example presented in appendix D. It represents a case in which upcoming experiments could detect the signatures of modulations in both the two-point function and the three-point function.

and the lower bound in (7.2). The result is

$$\mathcal{V}_E > 170 \left(\frac{g_s}{0.2}\right)^{1/2} \left(\frac{f_{res}}{10}\right)^{4/5} \left(\frac{10^{-4}}{bf}\right)^{4/5} . \tag{7.4}$$

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 §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 (7.1).

Finally, in figure 8 we show where a particular numerical toy example, with specific choices of the intersection numbers, lies in the $\{\log(f), \log(b)\}$ plane.

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 inflaton 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 nontrivial 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.* [22, 23, 8, 9, 10]. In light of our calculation of the power spectrum in terms of particle production (§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.

Acknowledgments

A   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[24] [33]

$$S_{10} = \frac{1}{(2\pi)^7 \alpha'^4} \int d^{10}x \sqrt{g_{string}} \left(e^{-2\Phi} R_{string} - \frac{1}{2} |dC_2|^2 \right) \tag{A.1}$$

---

[24]Remember that $2\kappa_{10}^2 = (2\pi)^7 \alpha'^4$.

which after the rescaling to the ten-dimensional Einstein-frame metric $e^{-\Phi/2}g_{string,MN} = g_{E,MN}$ becomes

$$S_{10} = \frac{1}{(2\pi)^7 \alpha'^4} \int d^{10}x \sqrt{g_E} \left(R_E - \frac{1}{2}g_s|dC_2|^2 \right),$$ (A.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

$$M_p^2 = \frac{\int_Y \sqrt{g_E}}{(2\pi)^6\alpha'^3}\frac{1}{\alpha'\pi} \equiv \frac{\mathcal{V}_E}{\alpha'\pi},$$ (A.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

$$\mathcal{V}_E = \frac{1}{6}\frac{\int J \wedge J \wedge J}{(2\pi)^6\alpha'^3} = \frac{1}{6}v^I v^J v^K \frac{\int \omega_I \wedge \omega_J \wedge \omega_K}{(2\pi)^6\alpha'^3} \equiv \frac{1}{6}v^I v^J v^K c_{IJK},$$ (A.4)

where ω_I for $I = 1, \ldots, h^{1,1}$ are a basis of the cohomology $H^2(Y, \mathbb{Z})$ normalized such that

$$\int_{\Sigma_I} \omega_J = (2\pi)^2\alpha'\delta_J{}^I$$ (A.5)

for a basis Σ_I of the dual homology $H_2(Y, \mathbb{Z})$. With the ansatz for the ten-dimensional RR two-form

$$C_2 = \frac{1}{2\pi}c(x)\omega,$$ (A.6)

for some base two-cycle ω, we get a four-dimensional axion $c(x)$ with periodicity[25] 2π, 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

$$\frac{f^2}{M_p^2} = \frac{g_s}{12\mathcal{V}_E(2\pi)^2}\left[\frac{\int \omega \wedge *\omega}{(2\pi)^6\alpha'^3}\right].$$ (A.7)

The four-dimensional $\mathcal{N} = 1$ Kähler potential for the Kähler moduli is

$$K = -2\log\mathcal{V}_E.$$ (A.8)

B   Induced Shift of the Four-Cycle Volume

In this appendix we address the issue raised in §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.

---

[25]Note that this choice differs from that in [5], where the axion periodicity was $(2\pi)^2$.

More specifically, if an NS5-brane wraps some cycle Σ, then a nonvanishing integral $\int_\Sigma C_2 \neq 0$ leads to the presence of energy that is localized near Σ 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 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.

