

Bayesian inference on Calabi–Yau moduli spaces and the axiverse: experimental data meets string theory

Mudit Jain ^{1,2,*} Elijah Sheridan ^{3,†} David J. E. Marsh ^{1,‡} Elli
Heyes ^{4,§} Keir K. Rogers ^{1,5,¶} and Andreas Schachner ^{3,6,**}

¹*Physics Department, King’s College London, Strand, London WC2R 2LS, UK*

²*Department of Physics and Astronomy, Rice University, Houston, Texas 77005, U.S.A.*

³*Department of Physics, Cornell University, Ithaca, NY 14853, USA*

⁴*Abdus Salam Centre for Theoretical Physics, Imperial College London, London SW7 2AZ, UK*

⁵*Department of Physics, Imperial College London,*

Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK

⁶*ASC for Theoretical Physics, LMU Munich, 80333 Munich, Germany*

We develop tools of Bayesian inference on the moduli space of Calabi–Yau (CY) manifolds. We sample from the invariant Weil–Petersson (WP) measure using Markov Chain Monte Carlo and normalising flows on Kähler moduli space with dimension up to $h^{1,1} = 30$, and present results on the spectrum of the CY volume and properties of divisors when the measure is restricted in physically meaningful ways. We furthermore present a theory-informed prior on axion masses and decay constants (m_a, f_a) marginalised over the WP measure for *all* inequivalent CYs constructable from the Kreuzer–Skarke database with $h^{1,1} \leq 5$. We then impose likelihoods based on axion physics. We demonstrate how detection of a relatively heavy QCD axion at small $h^{1,1}$, e.g. by ADMX, provides detailed information about CY geometry and topology. Finally, we compute a full forward model incorporating likelihoods from the cosmic microwave background and Lyman-alpha forest and find the maximum posterior probability region on the moduli space of a given CY favoured by a resolution of the tension in these data by an ultralight axion composing $\mathcal{O}(1\%)$ of the dark matter. This demonstration serves as a blueprint for future statistical analyses within string phenomenology. 

I. INTRODUCTION

String theory compactified on Calabi–Yau (CY) orientifolds is one of the best candidate theories to describe quantum gravity and particle physics in a unified manner [1]. Recent years have seen great advances in our ability to study CY manifolds, and the resulting physics, numerically [2]. Axions [3–5] are a ubiquitous feature of string theory compactifications [6–8], known as the *axiverse* [9–11], and the hunt for axions in the laboratory and in cosmology is currently undergoing intense progress [12–14]. Advances in the explicit constructions of CYs have led to similar progress in our detailed understanding of the axiverse [15–25]. The basic C_4 axion effective field theory (EFT) in type IIB string theory is well-understood: its kinetic matrix is polynomial in subvariety volumes and its potential is dominated by exponentials of divisor volumes. Divisor volumes are not, however, independent free parameters: in explicit CY geometries they are often strongly correlated, meaning that not all things possible from an EFT perspective are realised in string theory at the ensemble level. Outside specific CY models and beyond type IIB, other important recent progress includes Refs. [26–29].

The progress on axions in explicit CY ensembles has identified important general principles. Axion masses are approximately log-uniformly distributed and axion decay constants are approximately log-normally distributed. The scale of the largest mass and mean decay constant are set by the Hodge number $h^{1,1}$ of the CY (which also sets the number of axions), with both decreasing as $h^{1,1}$ increases [19]. To date, most of this research has been performed at a canonical but specific region in the moduli space: namely, along the ray generated by the “tip of the stretched Kähler cone” (see Section II) and relatively close to the origin (e.g., [15, 17–21]). In this work we lay the groundwork for a systematic and statistically rigorous approach to studying C_4 axion models in the full stretched Kähler cone.

A central challenge arising from such a program of study is the problem of *moduli*. The Kähler and complex structure moduli are scalar fields representing geometric deformations of the CY that preserve the defining property of Ricci flatness [30]. They are massless in the classical theory, but are believed to be stabilised, i.e., given masses and vacuum expectation values (vevs), by quantum effects (see Ref. [31] for a review). There has also been significant progress in understanding moduli stabilisation in explicit examples [32–38], yet even the most advanced models are still far from answering key questions about supersymmetry breaking and the cosmological constant in phenomenologically acceptable ways.

