

# Fencing in the swampland: quantum gravity constraints on large field inflation

---

**Jon Brown, William Cottrell, Gary Shiu and Pablo Soler**

*Department of Physics, University of Wisconsin,  
1150 University Avenue, Madison, WI, 53706 U.S.A.*

*E-mail:* [jhbrown2@wisc.edu](mailto:jhbrown2@wisc.edu), [wcottrell@wisc.edu](mailto:wcottrell@wisc.edu), [shiu@physics.wisc.edu](mailto:shiu@physics.wisc.edu),  
[psoler@physics.wisc.edu](mailto:psoler@physics.wisc.edu)

**ABSTRACT:** In this note we show that models of natural inflation based on closed string axions are incompatible with the weak gravity conjecture (WGC). Specifically, we use T-duality in order to map the bounds on the charge-to-mass ratio of particles imposed by the WGC, to constraints on the ratio between instanton actions and axion decay constants. We use this connection to prove that if the WGC holds, even when multiple axions are present and mix with each other, one cannot have large axion decay constants while remaining in a regime of perturbative control. We also discuss the extension of the WGC to discrete symmetries and its possible impact on models with axion monodromy, and the distinction between the strong and mild versions of the WGC. We argue that, if the strong version is violated, the constraints on large field inflation can be evaded while still satisfying the mild one. Finally, we offer some speculations regarding the import of these results to the general theory of inflation.

**KEYWORDS:** Cosmology of Theories beyond the SM, Superstring Vacua

**ARXIV EPRINT:** [1503.04783](https://arxiv.org/abs/1503.04783)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The weak gravity conjecture</b>	<b>3</b>
2.1	The conjecture	3
2.2	The convex hull condition	4
2.3	Generalization to $p$ -form symmetries	5
2.4	Discrete symmetries and the WGC	7
<b>3</b>	<b>Mapping instantons and axions to particles and gauge fields</b>	<b>7</b>
3.1	Setup	8
3.2	T-Duality	10
3.3	Black holes	11
<b>4</b>	<b>WGC constraints on axions</b>	<b>12</b>
4.1	General limits on axion inflation	13
4.2	A possible loophole	15
<b>5</b>	<b>Torsion cycles and generalized WGC</b>	<b>16</b>
<b>6</b>	<b>Tension with holography</b>	<b>18</b>
<b>7</b>	<b>Conclusions and speculations</b>	<b>19</b>

---

## 1 Introduction

Inflation is the leading theory solving a host of cosmological problems in the early universe, and it provides a seed for the observed large scale structure. Several inflationary models, such as chaotic [1] and natural [2] inflation, require the inflaton to traverse a large distance (in Planck units) in field space in order to produce the requisite number of e-folds. In the past decade, many such large field inflation models have been explored within string theory.<sup>1</sup> Guaranteeing a large field range presents problems of its own as Planck suppressed operators generically become sizable and thus the effective field theory would a priori lose perturbative control. The most popular solution to this problem is to consider axions as inflaton candidates since they enjoy a perturbative shift symmetry which suppresses quantum corrections. Instanton effects may then generate a sufficiently small potential while breaking the continuous shift symmetry to a discrete subgroup. Modding the axion

---

<sup>1</sup>The ultraviolet sensitivity of large field inflation motivates one to consider inflation within string theory. For a pedagogical review of string inflation, see e.g., [3].

moduli space by this subgroup, we find that the axion has a finite field range whose ‘radius’ is known as the ‘axion decay constant’. This decay constant will then set a bound on the effective field range of inflation.

From the perspective of effective field theory (EFT), requiring a large decay constant is simply a matter of tuning. However, for field values near the Planck scale, the theory’s embedding in a complete theory of quantum gravity becomes a relevant question. Banks et al. [4] argued via examples that, in stringy models, single-field large axion decay constants arise *only* when control of higher order instantons is lost and new light states appear in the theory. Thus, the form of the EFT axion potential is obscured. Most efforts in constructing large field inflation models have therefore been focused on evading this constraint by a variety of EFT mechanisms involving multiple axions [5–7]. The idea is essentially that while each axion decay constant must be sub-Planckian, the net effective decay constant may be larger, as in the alignment [6] or N-flation [7] scenarios.

Unfortunately, it is often very difficult to test directly whether these models have stringy embeddings, partially due to moduli stabilization issues. There have been several attempts [8–11] to engineer string realizations of the alignment scenario. In [8, 10], for instance the role of the axion is played by a p-form field reduced on a cycle in the internal manifold, with instantons generating the axion potential represented by Euclidean D-branes wrapping it. In general, one would expect that the tension of the D-brane shrinks the cycle it wraps and thus reduces the action of instantons wrapping the brane. This could very well lead to a loss of control as intimated by [4]. One needs a mechanism to stabilize the cycle’s contraction but incorporating such effects systematically in an inflationary background seems intractable. It was argued in Bachlechner et al. [11] that a ‘large’ axion field range can be obtained via multiple closed string axions using the semi-explicit construction of [12]. We will argue in this work that such a possibility is in conflict with general properties of quantum gravity. According to our results, we predict that a more careful analysis of the stabilization process, e.g.  $\alpha'$  and instanton corrections, or one-loop determinants will *either* lower the effective field range *or* increase the size of the instanton corrections. In anticipation of our results, we will obtain a precise numerical upper bound on the inflaton field range as our analysis does not allow for an  $\mathcal{O}(1)$  fudge factor.

The issue of determining when an EFT has a string embedding is precisely the problem of defining the ‘swampland’ [13, 14]. It would be very useful to have tools which allow one to tell, at a glance, whether a given EFT makes sense within the context of quantum gravity. One such criteria was proposed in [15], where it was conjectured that in any UV embeddable EFT ‘gravity is the weakest force’. This is known as the ‘Weak Gravity Conjecture’ (WGC). For example, electric repulsion always beats gravitational attraction between two electrons. In [16] an extension of this conjecture to theories with many U(1)’s and a general spectrum of charged particles was provided. The extension simply states that the convex hull of a certain charge matrix must contain the ‘unit ball’ (the precise formulation will be given later).

The main purpose of this work is to relate the WGC to a condition on the axion decay constant via T-duality and the convex hull condition of [16] to models with multiple axions, and to precisify the assumptions under which natural inflation is excluded. Specifically, we

will show that if the WGC holds, even when mechanisms like kinetic mixing, alignment, and large  $N$  enhancement are considered, one cannot get a large field range while staying in a regime of perturbative control.<sup>2</sup> Moreover, we will argue that the WGC implies precise numerical bounds on inflation models rather than fuzzy statistical bounds. These results apply to any model of natural inflation based on closed string axions.

An interesting exception could arise if we were to somehow suppress higher instanton corrections. One mechanism for doing this is to consider axions defined on torsional cycles. In fact, this line of thinking leads one directly to the F-term monodromy model considered in [18]. However, we will also discuss a possible generalization of the WGC to massive  $p$ -form fields, and argue that this model can be constrained as well.

This paper is structured as follows. In section 2 we review the basics of the weak gravity conjecture and describe its formalization in terms of the convex hull condition. Section 3 describes the T-duality between the usual WGC and the WGC as applied to axions, thus demonstrating that the conjecture is ‘covariant’ in an appropriate sense and also fixing the precise size of the convex hull. Section 4 proves that the convex hull condition does indeed prohibit natural inflation models in general (a caveat to our results will be discussed). Section 5 discusses an interesting class of models that falls outside our proof but nevertheless raise issues relevant to understanding the WGC. Next, in section 6 we point out a general tension between inflation models and holography. Finally, we offer our conclusions.

The results of [19, 20], which appeared while this paper was being finalized, contain a significant overlap with our results, which were presented at the MCTP Workshop “String/M-theory Compactifications and Moduli Stabilization”.

## 2 The weak gravity conjecture

In this section we review the original formulation of the WGC and some of the arguments that support it [15]. We also describe, along the lines of [16], its generalization to the case in which there are multiple gauge fields, and furthermore speculate about the possibility of applying the WGC to discrete gauge symmetries. We stress, nevertheless, that the goal of this paper is not to critically address the validity of the conjecture, but rather to extract consequences from it and apply them to models of inflation. Whether the conjecture is indeed a generic consistency requirement of quantum gravity is a question that will require further analysis.