The 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 Σ, with

$$\frac{1}{(2\pi)^2\alpha'} \int_\Sigma C_2 \equiv N_w \,, \tag{B.1}$$

as well as a distant NS5-brane wrapping a homologous cycle Σ', 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.[26]

_____

[26]More generally, one could consider different degrees of warping for each fivebrane; extending our consid-

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 [44], we recognize that the leading backreaction effect comes from the total D3-brane charge on each stack, while the brane-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

$$ds_6^2 = \sqrt{H^{-1}(y)}ds_4^2 + \sqrt{H(y)}ds_6^2 \tag{B.2}$$

$$\tilde{F}_5 = (1 + *)dH^{-1} \wedge \mathrm{Vol}_4\,, \quad \Phi = \mathrm{const}\,. \tag{B.3}$$

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.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.[27]

Then, in §B.2, §B.2.1 and §B.3, we describe the case of a resolved conifold using the solution of [52, 53]. We consider a particular holomorphic embedding of a four-cycle and compute numerically the inflaton-dependence of its warped volume.

B.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 B stack is located at some

---

erations to this case is straightforward.

[27]Although suppression from symmetry of the embedding appears unappealing because of the fine-tuning required, one should keep in mind that it could conceivably be enforced by a discrete symmetry of the compactification.
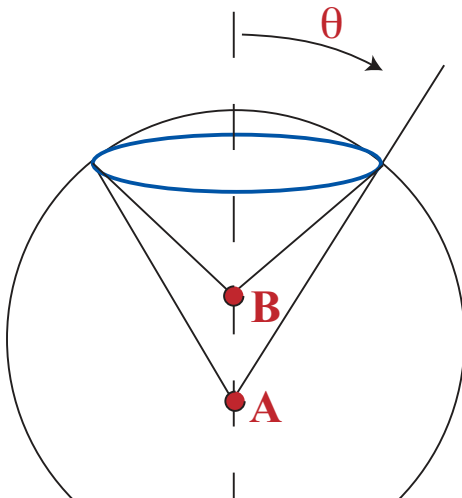
Figure 9: This diagram illustrates the positions of the A and B stacks of D3-branes in $\mathbb{R}^6$, the choice of the angular coordinate θ, and, in blue, a (topologically trivial) four-cycle.

position $(u, 0, 0, 0, 0, 0)$ for[28] $u \in \mathbb{R}^+$, where we have chosen spherical coordinates (see figure B.1) with the metric

$$ds_6^2 = dr^2 + r^2 \left(d\theta^2 + \sin^2\theta d\Omega_4^2 \right) , \tag{B.4}$$

where $d\Omega_4$ is the volume form of S^4. With the usual D3-brane ansatz (B.2) one finds the solution

$$H = 1 + \frac{R_A^4}{r^4} + \frac{R_B^4}{(r^2 + u^2 - 2ru\cos\theta)^2} , \tag{B.5}$$

$$R_{A,B} = 4\pi g_s \alpha'^2 \left(\frac{N}{2} \pm N_w \right) . \tag{B.6}$$

Let us consider a (topologically trivial) four-cycle Σ_4 defined by[29] $r = \mu$ 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

$$\int_{\Sigma_4} H(r, \theta) \sin^4\theta d\Omega_4 = V_4 H(\mu, \bar{\theta}) \tag{B.7}$$

$$= V_4 \left[1 + \frac{R^4}{\mu^4}\left(1 + \mathcal{O}\left(\frac{u}{\mu}\right) \right) - N_w \frac{u}{\mu}\left(4\cos\bar{\theta} - 2\frac{u}{\mu} + \mathcal{O}\left(\frac{u^2}{\mu^2}\right) \right) \right] .$$

---

[28]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.

[29]We choose the letter μ in analogy with the (usually complex) parameter appearing in other known embeddings of four-cycles in the conifold [54, 55].

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/μ 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.

B.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 [56, 53]. 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

$$\det W \equiv \det \begin{pmatrix} X & U \\ V & Y \end{pmatrix} = XY - VU = 0 \tag{B.8}$$

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

$$\begin{pmatrix} X & U \\ V & Y \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0 \tag{B.9}$$

where (λ_1, λ_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.8) and (B.9) are equivalent, but when $W = 0$, *i.e.* at the tip, (λ_1, λ_2) are arbitrary and (B.9) defines a $\mathbb{CP}_1 \simeq S^2$. The radial direction is defined by

$$\mathrm{Tr}W^\dagger W = r^2 \,. \tag{B.10}$$

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.9):

$$W = \begin{pmatrix} -\lambda U & U \\ -\lambda Y & Y \end{pmatrix} \,. \tag{B.11}$$