The values of the stabilised moduli fields (moduli vevs) can be seen as constant parameters of the four-dimensional EFT. The physics giving rise to moduli sta-

* mudit.jain@rice.edu

† es888@cornell.edu

‡ david.j.marsh@kcl.ac.uk

§ e.heyse@imperial.ac.uk

¶ keir.rogers@kcl.ac.uk

** as3475@cornell.edu

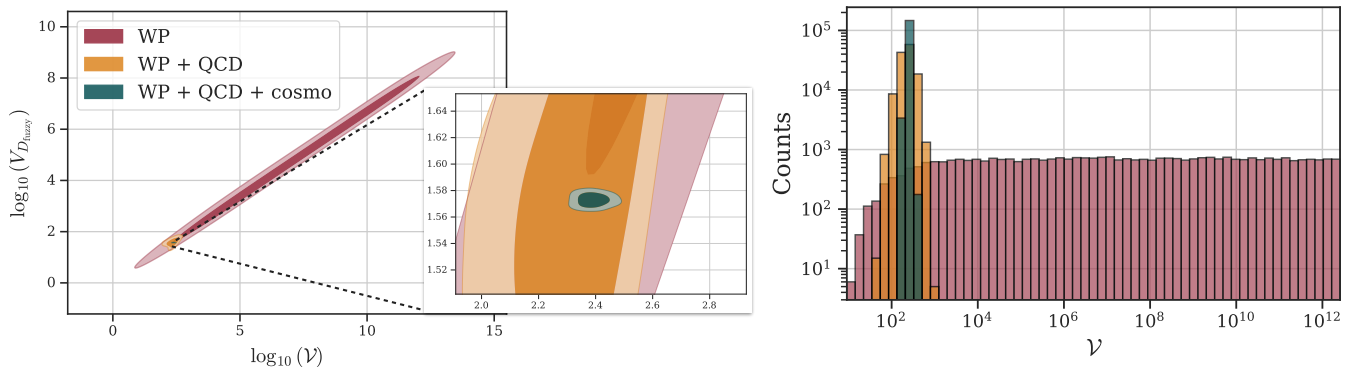


FIG. 1. Joint posterior distribution of $\log_{10} \mathcal{V}$ and $\log_{10}(\tau_{\text{fuzzy}})$ obtained by sampling the Kähler moduli of a Calabi–Yau with $h^{1,1} = 7$, where τ_{fuzzy} is the volume of the prime toric divisor hosting the instanton associated with the ultralight (fuzzy) axion. We show three cases: (i) the Weil–Petersson prior alone (with IR/UV cutoffs from EFT control and the KK scale), (ii) the Weil–Petersson prior supplemented by the requirement that at least one prime toric divisor attains a volume suitable for a QCD-like gauge coupling, and (iii) the same prior further weighted by CMB and Lyman- α likelihoods. The inset zooms into the small region where the cosmology-selected posterior concentrates. The right panel shows the corresponding marginal posterior of $\log_{10} \mathcal{V}$ for all three cases, illustrating how cosmological data lead to a sharply peaked posterior and a strongly localised region in moduli space.

bilisation, such as fluxes [39] and D-brane instantons [40] is, however, fundamentally discrete, meaning that the true spectrum of allowed vacua must also be discrete. Nevertheless, it has long been appreciated that the available choices are so numerous that this discrete spectrum can effectively be treated as continuous, or more precisely, as forming a dense “discretuum” [41, 42]. Following this reasoning, we approximate the vacuum expectation values of the Kähler moduli, $\langle T_i \rangle$, as continuous variables. There exist physical effects that may contribute to moduli stabilisation but cannot currently be computed with reliability, i.e., known unknowns, such as loop corrections to the Kähler potential; see, however, [43, 44]. In addition, there may be mechanisms for moduli stabilisation that remain entirely undiscovered, i.e., unknown unknowns.