### 2.1 The conjecture

The WGC simply states that gravity should be the weakest force in any consistent theory that includes quantum gravity. Stated another way, this just says that there should be a lower bound on the strength of gauge interactions. The simplest argument to understand this is to consider a  $U(1)$  symmetry with coupling constant  $g$ , and take  $g \rightarrow 0$ . In that

---

<sup>2</sup>In this paper we focus on constraints arising from applications of the WGC to the *electric* spectrum of  $U(1)$  theories. Further constraints can be obtained by imposing that the *magnetic* spectrum must satisfy the WGC as well [15, 17].

limit, the gauge bosons decouple and the symmetry becomes global, running afoul of ‘folk theorems’ of quantum gravity. As is well known, in theories with global continuous symmetries one can construct black holes with an arbitrary charge that cannot be radiated away, hence leading to an infinite number of remnants, in violation of the Covariant Entropy Bound (CEB) [21].

A more precise formulation of the WGC is that for each long range force there should exist at least one particle for which the gravitational attraction is compensated by a stronger gauge repulsion. Following [15], let us consider a U(1) gauge theory with a single family of particles with mass  $\mu$  and charge  $e$  which does not satisfy the WGC, i.e. for which the electric repulsion is overwhelmed by the gravitational attraction. In that case, one can form perfectly stable bound states of an arbitrary number  $N$  of particles with charge  $eN$  and a mass  $M < N\mu$  (the difference being the binding energy of the system). For large  $N$  these objects will form a stable black hole. Asymptotically, in the limit  $N \rightarrow \infty$ , the black hole will be extremal. This infinite number of safely bound perfectly stable states leads again to a violation of the CEB.<sup>3</sup>

The way to overcome this problem is to postulate that there must be a particle that allows the infinite series of stable states to decay. The charge to mass ratio of such a particle must then be greater than or equal to that of an extremal black hole

$$\frac{q}{m} \geq \left( \frac{Q}{M} \right)_{\text{EBH}} . \tag{2.1}$$

Now, there are two versions of the WGC, depending on which state is required to satisfy the condition (2.1). The *strong* version (strong-WGC) requires that it must be the lightest charged particle in the spectrum. The *mild* version (mild-WGC) puts no extra condition on the state that satisfies the inequality. In fact, mild-WGC can be rephrased as the statement that there should be only a finite number of exactly stable and safely bound states in the theory. In the presence of multiple U(1) symmetries, the requirement applies along each direction in charge space.

It is clear that the evidence for the strong-WGC in terms of the CEB are not as solid as those for the mild version. Nevertheless, as already mentioned in [15], there is no known counterexample to strong-WGC in string theory.<sup>4</sup> In section 6 we will argue that the covariant entropy bound does indeed provide support for the strong-WGC, however, it would be nice if one could produce some more direct and independent evidence.

## 2.2 The convex hull condition

We follow in this section the approach of [16] in order to understand the consequences of the WGC in setups in which there are several U(1) symmetries. For each particle  $i$  with charges  $\vec{q}_i$  and masses  $m_i$  in the system, it is convenient to introduce the vector  $\vec{z}_i = \vec{q}_i \frac{M_p}{m_i}$ . In the conventions described in section 3.3, these vectors are normalized such that for

---

<sup>3</sup>Strictly speaking, the results of [21] apply to an infinite number of stable states in a *finite* mass range, while a violation of the WGC results in an infinite number of states in an *infinite* mass range. The considerations of [21] can in principle be generalized to address this situation as well [22].

<sup>4</sup>With the possible exception of [11], to be discussed later.

macroscopic extremal black holes<sup>5</sup> in 4d  $|\vec{Z}|_{\text{EBH}} \equiv 1$ . Hence, for extremal black holes,  $\vec{Z}$  lies on the unit sphere, while for non extremal ones  $\vec{Z}$  lies inside the unit ball. (See figure 1.)

The WGC condition, which states that extremal black holes should be at most marginally stable, translates into a condition on the vectors  $\vec{z}_i$  of the particles in the spectrum. Consider an extremal black hole with charge  $\vec{Q}$ , mass  $M$  and vector  $\vec{Z} = \vec{Q} \frac{M_p}{M}$ , which by definition lies on the unit sphere  $|\vec{Z}| = 1$ . The (in)stability condition implies that there must exist a combination of  $n_i$  particles of species  $i$  into which the black hole can decay, i.e.  $\vec{Q} = \sum n_i \vec{q}_i$  and  $M \geq \sum n_i m_i$ . Define the mass fractions  $\sigma_i = n_i m_i / M$ . Energy and charge conservation imply  $\sum_i \sigma_i \leq 1$  and  $\vec{Z} = \sum_i \sigma_i \vec{z}_i$ . In other words,  $\vec{Z}$  is a subunitary weighted average of  $\vec{z}_i$ . Since extremal black holes can be created along any direction in charge space, the condition implies that the unit ball should be contained in the convex hull spanned by the vectors  $\pm \vec{z}_i$  of the system (see figure 1).

The mild-WGC states that along any given direction in charge space, there is *some* combination of particles with total mass  $m$  and total charge  $\vec{q}$  whose vector  $\vec{z} = \vec{q} \frac{M_p}{m}$  lies outside of the unit ball. It is natural to generalize the strong-WGC by requiring that the lightest (possibly multi-particle) state in any given direction in charge space satisfies the same condition, that is,  $|\vec{z}_{\text{lightest}}| \geq 1$ .<sup>6</sup> Unless otherwise stated (see section 4.2) our discussion applies to the mild-WGC, and hence automatically to the strong-WGC as well.

In figure 2 we show several examples that fail to satisfy the convex hull condition in some direction of charge space, even if there are several states for which  $|\vec{z}| \geq 1$ .

### 2.3 Generalization to $p$ -form symmetries

It was suggested in [15] that the WGC could be generalized to the case in which  $p$ -dimensional objects (e.g.  $D_{p-1}$  branes in string theory) couple to  $p$ -form gauge fields. The conjecture would imply that in theories of quantum gravity, there should exist in the spectrum at least one charged object whose tension to charge ratio satisfies

$$\frac{T}{gM_p} \leq 1 \tag{2.2}$$

where  $g$  is the charge density and “1” should be defined in terms of an extremal  $(p-1)$  brane.

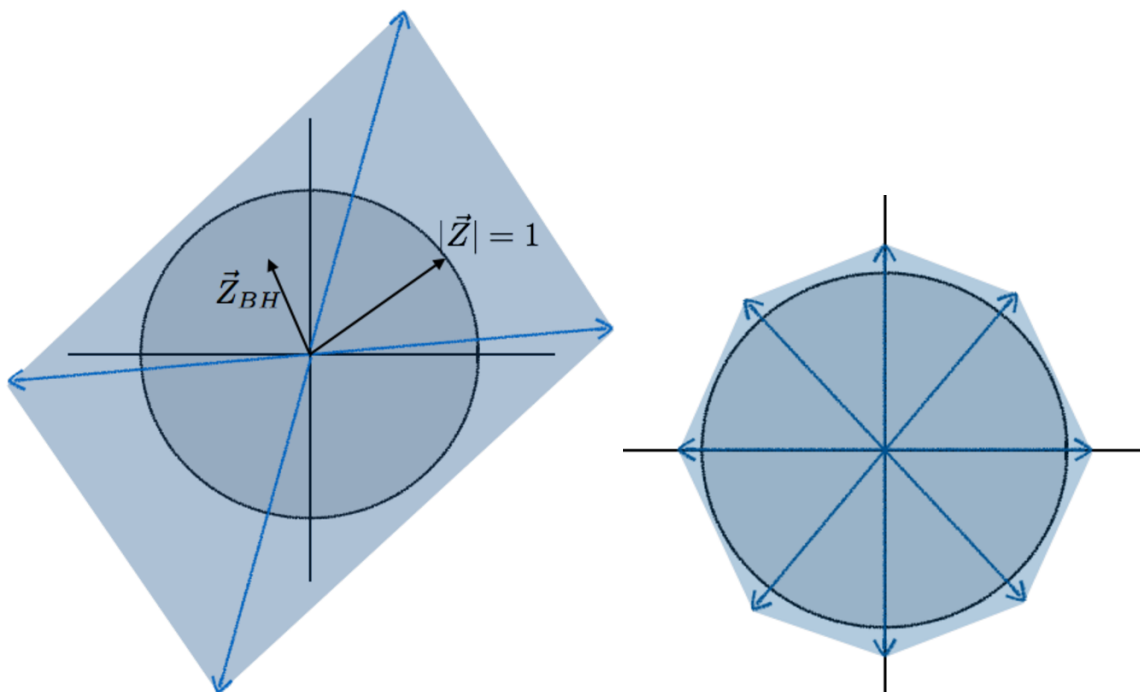
If such a generalization holds, it would have extremely important consequences for axions and the instantons which couple to them, which corresponds to the case  $p=0$ . The WGC would imply that axions with large (trans-Planckian) decay constants do not arise in perturbative regimes of theories of quantum gravity.