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:

$$W = \begin{pmatrix} X & -\tilde{\lambda}X \\ V & -\tilde{\lambda}V \end{pmatrix} \,. \tag{B.12}$$

The complex structure is given by

$$\Omega = dU \wedge dY \wedge d\lambda = dV \wedge dX \wedge d\tilde{\lambda}. \tag{B.13}$$

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.9) is given by

$$W_0 = \begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \lambda_0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \lambda = 0. \tag{B.14}$$

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,

$$L_i = \begin{pmatrix} \cos\frac{\theta_i}{2} e^{\frac{i}{2}(\psi_i+\phi_i)} & -\sin\frac{\theta_i}{2} e^{-\frac{i}{2}(\psi_i-\phi_i)} \\ \sin\frac{\theta_i}{2} e^{\frac{i}{2}(\psi_i-\phi_i)} & \cos\frac{\theta_i}{2} e^{-\frac{i}{2}(\psi_i+\phi_i)} \end{pmatrix}, \quad i = 1, 2 \tag{B.15}$$

written in terms of Euler angles. This gives

$$W = L_1 W_0 L_2^\dagger, \quad \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = L_2 \begin{pmatrix} \lambda_0 \\ 0 \end{pmatrix} \tag{B.16}$$

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 [53]

$$ds_6^2 = \kappa^{-1}(\rho)d\rho^2 + \frac{1}{9}\kappa(\rho)\rho^2 e_\psi^2 + \frac{1}{6}\rho^2(e_{\theta_1}^2 + e_{\phi_1}^2) + \frac{1}{6}\left(\rho^2 + 6u^2\right)(e_{\theta_2}^2 + e_{\phi_2}^2), \tag{B.17}$$

where

$$\kappa(\rho) = \frac{\rho^2 + 9u^2}{\rho^2 + 6u^2}. \tag{B.18}$$

Here, following [53], we have defined a new radial coordinate ρ by

$$r^4 = \frac{4}{9}\rho^4\left(\frac{2}{3}\rho^2 + 6u^2\right), \tag{B.19}$$

The explicit expression for the e's is

$$e_\psi = d\psi + \sum_{i=1}^{2}\cos\theta_i d\phi_i \quad , \quad e_{\theta_i} = d\theta_i \quad , \quad e_{\phi_i} = \sin\theta_i d\phi_i. \tag{B.20}$$

B.2.1 The λUY 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

$$r^2 = \left(1 + |\lambda|^2\right)\left(|U|^2 + |Y|^2\right) . \tag{B.21}$$

As a result, this embedding does not give us the dipole suppression factor analogous to the (u/μ) of appendix B.1. The next-simplest embedding (whose defining equation depends on r) is

$$\lambda UY = \mu^3 , \quad \mu \in \mathcal{R} , \tag{B.22}$$

which in real coordinates gives

$$\psi = 0 , \quad \sin(\theta_2)\sin(\theta_1) = 4\frac{\mu^3}{r^2} \sim \frac{\mu^3}{\rho^3} \quad \text{for large } r . \tag{B.23}$$

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^{ind} is independent of $\phi_{1,2}$ and reads

$$
\begin{aligned}
g_4^{ind} &= \frac{\rho^2 \csc^4(\theta_2)}{20736\left(9u^2 + \rho^2\right)^3\left(\rho^4(9u^2 + \rho^2)\sin^2(\theta_2) - 54\mu^6\right)} \times \\
&\quad \Bigg(\left(6u^2 + \rho^2\right)\left(9u^2\rho + \rho^3\right)^2 \cos(4\theta_2) \\
&\quad - 4\cos(2\theta_2)\left(486u^6\rho^2 + 189u^4\rho^4 + u^2\left(24\rho^6 - 324\mu^6\right) - 27\mu^6\rho^2 + \rho^8\right) \\
&\quad + 3\left(54\rho^2\left(9u^6 + 2\mu^6\right) + 189u^4\rho^4 + 864u^2\mu^6 + 24u^2\rho^6 + \rho^8\right)\Bigg)^2 .
\end{aligned}
\tag{B.24}
$$