Given the formidable challenge of determining where acceptable vacua populate moduli space, it may seem fruitless to address further questions of physics. In this work, however, we take the “modular approach” to phenomenology [31]. Although we are presently unable to solve the problem of vacuum distributions across moduli space for large ensembles of CYs, we can nevertheless compute certain features of low energy EFTs throughout moduli space. Our aim is therefore to *statistically sample* the physical properties of string compactifications. Provided that such sampling is computationally efficient, as well as *fair and unbiased* (in the Bayesian sense), moduli stabilisation may then be imposed *ex post facto*, allowing the resulting phenomenological consequences to be inferred.

We thus set aside the problem of explicit moduli stabilisation and instead endeavor to map out Kähler moduli space and the resulting physics statistically, for which we introduce tools of Bayesian inference and machine learn-

ing.¹ The *Weil–Petersson* (WP) measure, induced by the WP metric for Kähler moduli, is the natural measure on Kähler moduli space (see Section C for some details). This metric exhibits several useful features for our purposes. For example, in addition to being diffeomorphism invariant, the WP measure is also scale invariant, as expected for an uninformative prior in the absence of a preferred scale, as parametrised by e.g. the CY volume $\mathcal{V}$ (which could later be set by, for example, a particular moduli stabilisation scheme). The WP measure is thus similar to the Jeffreys prior, and is therefore attractive from a purely Bayesian perspective. Consequently, in this work we develop efficient methods for sampling the WP measure. Using the sampling techniques of Markov Chain Monte Carlo (MCMC) and normalising flows (NF), our goal is to compute axion physics observables, compare them with experimental or observational likelihoods, and perform Bayesian inference.

Since we are concerned with phenomenology, we impose extra restrictions on top of the WP prior. In particular, we impose *stretching*: i.e., the constraint that all holomorphic submanifolds have volume larger than unity in string units, thereby controlling the EFT. We also impose that the Kaluza-Klein (KK) scale must be above some physical energy scale to ensure consistency with observations, which implies a maximum CY volume. These two EFT restrictions break WP scale invariance, but have the welcome property of rendering the WP measure a proper (normalisable) prior. The shape of the region in moduli space satisfying our upper and lower bounds on

¹ We note that the complex structure moduli, at our level of approximation, do not affect the physics likelihoods studied here, see however Refs. [45–48] for progress in this direction.

the overall volume as well as our stretching conditions is sensitive to the topology of the CY manifold (such as its fibrations, see [Section II.C](#)). In particular, this region features limits near the boundary of the Kähler cone where divisors with particular topological properties have distinctive volume spectra (see [Section IV](#)), which we will connect to axion physics in [Section V.B](#). This also extends prior work relating divisor topology to statistics of volumes [\[23\]](#).

We assume that the Standard Model gauge group, and in particular quantum chromodynamics (QCD), is realised by a stack of D7-branes on some prime toric divisor — distinguished divisors on toric hypersurface Calabi–Yau threefolds — and we demand that the volume distribution of this divisor is relatively sharply peaked at 40. This ensures that the running of the SM gauge couplings to the KK scale is not drastically altered compared to that implied by the observed particle content [\[49\]](#). Allowing any such choice of prime toric divisor implicitly assumes the existence of an orientifold projection and a tadpole-cancelling D7/O7 configuration; given the variety of gauge realisations in F-theory, this is a reasonable idealisation. In this work we assume $h_-^{1,1} = 0$ and therefore restrict to the orientifold-even sector. This assumption should not be taken to imply that superpotential-relevant $O(1)$ instantons are generically available: their existence is subject to consistency requirements arising from fermionic zero modes and the Freed–Witten condition [\[50–52\]](#). For the purposes of this work we remain agnostic about these details and focus on developing and testing tools for efficiently sampling Kähler moduli space for Bayesian inference. A systematic treatment of fluxed instantons and the general case with $h_-^{1,1} \neq 0$ will be presented in forthcoming work [\[53\]](#).

We develop tools to efficiently sample the WP measure in geometries with $h^{1,1}$ as large as 30, and show that modest computational resources could extend this sampling to much larger values. We furthermore construct a theory-informed prior on axion parameters (m_a, f_a) from an *exhaustive* treatment of CYs from the Kreuzer–Skarke (KS) database with $h^{1,1} \leq 5$ under the WP measure.² We impose model likelihoods for future QCD axion detection in the laboratory, and demonstrate how this can restrict CY topology and the distribution on moduli space.