Although this generalization seems reasonable, and there is convincing circumstantial evidence that axions with large decay constant do not arise in string theory [4, 23], it is

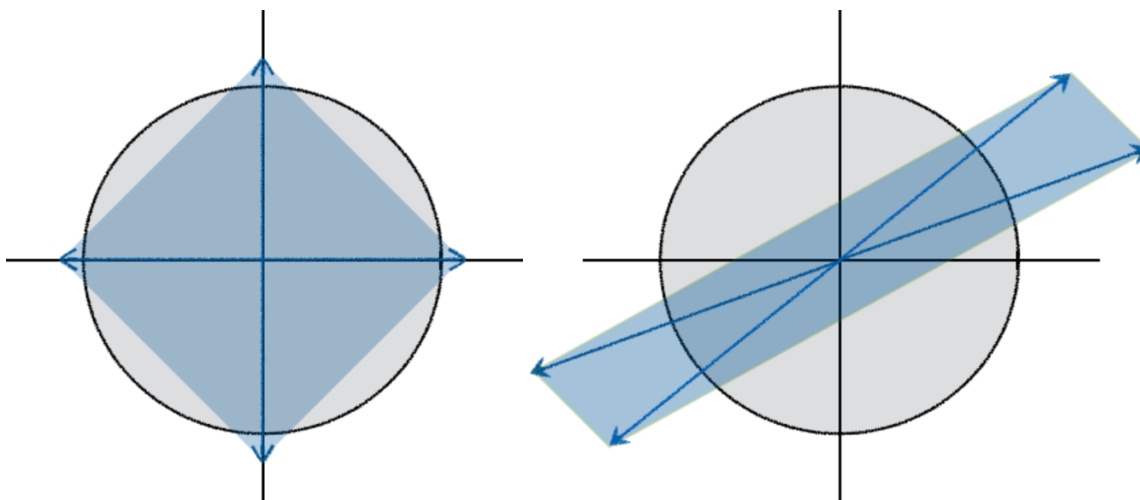
---

<sup>5</sup>The numerical constant,  $|\vec{Z}_{\text{EBH}}|$ , will change depending on the space-time dimension, as discussed in 3.3.

<sup>6</sup>Several slightly different generalizations are possible, all of which reduce to the standard one in the single U(1) case. The one formulated here is perhaps the most straightforward generalization of the original one as formulated in [15]. This is sufficient for the purposes of this work, for which the subtleties of the different formulations will not to play a role. Nevertheless, it would be interesting to understand which particular conditions are generically satisfied.



**Figure 1.** The convex hull condition as applied to two  $U(1)$  groups. Vectors corresponding to extremal black holes lie on the unit circle, while non-extremal black holes lie inside the unit disc. On the left (right) hand side, the convex hull defined by the  $\vec{z}$  vectors of two (four) charged particles and their antiparticles (blue arrows) ensure that the WGC is satisfied.



**Figure 2.** Two situations in which the convex hull condition is not satisfied. Black holes within the unit disc but outside of the convex hull of the two particles (blue areas) are stable and lead to an infinite number of remnants.

hard to establish a direct connection between axions and the standard form of the WGC in terms of black holes and the CEB. This connection would prove even more useful in the application of the convex hull condition to the case in which several axions mix with each other to generate a large effective decay constant. In the following sections, using

T-duality, we will provide such a connection for axions that arise in the closed string sector type II theories, and show how the WGC can be used to constrain a large class of models of inflation.

## 2.4 Discrete symmetries and the WGC

Before doing that, we would like to suggest a possible further generalization of the standard WGC to the case in which the symmetries are discrete. As we will see, if such a generalization holds, it could be used to constrain models of axion monodromy inflation, which cannot be related to the standard form of the WGC.

It is commonly assumed that discrete symmetries, just as continuous ones, should be gauged in a consistent theory of quantum gravity [24]. As we mentioned above, the evidence in the continuous case is that global symmetries seem to lead unavoidably to remnant black holes with an infinite entropy, or equivalently to an infinite number of remnants only labeled by their global charge. The case against global discrete symmetries, say  $\mathbb{Z}_k$ , is much weaker. In particular, the entropy or equivalently the number of remnants is determined by  $k$ , the order of the discrete symmetry. Although for a sufficiently large  $k$  one would indeed clash with the CEB, discrete symmetries of low order do not seem to be necessarily problematic. Nevertheless, it is commonly understood that all symmetries, including discrete ones of any order, are gauged. Indeed, as far as we know, all discrete symmetries found so far in string theory setups satisfy this assumption, giving circumstantial support to the conjecture.

Now, if a discrete symmetry, say  $\mathbb{Z}_k$ , is gauged, it must be a subgroup of a broken continuous gauge symmetry with coupling constant  $g$ . Just like we discussed in section 2.1, by sending  $g \rightarrow 0$  one recovers a global symmetry. If the common lore expressed in the previous paragraph is correct and discrete global symmetries are pathological, quantum gravity should prevent this limit from happening. This gives evidence that the WGC should be generalizable to the case of discrete symmetries.

In analogy with section 2.1, imagine one has a  $U(1)$  gauge symmetry with coupling constant  $g$  which is broken down to a discrete  $\mathbb{Z}_k$  subgroup, and a single family of particles charged under it. If the repulsive force mediated by the massive gauge bosons between two such particles is weaker than the gravitational attraction even at short distances, one could form bound states of arbitrary  $U(1)$  charge. The first  $k$  such states, distinguished by their discrete charge, would be exactly stable unless there is some other particle species with lower mass-to-charge ratio. For large enough  $k$ , one would find a conflict with the CEB. Although there is no fully convincing evidence that low values of  $k$  are problematic, it is not unreasonable to conjecture that the WGC applies to discrete symmetries, even to the lowest order  $k = 2$ . We are not aware of any counter example in string theory that contradicts this statement.

## 3 Mapping instantons and axions to particles and gauge fields

Now, we would like to verify that a restriction on large axionic field ranges does indeed follow from the standard WGC. We will do so by compactifying the theory on a circle to three dimensions, performing a T-duality and taking the decompactification limit in



the dual circle. Instantons coupled to axions in the original theory transform under this procedure into particles that couple to U(1) gauge fields, to which the WGC can be applied straightforwardly. The constraints on the spectrum of charges and masses from the WGC translate into bounds on the ratios of instanton actions and axion decay constants.

### 3.1 Setup

As an example we consider  $C_2$  axions in Type IIB, though the procedure below works for any closed string axion. In the conventions of [25, 26], we have the following action:

$$S_{\text{IIB}} \supset -\frac{1}{4\kappa_0^2} \int d^{10}x \sqrt{-g} \left( \sum_p |F_p|^2 \right) + \mu_p \int_{Dp} C_{p+1}. \quad (3.1)$$

The p-form equation of motion and charge quantization conditions are then:

$$\begin{aligned} d * F_{p+2} &= 2\kappa_0^2 \mu_p \delta^{9-p}(x_\perp) dV_\perp \\ \int_{S^{8-p}} *F_{p+2} &= 2\kappa_0^2 \mu_p. \end{aligned} \quad (3.2)$$

We may introduce a basis of 2-forms,  $w_a \in H^2(X, \mathbb{Z})$ , satisfying:

$$\int_X w_a \wedge *w_b = \frac{V_X}{\alpha'^2} K_{ab}. \quad (3.3)$$

The matrix  $K_{ab}$ , the Kähler moduli space metric, is chosen to be dimensionless and it may encode arbitrary kinetic mixing. We now define our axion fields as:

$$C_2 = \sum a^i w_i. \quad (3.4)$$

The fields  $a^i$  will have a natural periodicity of:

$$\Delta a^i = (2\pi l_s)^2. \quad (3.5)$$

This may be understood either as a consequence of quantizing the dual 5-brane charge, or as a result of instantons. To see how this emerges from the latter, note that the potential for  $a^i$  arises entirely from instantons in the form of euclidean  $D1$  branes wrapping a cycle  $\Sigma_k \in H_2(X, \mathbb{Z})$ . An axion background will contribute a term in the action proportional to  $\mu_1 q_k^i a^i$ , where  $q_k^i = \int_{\Sigma_k} w^i$  are integers. This can in turn induce a non-perturbative contribution to the 4d potential of the form:

$$V \supset \sum_k \Lambda^4 e^{-m_k} (1 - \cos(\mu_1 q_k^i a^i)). \quad (3.6)$$

When  $q_k^i = 1$  for at least one 2-cycle coupling to  $a^i$  then the periodicity is  $\Delta a^i = 2\pi/\mu_1 = (2\pi l_s)^2$  as claimed. Otherwise, the periodicity is given by:

$$\Delta a^i = \frac{(2\pi l_s)^2}{\text{gcd}(q_k^i)}. \quad (3.7)$$

The true field range, however, is only determined after we have properly normalized the kinetic terms in the action. The relevant terms are

$$S_{\text{IIB}} \supset \frac{M_p^2}{2} \int d^4x \sqrt{-g} \left( R + \frac{g_s^2}{2(\alpha')^2} K_{ij} \partial a^i \partial a^j \right) \quad (3.8)$$

$$\frac{M_p^2}{2} = \frac{V_X}{(2\pi)^7 (\alpha')^4 g_s^2}.$$

Let  $D_j^i$  be a matrix diagonalizing  $K$ , i.e.,  $D^\top K D = 1$ . Then the canonically normalized fields may be written as:

$$c^i = \frac{g_s M_p}{\sqrt{2\alpha'}} D_j^i a^j. \quad (3.9)$$

Written in terms of these fields, the potential is now:

$$V \supset \sum_k \Lambda^4 e^{-m_k} \left( 1 - \cos \left( \mathcal{Q}_k^i \frac{c^i}{M_p} \right) \right) \quad (3.10)$$

where

$$\mathcal{Q}_k^i = \frac{1}{\sqrt{2\pi g_s}} (D^{-1})_l^i q_k^l. \quad (3.11)$$

When the potential is written in the form (3.10) the field range is more easily discernible. Unfortunately, for general  $m_k$ ,  $\mathcal{Q}_k^i$  it is a bit difficult to define what is meant by ‘field range’ in the first place. If we inquire as to the range over which the potential is periodic, then it is trivial to pick a direction in field space for which the period is infinite. A better definition would be ‘the range over which the slow roll conditions are satisfied’. Sadly, this will, in general, be a quite complicated function of  $m_k$  and we are unaware of any useful expression that encodes this definition. One can, however, make progress by considering field range to be a concept defined purely in terms of the  $\mathcal{Q}_k^i$ ’s, in direct analogy with the single axion case. We thus choose a reasonable definition that is independent of  $m_k$ . For simplicity then, set all  $m_k = m$  and consider small perturbations around the minimum at  $c^i = 0$ . The potential is then:

$$V \supset \frac{1}{2} \Lambda^4 e^{-m} c \mathcal{Q}^\top \mathcal{Q} c. \quad (3.12)$$

We see then that  $\mathcal{Q}^\top \mathcal{Q}$  is proportional to the mass matrix near the origin and provides a good measure for when the potential is slowly varying. Thus, inverse eigenvalues are a good proxy for the square of the decay constant, i.e.,

$$\mathcal{Q}^\top \mathcal{Q} \tilde{c}^{(n)} = \lambda^{(n)} \tilde{c}^{(n)} \quad \Leftrightarrow \quad f_n = \frac{M_p}{\sqrt{\lambda^{(n)}}}. \quad (3.13)$$

Note that this reduces to the usual definition whenever there is a single instanton associated with each axion.<sup>7</sup>

---

<sup>7</sup>A different measure in terms of the maximum axionic field range, useful in the generic case with more instantons than axions, has been proposed in [11]. The methods presented here to constrain the parameters  $f_n$  can be easily generalized to constrain such a measure (and probably others) as well [27].

### 3.2 T-Duality

We now apply T-duality to this setup. First, we must compactify one of the non-compact directions, say,  $x^3$  along a circle of radius  $R$ .  $R$  is thus a modulus of the theory. When  $R$  takes sub-planckian values the theory is best described by IIA compactified on a radius of  $\tilde{R} = \alpha'/R$ . In this case the  $D1$  instanton maps to a massive particle represented by a  $D2$  brane wrapping the same cycle. The  $C_2$  axions become gauge fields with one index along  $x^3$ , from reduction of the type IIA  $C_3$  form along the same cycle. Other polarizations of the gauge field will come from fluctuations of  $C_4$

One might worry that Wilson lines from RR fields along the  $x^3$  direction might lead upon T-duality to a non-isotropic background about which we know too little to reliably construct a black hole solution. However, we may easily circumvent this problem by going to the limit where  $\tilde{R} \rightarrow \infty$ , in which case all anisotropies will vanish. We may thus approach the setup with  $U(1)$  gauge fields and charged scalars as studied in the original WGC.

Now our task is to relate the original parameters of type IIB to those of the final IIA theory. Standard arguments give the IIA Planck mass ( $M_p^{-2} = 8\pi G_4$ ) and dilaton as

$$\tilde{M}_P = \frac{M_{Pl_s}}{\tilde{R}} \quad \tilde{g}_s = \frac{g_s l_s}{R}. \quad (3.14)$$

For the  $U(1)$ 's, we work in a convention where the gauge coupling appears as part of the kinetic term and  $\int A \cdot j$  appears with integral coefficient. This convention implies

$$\oint A dx^3 \equiv \mu_2 \int_{\Sigma_i \times S^1} C_3. \quad (3.15)$$

Equivalently,

$$A_3^i = \frac{a_i}{(2\pi)^2 l_s^3}. \quad (3.16)$$

This determines the normalization of the gauge field. Standard compactification now gives the kinetic term:

$$S \supset \int \sqrt{-\tilde{g}} \frac{1}{4e^2} K_{ij} (F^i)_{\mu\nu} (F^j)^{\mu\nu} \quad (3.17)$$

where the parameter  $e$  is:

$$e = \frac{1}{(2\pi)^2 \tilde{M}_P \tilde{g}_s l_s}. \quad (3.18)$$

Notice that the kinetic matrix  $K_{ij}$  is not affected by the map we have performed, since the internal space remains untouched in the whole process. Finally, the masses of the charged particles are simply determined by requiring the action of an instanton be equivalent to the action of a particle running in a loop. This gives:

$$\tilde{M}_k = \frac{m_k}{2\pi \tilde{R}}. \quad (3.19)$$

We note that  $\tilde{M}_k$  ( $m_k$ ) is by definition the exact non-perturbative mass (action) of the particle (instanton). Thus, the arguments given here are immune to concerns regarding curvature and instanton corrections.

### 3.3 Black holes

Now we wish to see what constraints black hole physics in IIA places on the parameters of IIB. Unfortunately, there is a snag. Equation (3.14) tells us that when we go to the decompactification limit in IIA we are also approaching strong coupling. In this case, IIA gravity will no longer be a good description of the black hole solutions. Happily, we have another useful description in this regime; M-theory. In the M-theory description particles and gauge fields in IIA likewise lift to particles and gauge fields. The only difference is that we are now in a 5d setting where the radius of the fifth dimension is determined by  $\tilde{g}_s$ . In order to study possible constraints on our parameters we should therefore study 5d Reissner-Nordstrom black holes. For the sake of generality, we will consider black holes in arbitrary space-time dimension  $D = d + 1$ . The lagrangian truncated to the U(1)/gravity sector is:

$$S = \int d^D x \sqrt{-g_D} \left( \frac{M_D^2}{2} R_D - \frac{1}{4e_D^2} K_{ij} (F^i)^{\mu\nu} (F^j)_{\mu\nu} \right) + \int A^i \wedge *j_i^{(D)}. \quad (3.20)$$

For the case of interest, we may compactify on a circle of radius  $r_5$  to get a relationship between the M-theory parameters and the IIA parameters:

$$\begin{aligned} \frac{2\pi r_5}{e_5^2} &= \frac{1}{e^2} \\ \frac{r_5}{4G_5} &= \tilde{M}_P^2 = \frac{1}{8\pi\tilde{G}_4} \end{aligned} \quad (3.21)$$

where we have defined  $M_D^{-2} = 8\pi G_D$ . Now, a charged black hole has the solution:

$$\begin{aligned} ds^2 &= -h dt^2 + h^{-1} dr^2 + r^2 d\Omega_3^2 \\ h &= 1 - \frac{2\tilde{M}}{(d-1)M_D^2 \Omega_{d-1} r^{d-2}} + \frac{e_D^2 |n|^2}{(d-1)(d-2)M_D^2 \Omega_{d-1}^2 r^{2d-4}} \\ A_0^i &= \frac{e_D^2 n^i}{\Omega_{d-1} r^{d-2}}. \end{aligned} \quad (3.22)$$

We have introduced the notation  $|n|^2 \equiv K_{ij}^{-1} n^i n^j$ . The general extremality condition is:

$$\tilde{M} = e_D |n| \sqrt{\frac{d-1}{d-2}} M_D. \quad (3.23)$$

Specializing to 5d and translating to IIA:

$$\frac{\tilde{M}}{\tilde{M}_P} = \sqrt{\frac{3}{2}} e |n|. \quad (3.24)$$

We note that the extremality condition makes sense both in the M-theory language and in the IIA language.