We see that there is a boundary beyond which the sign of the determinant becomes negative, which thus defines the integration boundary in ρ, θ_2-space:

$$
\rho_{min}(\theta_2) = \sqrt{3}\,\mu\,\sqrt{A - \frac{u^2}{\mu^2}\cdot\left(1 - \frac{u^2}{\mu^2}\frac{1}{A}\right)} \to \sqrt{3 \cdot 2^{1/3}}\,\mu\,\csc^{1/3}(\theta_2) \quad \text{for } \frac{u}{\mu} \ll 1
$$

$$
\text{with}: \quad A = \sqrt[3]{\sqrt{-\csc^2(\theta_2)\left(2\frac{u^6}{\mu^6} - \csc^2(\theta_2)\right)} - \frac{u^6}{\mu^6} + \csc^2(\theta_2)} . \tag{B.25}
$$

B.3 The shift of the four-cycle volume

The solution with the branes smeared over the S^2 was obtained in [53]. Later, the solutions with pointlike sources were given in [52]. 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

$$H = \sum_l (2l+1)H_l(\rho)\left[L_A^4 P_l(\cos(\theta_2)) + L_B^4 P_l(\cos(\theta_2))(-1)^l\right], \tag{B.26}$$

$$H_l = \frac{2}{9u^2}\frac{C_\beta}{\rho^{2+2\beta}} \,{}_2F_1\left(\beta, 1+\beta, 1+2\beta, -\frac{9u^2}{\rho^2}\right), \tag{B.27}$$

$$C_\beta = \frac{(3u)^{2\beta}\Gamma(1+\beta)^2}{\Gamma(1+2\beta)} \quad \beta = \sqrt{1+(3/2)l(l+1)}, \tag{B.28}$$

$$L_{A,B} = \frac{27}{16}4\pi g_s(\alpha')^2(N \mp N_w), \tag{B.29}$$

where ${}_2F_1$ is a hypergeometric function. We want to integrate this warp factor on some supersymmetric four-cycle Σ_4. This gives us an estimate of the inflaton dependence of the gauge kinetic function of a stack of D7-branes wrapping Σ_4.

Using this information and (B.26) and (B.27) we can now calculate the integral

$$\mathcal{V}_{warped} = \int_{\Sigma_4} d\rho\, d\theta_2\, d\phi_1\, d\phi_2 \sqrt{-g_4^{ind}}H(\rho,\theta_2) = 4\pi^2 \int_{\Sigma_4} d\rho\, d\theta_2 \sqrt{-g_4^{ind}}H(\rho,\theta_2) \tag{B.30}$$

numerically, as a function of μ. To facilitate this we will expand (B.26) up to $\ell = 1$, the dipole term, and take the large ρ limit

$$H(\rho,\theta_2) = \frac{L^4}{2\rho^4}\left[1 + 3(2\ell+1)\frac{N_w}{N}\frac{u^2}{\rho^2}P_\ell(\cos\theta_2)\Big|_{\ell=1}\right] = H_{\ell=0}(\rho) + \delta H_{\ell=1}(\rho,\theta_2) \quad . \tag{B.31}$$

We can now calculate

$$\mathcal{V}_{warped}^{(0)} = 4\pi^2 \int_0^\pi d\theta_2 \int_{\rho_{min}(\theta_2)}^{\rho_R} d\rho\sqrt{-g_4^{ind}}H_{\ell=0}(\rho) \tag{B.32}$$

$$\delta\mathcal{V}_{warped}^{\ell=1}(\delta\theta_2) = 4\pi^2 \int_0^\pi d\theta_2 \int_{\rho_{min}(\theta_2)}^{\rho_R} d\rho\sqrt{-g_4^{ind}}\delta H_{\ell=1}(\rho,\theta_2+\delta\theta_2) \tag{B.33}$$