As a representative of our results, [Fig. 1](#) summarizes a key achievement of this work. The figure shows the posterior probability on CY moduli space derived from three separate combinations of priors and likelihoods for a specific model with $h^{1,1} = 7$ (see [Sec. 4.2.3](#) in [Ref. \[16\]](#) for more details of the CY). Sampling from only the WP

prior gives a log-flat spectrum for the Calabi–Yau volume $\mathcal{V}$ – scale invariant – between a minimum value set by control over the EFT, and a maximum set by the KK scale. The second set of samples adds to this a prior that there exists at least one prime toric divisor with volume in a range around 40 in string units, such that it is possible in principle to host the QCD gauge group on a stack of D7-branes with a realistic gauge coupling in the UV, which further restricts $\mathcal{V} \lesssim 10^3$. Finally, we show the result of imposing cosmological likelihoods from the CMB and the Lyman alpha forest in a full forward model. The preference in this data combination for an ultralight axion dark matter (DM) component [\[57\]](#) strongly restricts the region in moduli space, leading to a sharply peaked posterior, visible here as a preferred value of the marginal distribution on the CY volume, and even more strikingly in the joint posterior on the CY volume and a selected divisor volume.

This paper is organised as follows. In [Section II](#) we discuss the Kähler cone and WP metric of CY threefolds, as well as other relevant aspects of Calabi–Yau topology and geometry. In [Section III](#) we develop efficient sampling methods for the WP measure, including MCMC and normalising flows. In [Section IV](#) we present geometric results under the WP prior, including the behaviour and distribution of volumes of the various divisor classes across different geometries. In [Section V](#) we apply our sampling methods to axion physics, construct theory-informed priors for (m_a, f_a) , incorporate QCD-like and cosmological likelihoods, and analyse the resulting constraints on CY topology and moduli space. We summarise our findings in [Section VI](#). The appendices contain technical material on the axion EFT, the normalising-flow architecture, aspects of the WP measure, and Calabi–Yau topology and geometry. All CY computations in this work employ CYTOOLS [\[2\]](#). Our results are publicly available on GitHub [\[2\]](#).

II. KÄHLER MODULI SPACES OF CALABI–YAU THREEFOLDS

A. Kähler moduli and the Kähler metric

We are here concerned with the question of how physics varies across the moduli space of CY compactifications of string theory [\[30\]](#). We consider the Kähler moduli space of a CY orientifold X given by the Kähler parameters t^i of dimension $h^{1,1} = h_+^{1,1}$. These parameters take values in the Kähler cone $K_X \subset H^2(X, \mathbb{R})$ corresponding to the region where volumes of subvarieties are positive. The complexified Kähler moduli space of type IIB string theory can be parametrised by coordinates

$$T_i = \tau_i + i\phi_i = \frac{1}{2} \int_X J \wedge J \wedge \omega_i + i \int_X C_4 \wedge \omega_i, \quad (\text{II.1})$$

where C_4 is the four-form with the associated axions ϕ_i , $J = t^i \omega_i$ is the Kähler form on X in a basis

² The Kreuzer–Skarke database of CY threefolds arising from fine, regular, star triangulations (FRSTs) of four-dimensional reflexive polytopes is quite degenerate, in that many FRSTs induce the same CY manifold. Using the methods of [\[54\]](#), we can efficiently remove a large fraction of this degeneracy, but not all of it. Removing all degeneracy is a difficult problem [\[55, 56\]](#) that we will not undertake in this work.

$\omega_i \in H^2(X, \mathbb{Z})$ and

$$\tau_i = \frac{1}{2} \kappa_{ijk} t^j t^k, \quad \kappa_{ijk} = \int_X \omega_i \wedge \omega_j \wedge \omega_k. \quad (\text{II.2})$$

Here, τ_i correspond to the volumes of the 4-cycles Poincaré dual to the ω_i and κ_{ijk} are the triple intersection numbers of X .