We have managed to translate all the relevant quantities of the original type IIB theory with D1-instantons and  $C_2$  axions to those of M-theory with M2-particles coupled to gauge bosons. It is worth stressing that these theories are not equivalent, they are not

different descriptions of the same physical setup. Our point of view is rather that for any consistent compactification of type IIB on a given background, one should find a consistent 5d compactification of M-theory on the same internal space, with the parameters on both theories related by the dictionary just provided. If we can prove that certain parameter ranges are not allowed in any consistent compactification of M-theory, we can conclude that the translated parameters will not arise in the type IIB compactification. In this sense, we have managed to translate constraints that the standard WGC puts on 5d charged black holes to constraints on  $C_2$  axions and their field ranges in 4d type IIB models. In the following, we will use this rationale to constrain models of large field inflation.

#### 4 WGC constraints on axions

All of the original WGC arguments may be applied now to the 5d theory obtained by truncating M-theory to the U(1) sector. The spectrum of particles must thus obey a convex hull condition in order to prohibit the formation of an infinite number of remnants. The size of the ‘unit ball’ which the convex hull must contain is determined by the BPS condition obtained in (3.24). Having derived the appropriate condition in M-theory, we may now just as well use the IIA theory to discuss the resulting bound. The IIA theory contains particles with masses given by (3.19) and charges determined as  $Q_k^i = eq_k^i$ . We may thus state the convex hull condition as:

**Convex hull condition (IIA).** *Define the set of vectors  $\vec{w}_k = e\vec{q}_k \tilde{M}_P / \tilde{M}_k$ , where  $(\vec{q}_k)^i = q_k^i$ . Then, the convex hull generated by  $\vec{w}_k$  must contain the ball defined by  $K_{ij}^{-1} x^i x^j \leq \sqrt{2/3}$ .*

We now translate this to the original IIB theory using (3.11), (3.14) and (3.19).

**Convex hull condition (IIB).** *Define the convex hull generated by  $\vec{z} = \vec{Q}_k / m_k$ . Then, the convex hull generated by  $\vec{z}_k$  must contain the ball of radius  $2/\sqrt{3}$ .*

Notice that all dependence on  $R$  has dropped out of the final statement.

Although we have so far focused on axions coming from  $C_2$  in type IIB theories, the results obtained above may easily be generalized to other p-form axions. In all cases, the strategy is to relate the axion couplings/actions to particle charges/masses in a weakly coupled theory via dualities. In all cases we find that the moduli (e.g., the compact radius  $r$ ) drop out of the final condition and so we arrive at a sensible condition in the original axionic variables. In general, we may make a statement of the following form:

**General convex hull condition.** *Let  $c^i$  be a set of canonically normalized axions coming from a p-form compactified down to a  $D = d + 1$  dimensional Minkowski space. Let the instanton couplings and actions be  $Q_k^i$  and  $m_k$  respectively. Define the convex hull,  $H$ , generated by  $\vec{z} = \vec{Q}_k / m_k$ . Then,  $H$  must contain the ball of radius  $r_{(p,d)}$ , where  $r_{(p,d)}$  is a specific  $\mathcal{O}(1)$  number.*

Below, we run through some examples:

- The  $B_2$  axion obeys the same constraint as the  $C_2$  axion, being related to it via  $SL(2, \mathbb{Z})$  dualities.

- The  $C_4$  axion in IIB dualizes to  $C_5$  in IIA, which lifts to a 5-form in M-theory. The instanton dualizes to a particle in IIA and to a string in the  $5d$  lift. This string behaves just like a black hole in  $(3+1)d$  and so the usual BPS bound,  $M \geq \sqrt{2}Q$ , applies. This leads to  $r_{(4,3)} = 1$ .
- Starting from a IIA axion, we may T-dualize to IIB and then apply S-duality to reach weak coupling. One then ends up with black holes in  $(3+1)d$  and so  $r_{(\text{odd},3)} = 1$ .
- More generally, an axion in IIA compactified down to  $D = d + 1$  dimensions is related to a IIB black hole in the same number of dimensions. The black hole obeys a BPS bound:  $M \geq \sqrt{(d-1)/(d-2)}Q$ . From this we deduce that  $r_{(\text{odd},d)} = \sqrt{2(d-2)}/\sqrt{(d-1)}$ .
- Similarly, a  $C_2$  axion in  $d + 1$  dimensions relates to a black hole in  $d + 2$  dimensions, so,  $r_{(2,d)} = \sqrt{2(d-1)}/\sqrt{d}$ .

#### 4.1 General limits on axion inflation

It is now easy to show with a few examples how naive models of effective field theories of inflation with several axions fail to satisfy the convex hull condition [20, 28]. In fact figure 2 illustrates two prototypical cases. The left figure corresponds to models of N-flation (for the case  $N = 2$ ), while the right figure corresponds to models with aligned axions.

We now wish to give a more general proof that the the convex hull condition strictly forbids controlled axion inflation. The argument is as follows. First, we (arbitrarily) declare our perturbative tolerance to be  $e^{-\frac{1}{r_{(p,d)}}}$ , or,  $e^{-\sqrt{3}/2} \sim 0.421$  for the case of  $C_2$  axions. In other words, we don't allow any of the  $m_k$  to sink below  $1/r_{(p,d)}$ . With this constraint, we wish to show that the axion decay constant as defined by (3.13) is bounded by  $M_P$ . To simplify matters we may consider only the boundary of the region of perturbative control, i.e.,  $\forall k, m_k = 1/r_{(p,d)}$ . We intend to show that *even* on the boundary of the region of perturbative control one cannot obtain decay constants greater than  $M_P$ . When the convex hull condition is satisfied, increasing  $m_k$  only places more stringent restrictions on  $\vec{Q}_k$  and so we claim that proving the claim on the boundary is sufficient.

On the boundary of the perturbative region, the convex hull condition simplifies to the requirement that the convex hull of  $\vec{Q}_k$  contains the unit ball. We thus need to prove that under this condition *all* of the eigenvalues  $\lambda^{(n)}$  of  $\mathcal{Q}^\top \mathcal{Q}$  are greater than 1. This will immediately imply that  $f_n = M_P/\sqrt{\lambda^{(n)}}$  is sub-Planckian.

In order to obtain the desired result, it is useful to restate the convex hull condition as follows: *The convex hull of  $\vec{Q}_k$  contains the unit ball  $\iff$  for any vector  $\vec{c}$  such that  $|\vec{c}| = 1$  there is a unique set of numbers  $\alpha_k$  and  $\rho$  such that  $\sum_k \alpha_k = 1$ ,  $\rho > 1$ , and  $\sum_k \alpha_k \vec{Q}_k = \rho \vec{c}$ .*

It is important to note that the numbers  $\alpha_k, \rho$  defined above are unique. For a square charge matrix this may be seen immediately from equation counting. With rectangular matrices, we need a more refined argument. Suppose, on the contrary, that the parameters  $\alpha_k$  and  $\rho$  where not unique. We may then reach the same charge vector in two different ways, i.e.:

$$\sum_k^n \alpha_k \vec{z}_k = \rho \vec{c}, \quad \sum_j^m \tilde{\alpha}_j \vec{z}_j = \tilde{\rho} \vec{c} \tag{4.1}$$

where  $\vec{z}$  and  $\vec{\tilde{z}}$  are generators of the convex hull. Without loss of generality, we assume  $\tilde{\rho} \leq \rho$ . At least one of the  $\vec{\tilde{z}}_j$  vectors must be distinct from all the  $\vec{z}_k$ 's. Let us suppose that this is  $\vec{\tilde{z}}_m$ . Using the equations above, we may algebraically solve for  $\vec{\tilde{z}}_m$  by eliminating  $\vec{c}$  between the two equations. Then, it is straightforward to check that  $\vec{\tilde{z}}_m$  is actually *inside* the convex hull generated by the other vectors. (If  $\rho \leq \tilde{\rho}$  then we must choose  $\vec{z}_n$  to eliminate instead.) We can safely discard this particle from the spectrum, and the convex hull of the system will not be affected. Proceeding in this manner, we may reach a minimal convex hull such that all vectors have a unique representation as stated above.<sup>8</sup> By throwing away these extra charge vectors we are simply focusing on a certain set of instanton corrections and we will be showing that, if the convex hull contains the unit ball, large instanton corrections may be encountered already within this subset.