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 μ between $\delta\mathcal{V}_{warped}(\delta\theta_2)$ and $\mathcal{V}_{warped}^{(0)}$. This is displayed in Fig. 10. We see clearly that the leading $\ell = 0$ term scales as $u^0\mu^0 = const.$ 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
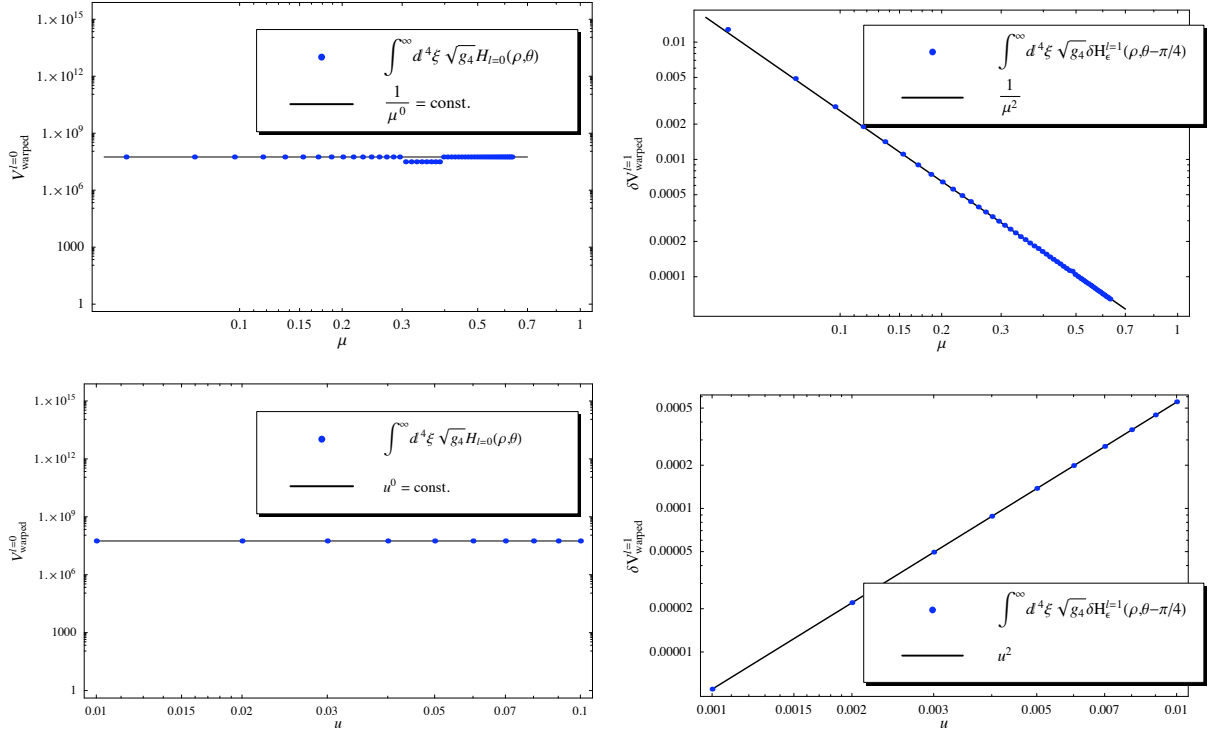
Figure 10: 1st row: Plot of $\mathcal{V}^{(0)}_{warped}$ and $\delta\mathcal{V}_{warped}(\delta\theta_2)$ as functions of μ at constant $u = 0.01$. 2nd row: Plot of $\mathcal{V}^{(0)}_{warped}$ and $\delta\mathcal{V}_{warped}(\delta\theta_2)$ as functions of u at constant $\mu = 0.1$. The leading $\ell = 0$ term scales as $u^0\mu^0 = const.$ 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 $\rho \to \infty$), while for a finite cutoff, resembling a crude approximation to a compact setting, there remains a weak dependence of the $\ell = 0$ term on μ, of the form $(u/\mu)^\delta$, where $\delta \to 0$ for $\rho_{bulk} \to \infty$. For the example we have chosen $\epsilon \equiv N_w/N = 0.1$.