The classical Kähler potential for the Kähler moduli is³

$$\mathcal{K} = -2 \ln(\mathcal{V}), \quad \mathcal{V} = \frac{1}{6} \kappa_{ijk} t^i t^j t^k \quad (\text{II.3})$$

where $\mathcal{V}$ is the CY volume in string units [58]. From the Kähler potential, the Kähler metric is given by $K^{ij} = \partial^2 \mathcal{K} / \partial t_i \partial t_j$,⁴ and the four-dimensional kinetic term for the divisor volumes τ_i is

$$\mathcal{L}_{\text{kin}} \supset K^{ij} d\tau_i \wedge \star d\tau_j. \quad (\text{II.4})$$

The classical metric on moduli space is independent of the axions, see also Section A. Using (II.2), the kinetic terms for the Kähler parameters t^i become

$$\mathcal{L}_{\text{kin}} \supset M_{ij} dt^i \wedge \star dt^j, \quad \text{where} \\ M_{ij} = K^{ab} \kappa_{aci} \kappa_{bdj} t^c t^d = \frac{1}{4\mathcal{V}^2} (2\tau_i \tau_j - \mathcal{V} \kappa_{ijk} t^k). \quad (\text{II.5})$$

Therefore, a natural metric defined on the space of Kähler parameters t^i is simply M_{ij} referred to as the *Weil–Petersson* (WP) metric (see also Section C),⁵ giving the following probability measure

$$D_{\text{WP}} := \sqrt{|\det M|} \prod_i dt^i. \quad (\text{II.6})$$

This natural measure on the Kähler cone K_X is known as the WP measure.

B. Parametrising the Kähler moduli space

Let us begin by parametrising the Kähler cone K_X . It can be constructed as

$$K_X = \{t^i \mid t^i = \alpha^\mu \mathcal{G}_\mu^i; \alpha^\mu > 0\}, \quad (\text{II.7})$$

³ The Kähler potential in Eq. (II.3) and the real parts of the Kähler coordinates in Eq. (II.1) receive both perturbative and non-perturbative corrections, see e.g. [37] for a discussion. These effects modify the classical WP metric, yielding a quantum-corrected geometry which becomes essential once the moduli are taken outside the strict geometric regime, i.e., outside the SKC. We will at points consider values of tree-level (i.e. purely geometric) quantities outside of the SKC (such as when diagnosing the existence of fibrations), but this is merely to infer topological properties of the geometry rather than physical quantities.

⁴ We note that this convention differs from other references, such as [16, 17], which denote the Kähler metric with lowered indices and its inverse with upper indices.

⁵ The WP metric can be shown to be equivalent to other more fundamental metrics like the Zamolodchikov metric [59] which is a special case of the information metric, see e.g. [60].

where $\mathcal{G}_\mu^i$ is the $(N_g \times h^{1,1})$ -dimensional matrix whose rows are the extremal generators of K_X with N_g denoting their total number. The measure induced by WP on the space of α coefficients becomes

$$D_{\text{WP}}^\mathcal{G} = |\det \mathcal{G}| \sqrt{|\det M|} \prod_{\mu=1}^{N_g} d\alpha^\mu \Theta(\alpha^\mu > 0). \quad (\text{II.8})$$

For the purposes of traversing the cone efficiently, we further re-parametrise the positive α coefficients space, $\mathbb{R}_+^{N_g}$, into reals, $\mathbb{R}^{N_g}$, by exponentiation $\alpha^\mu = \exp(\theta^\mu)$. Such a parametrisation and scanning in the space of θ 's results in robust and efficient sampling of the Kähler cone, since far away points that are deep within the cone — points with exponentially large CY volumes — are easily accessible: in particular, multiplicative separations in the α 's become additive separations in the θ 's, so that regions corresponding to exponentially large α^μ lie only a linear distance $|\Delta\theta^\mu|$ away. This parametrisation also removes the hard boundary at $\alpha^\mu = 0$ and allows proposal steps to explore large-volume interior regions with uniform numerical efficiency. The WP measure in terms of the target variables θ^μ is therefore

$$D_{\text{WP}}^\mathcal{G} = |\det \mathcal{G}| \sqrt{|\det M|} \prod_{\mu=1}^{N_g} d\theta^\mu \alpha^\mu. \quad (\text{II.9})$$