Now, let us suppose that  $\vec{c}$  is a normalized eigenvector of  $Q^\top Q$  with eigenvalue  $\lambda$ . Define  $\tilde{\alpha}_k = Q_k^i c^i$ . Dotting both sides of the eigenvalue equation with  $\vec{c}$  one immediately finds:

$$\lambda = \sum_k \tilde{\alpha}_k^2. \tag{4.2}$$

Now, define,  $\alpha_k$  as:

$$\alpha_k = \frac{\tilde{\alpha}_k}{\sum_j \tilde{\alpha}_j}. \tag{4.3}$$

These sum to one by construction. If we now divide both sides of the eigenvalue equation by  $\sum_k \tilde{\alpha}_k$  and use the equations above we get:

$$\sum \vec{Q}_k \alpha_k = \left( \frac{\sum_k \tilde{\alpha}_k^2}{\sum_k \tilde{\alpha}_k} \right) \vec{c}. \tag{4.4}$$

The uniqueness property mentioned above now implies that:

$$\left( \frac{\sum_k \tilde{\alpha}_k^2}{\sum_k \tilde{\alpha}_k} \right) = \rho > 1. \tag{4.5}$$

From this it follows that at least one of the  $\tilde{\alpha}_k$ 's are greater than one, otherwise we would have a contradiction. We thus see from (4.2) that  $\lambda$  must also be greater than one, which is what we were trying to show. This implies that whenever we are in the regime of perturbative control, all decay constants satisfy:

$$f_n \leq M_p. \tag{4.6}$$

One might ask how to interpret these results in light of works such as [11], which seem to be in conflict with the WGC. If one does not give up the idea that WGC must be satisfied in string theory, we are predicting that these constructions will suffer corrections that bring them into line with the WGC. The models in fact include small (often sub-Planckian) volumes on which large corrections could arise. In particular, even if an instanton brane wrapping such a cycle does not contribute to the superpotential it can show up in the *scalar* potential and induce large deviations. Other sources of concern would be *KK*

---

<sup>8</sup>The minimal convex hull need not correspond to a square matrix.

and  $\alpha'$  corrections, one-loop determinants and the backreaction of instanton branes, which is particularly significant in models with small cycles. If such corrections are ultimately absent, this would be an indication that the instantons considered do not satisfy the WGC. The conclusion would be that either the WGC is violated (in both its mild and strong forms) or that there exists some other instantons not arising from Euclidean D-branes (or gaugino condensates) that satisfy the WGC constraints, perhaps realizing the loophole described in the following section. The origin of such instantons would be nevertheless extremely mysterious based on our current understanding of string theory.

## 4.2 A possible loophole

From the above discussion it seems clear that the models with several axions so far proposed in the literature are incompatible with the WGC, both in its strong and its mild formulations. As we describe next, the mild form could still allow for a certain enhancement of the axion decay constant. It is nevertheless worth mentioning beforehand that in string theory the strong version of the WGC seems to hold, and the mechanism we discuss below has not yet found its realization in the literature.<sup>9</sup>

For simplicity we will restrict our discussion to the case of a single U(1) group (a single axion on the dual side), the generalization to more general cases should be simple. For the argument to hold, we need to assume that there are several families of particles (instantons) charged under the U(1) group. As we mentioned in section 2, the particle  $(M, Q)$  that satisfies the constraint  $Q M_p/M \geq 1$  does not need to be the lightest one in the spectrum if only the mild version of the WGC holds. Let us assume for example that there is another particle  $(m, q)$ , which is lighter ( $M > m$ ), has a smaller charge, ( $Q = kq$  for some  $k \in \mathbb{Z}$ ), and does not satisfy the WGC bound, i.e.,  $qM_p/m < 1$ . Although this situation is possible, it leads to the presence of at least  $k$  stable bound states. Hence  $k$  cannot be made parametrically large without running into problems with the CEB.

Translated to the T-dual language of instantons and axions, one would have two instantons coupled to a single axion. The dominant contribution comes from  $(m, q)$ , which is suppressed by  $e^{-m}$  and has the longer periodicity  $f \sim 1/q$ . The other instanton has a suppression of  $e^{-M} \ll e^{-m}$  and a shorter modulation  $F \sim f/k$ . The non-perturbative potential for the axions has the form:

$$V = \Lambda_1^4 e^{-m} \left[ 1 - \cos\left(\frac{\phi}{f}\right) \right] + \Lambda_2^4 e^{-M} \left[ 1 - \cos\left(\frac{k\phi}{f}\right) \right]. \quad (4.7)$$

The second instanton contribution is the one that ensures that the mild version of the WGC is satisfied. Nevertheless, if  $m < M$ , this term can be highly suppressed and negligible in comparison with the first one, which can be used to generate the inflaton potential. Notice that  $m$  still needs to be bigger than  $\sim 1$  to suppress higher order corrections. The requirement of trans-Planckian field range  $f > M_p \implies f \cdot m > M_p$ , together with the WGC constraint satisfied by the second instanton leads to the (conservative) bound  $M > m > M/k$ . Since  $k$  cannot be arbitrarily large, the enhancement of the decay constants cannot be made parametrically large.

---

<sup>9</sup>A similar idea has been considered in a particular model of axion alignment in [20].



This seems to be an interesting possibility by which the WGC allows for a trans-Planckian decay constant, although it only holds if the strong-WGC is violated. The key point here is to assume that the convex hull condition is fulfilled by a set of instantons whose contributions to the potential are negligible in comparison with other instantons, which by themselves would not satisfy the WGC. Hence, for every proposal that generates a long-range inflaton potential, naively violating the WGC, one would need to ensure the existence of negligible instantons whose actions and charges satisfy the convex hull condition. This is a rather trivial step in field theory, but is highly non-trivial in explicit string embeddings. In fact, we are not aware of any example which violates the strong-WGC.

## 5 Torsion cycles and generalized WGC

We have seen that the weak gravity conjecture prevents us from having large field ranges while keeping higher order instanton corrections under control. These higher instantons come from multi-wrapped branes living in  $H_p(X, \mathbb{Z})$ . One natural idea is to try and eliminate higher instantons from the spectrum by considering torsion cycles. Suppose  $X$  is a 6d compact manifold with:

$$H_r(X, \mathbb{Z}) = \mathbb{Z}^{b_r} \oplus \mathbb{Z}_{k_1} \oplus \mathbb{Z}_{k_2} \oplus \dots \oplus \mathbb{Z}_{k_n}. \quad (5.1)$$

Consider an instanton represented by a  $D_{r-1}$  brane wrapping a torsion cycle in such a manifold. Such an instanton cannot generate a potential for any massless closed string axion. To see this let  $\Sigma$  be a torsion cycle such that  $k\Sigma = \partial\Omega$  and suppose that  $w_r$  is a harmonic (hence closed)  $r$ -form ‘axion’ for which we would like to generate a potential. Then:

$$k \int_{\Sigma} w_r = \int_{k\Sigma} w_r = \int_{\partial\Omega} w_r = \int_{\Omega} dw_r = 0. \quad (5.2)$$

Thus,  $\int_{\Sigma} w_r = 0$ . In general, the massless closed string RR field will not see the instanton at all [29]. The appropriate generalization must therefore begin with massive RR fields. This leads us naturally to the setup discussed in [18], which provided a concrete realization of models proposed in [30, 31]. As discussed in this reference, the general theory of torsion cycles tells us that:

$$TorH_p(X, \mathbb{Z}) = TorH_{5-p}(X, \mathbb{Z}) = TorH^{p+1}(X, \mathbb{Z}) = TorH^{6-p}(X, \mathbb{Z}). \quad (5.3)$$

Let us suppose that we have a  $C_p$  RR form and let us call the fields obtained by reduction along the  $C_p$  form ‘electric’. The dual fields obtained by reduction of the  $C_{8-p}$  form will be labelled as ‘magnetic’. The cohomological relations above guarantee us a non-trivial element of  $TorH^{p+1}(X, \mathbb{Z})$ . Call this element  $\omega_{p+1}$ . Then, for some  $k \in \mathbb{Z}$  and some  $p$ -form  $\eta_p$ :

$$d\eta_p = k\omega_{p+1}. \quad (5.4)$$

We may now define a massive inflaton field by reducing  $C_p$  along  $\eta_p$ :

$$C_p \equiv \phi\eta_p, \quad F_{p+1} = d\phi \wedge \eta_p + k\phi\omega_{p+1}. \quad (5.5)$$

Dimensional reduction of the kinetic terms reproduces the monodromy model in [32, 33]. We also have domain walls and strings in the 4d effective theory. Again, this may be seen as a consequence of the special relations between the homology groups for torsion cycles. A  $D_{7-p}$  brane wrapped on a  $TorH_{5-p}(X, \mathbb{Z})$  cycle (which is guaranteed to exist by (5.3)) will give a domain wall. Moreover, if  $\Sigma_{5-p} \in TorH_{5-p}(X, \mathbb{Z})$ , then  $k\Sigma_{5-p} = \partial\Omega_{6-p}$  for some  $\Omega_{6-p}$ . The  $D_{7-p}$  brane wrapping this  $\Omega_{6-p}$  will give a string, upon which  $k$  copies of the domain wall may end.