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 μ/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/ρ^2 should be replaced by the scaling u^2/μ^2.

C   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.

C.1 The effective theory

The DBI action for a D5-brane is

$$S = T_5 \int d^4x \, dy \, dz \sqrt{-\det\left(G_{ab}^{ind} + \mathcal{F}_{ab}\right)}, \tag{C.1}$$

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 $a, b = 0, \ldots, 5$; spacetime indices are $\mu, \nu = 0, \ldots, 3$ as usual; ten-dimensional indices are $M, N = 0, \ldots, 9$; six-dimensional compact indices are $m, n = 4, \ldots, 9$; and indices transverse to the D5-brane are $i, j = 6, \ldots, 9$. We first expand the square root using

$$\sqrt{\det(M_0 + \delta M)} = \sqrt{\det M_0} \left\{ 1 + \frac{1}{2}\text{Tr}(M_0^{-1}\delta M) + \frac{1}{8}[\text{Tr}(M_0^{-1}\delta M)]^2 \right. \tag{C.2}$$

$$\left. -\frac{1}{4}\text{Tr}(M_0^{-1}\delta M M_0^{-1}\delta M) + \ldots \right\}. \tag{C.3}$$

We will consider a background with two-form flux on the two-cycle

$$\int \mathcal{F} = \int B = \int dy \wedge dz B_{yz}(x, y, z) = b(x) = b, \tag{C.4}$$

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

$$B_{MN} = b\delta_{My}\delta_{Nz} - b\delta_{Mz}\delta_{Ny}, \quad F_{ab} = 0, \tag{C.5}$$

$$ds_{10}^2 = g_{\mu\nu}dx^\mu\nu + g_{yy}dy^2 + g_{zz}dz^2 + 2g_{yz}dydz + g_{ij}dy^idy^j. \tag{C.6}$$

Hence

$$(M_0)_{ab} = \begin{pmatrix} g_{\mu\nu} & & \\ & g_{yy} & g_{yz} + b \\ & g_{zy} - b & g_{zz} \end{pmatrix}. \tag{C.7}$$

The perturbations are

$$(\delta M)_{ab} = \partial_a X^i \partial_b X^j (g_{ij} + B_{ij}) + F_{ab} + \delta B_{ab}. \tag{C.8}$$

The calculation is simplified by the block-diagonal form of the background M_0. The 2×2 block is the sum of a symmetric and an antisymmetric piece that we call S and A respectively. We have that

$$\det(A + S) = \det(A) + \det(S), \tag{C.9}$$

$$(S + A)^{-1} = S^{-1}\frac{\det(S)}{\det(A) + \det(S)} + A^{-1}\frac{\det(A)}{\det(A) + \det(S)}, \tag{C.10}$$

which substantially simplifies the calculation. Using (C.2) we get at leading order

$$S \ = \ T_5 \int d^4x \, dy \, dz \sqrt{-g_4} \sqrt{g_2 + b^2} \Big[1 + \frac{1}{2} \partial_\mu X^i \partial^\mu X_i \tag{C.11}$$

$$+ \frac{1}{2} \frac{g_2}{g_2 + b^2} \left(\partial_y X^i \partial^y X_i + \partial_z X^i \partial^z X_i \right) \tag{C.12}$$

$$+ \frac{1}{2} \frac{2b}{g_2 + b^2} \left(\partial_y X^i \partial_z X^j \delta B_{ij} + F_{yz} + \delta B_{yz} \right) \Big] + \dots, \tag{C.13}$$

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

$$m_{bKK}^2 = \frac{g_2}{g_2 + b^2} m_{KK}^2 \,, \tag{C.14}$$

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 $\sqrt{g_2}/b \simeq L^2/b \ll 1$.[30] 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.

C.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 (C.14) 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.

   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.

---

[30]We have assumed for simplicity that the internal space is isotropic, with typical size $L\sqrt{\alpha'}$.

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 $\dot{m}_{bKK}/m^2_{bKK}$ and substituting the constraints on the volume, and hence on the Kaluza-Klein mass, from §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 [57, 58] 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.