The 4d effective potential for the axion,  $\phi$ , takes the form

$$V(\phi) = \frac{1}{2}(q + \mu\phi)^2 \quad (5.6)$$

where  $\mu$  now contributes to the mass of the axion. The parameter  $q$  has dynamics sourced by the domain wall and so will jump in discrete steps across one of these domain walls. It is also, then, quantized in terms of the brane charge  $e$  such that  $q = ne$  for some  $n \in \mathbb{Z}$ . This gives the potential a branching structure with discrete quadratic branches labelled by  $n$  [33].

A potential problem with this scenario is that the inflaton could tunnel between different branches of the potential thus destabilizing the vacua on timescales too short to support inflation. This tunnelling event is marked with the nucleation of a domain wall separating regions of space in different branches of the potential. In order to maintain the monodromy, the nucleation rate of these domain walls would need to be suppressed. The tunnelling amplitude is given in terms of the tension of the domain wall,  $\sigma$ , and the potential difference between the two sides of the domain wall,  $\Delta V$ , as [34, 35]

$$P \sim e^{-S_{CD}} \quad (5.7)$$

$$S_{CD} = \frac{B_0}{(1 + A^2)^2}$$

$$B_0 = \frac{27\pi^2\sigma^4}{2(\Delta V)^3}, \quad A = \frac{\sqrt{3}\sigma}{\sqrt{2\Delta V}M_p}.$$

We parameterize the vacuum energy by  $V = \beta M_p^4$  and assume that  $V \sim \Delta V$ . If we impose the standard hierarchy of scales;  $H < V^{1/4} \sim \Delta V^{1/4} < M_{KK} < M_s < M_p$ , then one may estimate that [36]:

$$S_{CD} \sim g_s^8 \left( \frac{M_{KK}}{M_s} \right)^{44-4p} \left( \frac{1}{\beta} \right)^3. \quad (5.8)$$

Note that  $M_{KK} < M_s$  and so maintaining even a modest hierarchy puts a very stringent lower bound on  $g_s$ . Thus, it seems difficult to find a parametrically large window of perturbative control where tunneling is suppressed. To put this problem in the context of the WGC, let us suppose that WGC is strongly violated. Notice that  $\Delta V \sim q^2$ , and so  $A \sim \sigma/qM_p$ . Thus, strongly violating the WGC is equivalent to the statement that  $A \gg 1$ . In this regime:

$$S_{CD} \sim \frac{1}{\beta} \gg 1. \quad (5.9)$$

We therefore see that tunneling *is* parametrically suppressed when WGC is violated and only marginally suppressed when it is satisfied. We suspect that this statement may be made more sharp, although this would require a reassessment of the validity of the Coleman-de Luccia formula (5.7) for non-metastable vacua [36].

## 6 Tension with holography

So far, we have shown that models with large axion decay constants are related through T-duality to models producing black-hole remnants. Should there not be some sort of pathology even before performing T-duality? In fact, such pathologies have been noted in other contexts. In [37], it was found that constructing a certain class of Euclidean worm-holes in the supergravity limit of string theory required one to have large axionic field ranges. Since such large field ranges are unachievable with single axions, the authors resorted to a mechanism akin to N-flation in order to achieve the desired range. These wormholes are non-singular, yet are strong curvature solutions and therefore susceptible to  $\alpha'$  corrections. Indeed, if we were to take these solutions at face value then they would violate the principle of locality and also cluster decomposition whenever there is a good field theory dual. Thus, we should rigorously exclude such wormholes on first principles. It would be interesting to verify in more generality that large axion field ranges are directly tied to non-locality and may thus be excluded out of this (sacred) principle.

One might also suspect that the ultimate sickness of theories with strong gravity is that they violate the covariant entropy bound (CEB) [38] and that remnants are just a manifestation of this. We can look for conflict with CEB in models with large decay constants as well. In fact, it has been argued that general inflationary models may be in tension with the holographic principle [39]. To see this, note that during inflation, the entropy of a region should be bounded by the corresponding ‘light sheet’, whose area goes like  $a^2 \sim e^{2Ht}$ . However, upon reheating, we find that semiclassical physics gives an entropy on a spatial slice proportional to the volume, or  $a^3 \sim e^{3Ht}$ . Clearly, the entropy measured in this way cannot be bound by the area while achieving a large number of e-folds.

There is, however, a subtlety in this argument. The covariant entropy bound, strictly speaking, should apply to the entropy of the light cone itself, rather than to the matter system on a spatial slice. When the system is bounded by future directed light cones (e.g., when the strong energy condition is satisfied) then one may also obtain a bound on the entropy of a spatial slice. On the other hand, in an inflationary process we have only past directed light cones and so we must try and confront the light-cone entropy directly. Indeed, just such an attempt was made in [40]. In this work, the author found that in order to satisfy CEB (with no kinetic mixing) the axion decay constants must satisfy:

$$\sum_i f_i^2 \leq \mathcal{O}(1)M_p^2. \quad (6.1)$$

This just says that even when moving along the diagonal in axion field space, one still finds a decay constant bounded by  $M_p$ , in contrast with N-flation scenarios. This is exactly the same result, (4.6), we have found from the WGC. It thus seems that nature’s stricture

against large decay constants is intimately tied to the principle of holography. One could further argue that this provides some evidence for the strong-WGC above and beyond that which supports the mild-WGC. If only the mild version were valid, then one could take advantage of the loophole discussed in section 4.2 and construct inflationary theories which are in violation of the CEB. Therefore, it seems that the CEB may actually require the strong WGC in order to be consistent.

Finally, independent of the CEB, it also seems to be rather difficult to understand how the entropy at reheating may be generated in a unitary manner within a semiclassical description of inflation. Suppose, for instance, that the universe began inflating in a pure state. In this case, the observed entropy upon reheating must be due entirely to correlation entropy between local and very distant regions of the universe. This would seem to require non-local correlations amongst particles which were never in causal contact, in direct analogy with the black hole information paradox. It would be interesting to see if this observation could be put on firmer footing.

## 7 Conclusions and speculations

In this note we have shown that many models of axion inflation are in violation of the weak gravity conjecture and are therefore unlikely to have a consistent embedding within string theory, or any other consistent theory of quantum gravity. The ultimate origin of the weak gravity conjecture may be found within the covariant entropy bound, which codifies the holographic nature of quantum gravity. However, if we really take the covariant entropy bound seriously, we find tension not just with axion inflation, but with *many* inflation models.

This tension would seem to indicate that there is some basic gap in our understanding of inflation, analogous to the information problem surrounding black holes. If we must indeed hold the covariant entropy bound sacrosanct, one possibility is to replace inflationary models with something analogous to a fuzzball structure [41], i.e. a ‘fuzzbang’. In this scenario, the universe would actually satisfy the no boundary proposal [42] and each microstate of the universe would correspond to a particular internal geometry at the time of reheating. If this were true, then inflation would be essentially a mirage which appears to be valid in the low-energy effective field theory description. Again, this is analogous to the fuzzball paradigm for blackholes, where the geometry may appear to be Schwarzschild outside the horizon but the interior region is simply absent. One may also speculate along the lines of AMPS [43] that semiclassical physics breaks down upon reheating, thus thwarting one’s attempt to formulate a sharp paradox with unitarity.

Thus, from the perspective of holography, the Big Bang and the Firewall appear to share some interesting similarities. On the other hand, a truly explicit parametrically controlled string embedding of inflation would tell us that holography is overstepping its bounds in this setting. We have, in fact shown that there are several possible caveats which would allow one to evade some of the constraints of the WGC, e.g. if the strong or the discrete versions of the WGC are violated. As with any “no-go theorem”, the most interesting thing to understand is how to get around it. Moreover, the tension between

holography and large field inflation models is perhaps pointing us to a deeper understanding of quantum gravity and deserves further study.