D   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 §6.5, finding numerical values leading to the dot in figure 8.

D.1   Intersection numbers: set I

We consider as a toy-model Calabi-Yau manifold one with $H^{1,1}_+ = \mathrm{span}(\omega^L, \omega^+)$ for the orientifold-even homology two-cycles and $H^{1,1}_- = \mathrm{span}(\omega^-)$ for the orientifold-odd homology two-cycles. We assume the following simple set of intersection numbers

$$c_{LLL} = c_{LL+} = c_{+--} = 1\,, \tag{D.1}$$

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

$$\mathcal{V}_E \;=\; \frac{1}{6}c_{\alpha\beta\gamma}v^\alpha v^\beta v^\gamma\,, \quad \tau_\alpha = \partial_{v^\alpha}\mathcal{V}_E = \frac{1}{2}c_{\alpha\beta\gamma}v^\beta v^\gamma\,, \tag{D.2}$$

one finds

$$v_L = \sqrt{2\tau_+}\,, \; v_+ = \frac{\tau_L - \tau_+}{\sqrt{2\tau_+}}\,, \tag{D.3}$$

and

$$\mathcal{V}_E = \frac{\sqrt{2\tau_+}}{2}\tau_L - \frac{\sqrt{2}}{6}\tau_+^{3/2}\,. \tag{D.4}$$

D.2 Intersection numbers: set II

Again assuming $H_+^{1,1} = \mathrm{span}(\omega^L, \omega^+)$ and $H_-^{1,1} = \mathrm{span}(\omega^-)$, we consider the intersection numbers

$$c_{LLL} = c_{L++} = c_{+--} = 1\,, \tag{D.5}$$

with all the others vanishing. We find

$$v_L = \frac{1}{\sqrt{2}}\left((\tau_L + \tau_+)^{1/2} + (\tau_L - \tau_+)^{1/2}\right)\,, \; v_+ = \frac{1}{\sqrt{2}}\left((\tau_L + \tau_+)^{1/2} - (\tau_L - \tau_+)^{1/2}\right)\,, \tag{D.6}$$

and

$$\mathcal{V}_E = \frac{1}{3\sqrt{2}}\left((\tau_L + \tau_+)^{3/2} + (\tau_L - \tau_+)^{3/2}\right)\,. \tag{D.7}$$

References

[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] E. Silverstein and A. Westphal, "Monodromy in the CMB: Gravity Waves and String Inflation," arXiv:0803.3085 [hep-th].

[5] L. McAllister, E. Silverstein and A. Westphal, "Gravity Waves and Linear Inflation from Axion Monodromy," arXiv:0808.0706 [hep-th].

[6] 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]].

[7] 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]].

[8] J. Martin and C. Ringeval, "Superimposed Oscillations in the WMAP Data?," Phys. Rev. D **69**, 083515 (2004) [arXiv:astro-ph/0310382].

[9] J. Martin and C. Ringeval, "Exploring the superimposed oscillations parameter space," JCAP **0501**, 007 (2005) [arXiv:hep-ph/0405249].

[10] 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].

[11] X. Chen, R. Easther and E. A. Lim, "Large non-Gaussianities in single field inflation," JCAP **0706** (2007) 023 [arXiv:astro-ph/0611645].

[12] 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].

[13] 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]].

[14] C. Pahud, M. Kamionkowski and A. R. Liddle, "Oscillations in the inflaton potential?," arXiv:0807.0322 [astro-ph].

[15] D. Baumann and L. McAllister, "Advances in Inflation in String Theory," arXiv:0901.0265 [hep-th].

[16] K. Freese and D. Spolyar, "Chain inflation: 'Bubble bubble toil and trouble'," JCAP **0507**, 007 (2005) [arXiv:hep-ph/0412145].

[17] K. Freese, J. T. Liu and D. Spolyar, "Chain inflation via rapid tunneling in the landscape," arXiv:hep-th/0612056.

[18] D. Chialva and U. H. Danielsson, "Chain inflation revisited," JCAP **0810**, 012 (2008) [arXiv:0804.2846 [hep-th]].

[19] S. Weinberg, *Cosmology*, Oxford University Press, Oxford, UK (2008).

[20] Bateman Manuscript Project, *Higher Transcendental Functions, Volume 2*, McGraw-Hill, New York, USA, (1953-55).

[21] 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]].

[22] 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].

[23] 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].

[24] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Butterworth-Heinemann, Oxford, UK (2002). "Course of Theoretical Physics. Vol. 3: Quantum Mechanics,"

[25] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," JHEP **0305** (2003) 013 [arXiv:astro-ph/0210603].

[26] 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]].

[27] 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].

[28] 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].

[29] The code is available at http://camb.info/

[30] http://lambda.gsfc.nasa.gov/

[31] 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].

[32] The code is available at http://cosmologist.info/cosmomc/

[33] J. Polchinski, *String Theory. Vol. 2: Superstring Theory and Beyond*, Cambridge University Press, Cambridge, UK (1998).

[34] 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].

[35] 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].

[36] X. G. Wen and E. Witten, "World Sheet Instantons And The Peccei-Quinn Symmetry," Phys. Lett. B **166**, 397 (1986).

[37] M. Dine and N. Seiberg, "Nonrenormalization Theorems in Superstring Theory," Phys. Rev. Lett. **57**, 2625 (1986).

[38] E. Pajer, "Inflation at the Tip," JCAP **0804**, 031 (2008) [arXiv:0802.2916 [hep-th]].

[39] 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].

[40] P. Svrcek and E. Witten, "Axions in string theory," JHEP **0606** (2006) 051 [arXiv:hep-th/0605206].

[41] 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].

[42] 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].

[43] 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]].

[44] 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]].

[45] D. J. Gross and J. H. Sloan, "The Quartic Effective Action for the Heterotic String," Nucl. Phys. B **291**, 41 (1987).

[46] 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].

[47] 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].

[48] E. Witten, "Non-Perturbative Superpotentials In String Theory," Nucl. Phys. B **474**, 343 (1996) [arXiv:hep-th/9604030].

[49] T. W. Grimm, "Non-Perturbative Corrections and Modularity in N=1 Type IIB Compactifications," JHEP **0710**, 004 (2007) [arXiv:0705.3253 [hep-th]].

[50] D. Robles-Llana, M. Rocek, F. Saueressig, U. Theis and S. Vandoren, "Nonperturbative corrections to 4D string theory effective actions from SL(2,Z) duality and supersymmetry," Phys. Rev. Lett. **98**, 211602 (2007) [arXiv:hep-th/0612027].

[51] V. Balasubramanian, P. Berglund, J. P. Conlon and F. Quevedo, JHEP **0503**, 007 (2005) [arXiv:hep-th/0502058].

[52] I. R. Klebanov and A. Murugan, "Gauge/Gravity Duality and Warped Resolved Conifold," JHEP **0703**, 042 (2007) [arXiv:hep-th/0701064].

[53] L. A. Pando Zayas and A. A. Tseytlin, "3-branes on resolved conifold," JHEP **0011**, 028 (2000) [arXiv:hep-th/0010088].

[54] S. Kuperstein, "Meson spectroscopy from holomorphic probes on the warped deformed conifold," JHEP **0503**, 014 (2005) [arXiv:hep-th/0411097].

[55] P. Ouyang, "Holomorphic D7-branes and flavored N = 1 gauge theories," Nucl. Phys. B **699**, 207 (2004) [arXiv:hep-th/0311084].

[56] P. Candelas and X. C. de la Ossa, "Comments on Conifolds," Nucl. Phys. B **342**, 246 (1990).

[57] N. Barnaby, C. P. Burgess and J. M. Cline, "Warped reheating in brane-antibrane inflation," JCAP **0504**, 007 (2005) [arXiv:hep-th/0412040].

[58] L. Kofman and P. Yi, "Reheating the universe after string theory inflation," Phys. Rev. D **72**, 106001 (2005) [arXiv:hep-th/0507257].