## Acknowledgments

We would like to thank Frederik Denef, Daniel Junghans, Fernando Marchesano, Thomas van Riet, Hagen Triendl, Angel Uranga and Timm Wrase for useful discussions and suggestions. We would also like thank the Michigan Center for Theoretical Physics and the organizers and participants of the workshop “String/M-theory Compactifications and Moduli Stabilization” where the results of this work were originally presented. This work is supported in part by the DOE grant DE-FG-02-95ER40896 and the HKRGC grants HUKST4/CRF/13G, 604231 and 16304414.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] A.D. Linde, *Eternally Existing Selfreproducing Chaotic Inflationary Universe*, *Phys. Lett. B* **175** (1986) 395 [[INSPIRE](#)].
- [2] K. Freese, J.A. Frieman and A.V. Olinto, *Natural inflation with pseudo Nambu-Goldstone bosons*, *Phys. Rev. Lett.* **65** (1990) 3233 [[INSPIRE](#)].
- [3] D. Baumann and L. McAllister, *Inflation and String Theory*, [arXiv:1404.2601](#) [[INSPIRE](#)].
- [4] T. Banks, M. Dine, P.J. Fox and E. Gorbatov, *On the possibility of large axion decay constants*, *JCAP* **06** (2003) 001 [[hep-th/0303252](#)] [[INSPIRE](#)].
- [5] K. Choi, H. Kim and S. Yun, *Natural inflation with multiple sub-Planckian axions*, *Phys. Rev. D* **90** (2014) 023545 [[arXiv:1404.6209](#)] [[INSPIRE](#)].
- [6] J.E. Kim, H.P. Nilles and M. Peloso, *Completing natural inflation*, *JCAP* **01** (2005) 005 [[hep-ph/0409138](#)] [[INSPIRE](#)].
- [7] S. Dimopoulos, S. Kachru, J. McGreevy and J.G. Wacker, *N-flation*, *JCAP* **08** (2008) 003 [[hep-th/0507205](#)] [[INSPIRE](#)].
- [8] C. Long, L. McAllister and P. McGuirk, *Aligned Natural Inflation in String Theory*, *Phys. Rev. D* **90** (2014) 023501 [[arXiv:1404.7852](#)] [[INSPIRE](#)].
- [9] G. Shiu, W. Staessens and F. Ye, *Widening the Axion Window via Kinetic and Stückelberg Mixings*, [arXiv:1503.01015](#) [[INSPIRE](#)].
- [10] G. Shiu, W. Staessens and F. Ye, *Large Field Inflation from Axion Mixing*, *JHEP* **06** (2015) 026 [[arXiv:1503.02965](#)] [[INSPIRE](#)].
- [11] T.C. Bachlechner, C. Long and L. McAllister, *Planckian Axions in String Theory*, [arXiv:1412.1093](#) [[INSPIRE](#)].
- [12] F. Denef, M.R. Douglas, B. Florea, A. Grassi and S. Kachru, *Fixing all moduli in a simple F-theory compactification*, *Adv. Theor. Math. Phys.* **9** (2005) 861 [[hep-th/0503124](#)] [[INSPIRE](#)].

- [13] C. Vafa, *The String landscape and the swampland*, [hep-th/0509212](#) [INSPIRE].
- [14] H. Ooguri and C. Vafa, *On the Geometry of the String Landscape and the Swampland*, *Nucl. Phys. B* **766** (2007) 21 [[hep-th/0605264](#)] [INSPIRE].
- [15] N. Arkani-Hamed, L. Motl, A. Nicolis and C. Vafa, *The String landscape, black holes and gravity as the weakest force*, *JHEP* **06** (2007) 060 [[hep-th/0601001](#)] [INSPIRE].
- [16] C. Cheung and G.N. Remmen, *Naturalness and the Weak Gravity Conjecture*, *Phys. Rev. Lett.* **113** (2014) 051601 [[arXiv:1402.2287](#)] [INSPIRE].
- [17] A. de la Fuente, P. Saraswat and R. Sundrum, *Natural Inflation and Quantum Gravity*, *Phys. Rev. Lett.* **114** (2015) 151303 [[arXiv:1412.3457](#)] [INSPIRE].
- [18] F. Marchesano, G. Shiu and A.M. Uranga, *F-term Axion Monodromy Inflation*, *JHEP* **09** (2014) 184 [[arXiv:1404.3040](#)] [INSPIRE].
- [19] M. Montero, A.M. Uranga and I. Valenzuela, *Transplanckian axions!?*, *JHEP* **08** (2015) 032 [[arXiv:1503.03886](#)] [INSPIRE].
- [20] T. Rudelius, *Constraints on Axion Inflation from the Weak Gravity Conjecture*, *JCAP* **09** (2015) 020 [[arXiv:1503.00795](#)] [INSPIRE].
- [21] L. Susskind, *Trouble for remnants*, [hep-th/9501106](#) [INSPIRE].
- [22] J. Brown, W. Cottrell, G. Shiu and P. Soler, work in progress.
- [23] P. Svrček and E. Witten, *Axions In String Theory*, *JHEP* **06** (2006) 051 [[hep-th/0605206](#)] [INSPIRE].
- [24] T. Banks and N. Seiberg, *Symmetries and Strings in Field Theory and Gravity*, *Phys. Rev. D* **83** (2011) 084019 [[arXiv:1011.5120](#)] [INSPIRE].
- [25] J. Polchinski, *String Theory. Volumes 1. Superstring Theory and Beyond*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (1998).
- [26] J. Polchinski, *String Theory. Volumes 2. Superstring Theory and Beyond*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (1998).
- [27] J. Brown, W. Cottrell, G. Shiu and P. Soler, *On Axionic Field Ranges, Loopholes and the Weak Gravity Conjecture*, [arXiv:1504.00659](#) [INSPIRE].
- [28] T. Rudelius, *On the Possibility of Large Axion Moduli Spaces*, *JCAP* **04** (2015) 049 [[arXiv:1409.5793](#)] [INSPIRE].
- [29] P.G. Camara, L.E. Ibáñez and F. Marchesano, *RR photons*, *JHEP* **09** (2011) 110 [[arXiv:1106.0060](#)] [INSPIRE].
- [30] L. McAllister, E. Silverstein and A. Westphal, *Gravity Waves and Linear Inflation from Axion Monodromy*, *Phys. Rev. D* **82** (2010) 046003 [[arXiv:0808.0706](#)] [INSPIRE].
- [31] E. Silverstein and A. Westphal, *Monodromy in the CMB: Gravity Waves and String Inflation*, *Phys. Rev. D* **78** (2008) 106003 [[arXiv:0803.3085](#)] [INSPIRE].
- [32] N. Kaloper and L. Sorbo, *A Natural Framework for Chaotic Inflation*, *Phys. Rev. Lett.* **102** (2009) 121301 [[arXiv:0811.1989](#)] [INSPIRE].
- [33] N. Kaloper, A. Lawrence and L. Sorbo, *An Ignoble Approach to Large Field Inflation*, *JCAP* **03** (2011) 023 [[arXiv:1101.0026](#)] [INSPIRE].

- [34] S.R. Coleman, *The Fate of the False Vacuum. 1. Semiclassical Theory*, *Phys. Rev. D* **15** (1977) 2929 [Erratum *ibid.* **D 16** (1977) 1248] [[INSPIRE](#)].
- [35] S.R. Coleman and F. De Luccia, *Gravitational Effects on and of Vacuum Decay*, *Phys. Rev. D* **21** (1980) 3305 [[INSPIRE](#)].
- [36] J. Brown, I. Garcia-Etxebarria, F. Marchesano and G. Shiu, work in progress.
- [37] N. Arkani-Hamed, J. Orgera and J. Polchinski, *Euclidean wormholes in string theory*, *JHEP* **12** (2007) 018 [[arXiv:0705.2768](#)] [[INSPIRE](#)].
- [38] R. Bousso, *The Holographic principle*, *Rev. Mod. Phys.* **74** (2002) 825 [[hep-th/0203101](#)] [[INSPIRE](#)].
- [39] N. Kaloper and A.D. Linde, *Cosmology versus holography*, *Phys. Rev. D* **60** (1999) 103509 [[hep-th/9904120](#)] [[INSPIRE](#)].
- [40] J.P. Conlon, *Quantum Gravity Constraints on Inflation*, *JCAP* **09** (2012) 019 [[arXiv:1203.5476](#)] [[INSPIRE](#)].
- [41] S.D. Mathur, *The Fuzzball proposal for black holes: An Elementary review*, *Fortsch. Phys.* **53** (2005) 793 [[hep-th/0502050](#)] [[INSPIRE](#)].
- [42] J.B. Hartle and S.W. Hawking, *Wave Function of the Universe*, *Phys. Rev. D* **28** (1983) 2960 [[INSPIRE](#)].
- [43] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **02** (2013) 062 [[arXiv:1207.3123](#)] [[INSPIRE](#)].