In general, especially as $h^{1,1}$ increases, the majority of Kähler cones are non-simplicial ($N_g > h^{1,1}$). Thus, due to the large number of additional parameters, it becomes computationally difficult to sample non-simplicial, and consequently narrow, cones. Furthermore, it becomes computationally challenging to even determine all of the cone generators. Finally, for fibered CY manifolds, the support of the WP measure has non-compact regions (as we discuss in Section II.C) which can be narrow depending on the choice of basis, which also poses a challenge for sampling.

To circumvent these issues, we make use of the dual parametrisation of Kähler cones, i.e., using their extremal facets (extremal codimension-one faces, or extremal rays of the dual cone known as the Mori cone). Let $\mathcal{H}$ be the $(N_h \times h^{1,1})$ -dimensional matrix whose rows are the extremal facet normal vectors that carve out the Kähler cone; that is,

$$K_X = \{t^i \mid \mathcal{H}_i^\mu t^i > 0\}. \quad (\text{II.10})$$

By extremal we mean that for any $h^{1,1} \times h^{1,1}$ sub-block that is a full rank square matrix, no other rows are expressible as *positive* linear superpositions of the rows of this block. We pick the first such full rank $h^{1,1}$ rows of $\mathcal{H}$, call it $\tilde{\mathcal{H}}_i^j$, and perform a change of basis:

$$\bar{t}^j = \tilde{\mathcal{H}}_i^j t^i, \quad (\text{II.11})$$

In the new basis, the hyperplane matrix takes the form

$$\bar{\mathcal{H}}_i^\mu = \mathcal{H}_j^\mu [\tilde{\mathcal{H}}^{-1}]_i^j = \begin{pmatrix} \mathbb{I}_{h^{1,1}} \\ \mathcal{B} \end{pmatrix} \quad (\text{II.12})$$

where $\mathcal{B}$ is a $(N_h - h^{1,1}) \times h^{1,1}$ dimensional matrix with N_h being the number of extremal facets, rendering the Kähler cone as

$$\bar{K}_X = \{\bar{t}^i \mid \bar{t}^i > 0 \text{ and } \mathcal{B}_i^a \bar{t}^i > 0\}. \quad (\text{II.13})$$

That is, the transformed Kähler cone is guaranteed to be contained within the positive orthant $\mathbb{R}_+^{h^{1,1}}$. As before, we can then parametrise the Kähler parameters using an exponential map, $\bar{t}^i := \exp(\bar{\theta}^i)$, in order to efficiently traverse it. The WP measure in the new basis and after using $d\bar{t}^i = d\bar{\theta}^i \bar{t}^i$, is simply

$$D_{\text{WP}}^{\mathcal{H}} := \frac{\sqrt{|\det M|}}{|\det \mathcal{H}|} \prod_{a=1}^{N_h - h^{1,1}} \Theta(\mathcal{B}_i^a \bar{t}^i) \prod_{i=1}^{h^{1,1}} d\bar{\theta}^i \bar{t}^i \quad (\text{II.14})$$

With the above parametrisations of the Kähler cone and the associated WP measures (priors), we can impose additional constraints for physical/phenomenological considerations. In the context of a string compactification, we restrict ourselves to the *stretched Kähler cone* (SKC) [15] — the region of moduli space where volumes of all holomorphic curves are larger than unity (in string units). In fact, we shall also require that we remain within the *divisor stretched Kähler cone* (DSKC) — the region of moduli space where all prime toric divisor volumes are also larger than unity (in string units). It is within this divisor stretched cone that we can be most confident that the four dimensional EFTs (see Section A) are under perturbative control. From the Kähler cone K_X , the DSKC can be obtained as

$$K_{X,s} = K_X \cap \{t^i \mid H_i^a t^i \geq 1\} \cap \{\tau_I > 1\}, \quad (\text{II.15})$$

where H is the matrix of hyperplane normal vectors associated with *all* holomorphic curve constraints (not only the extremal facets), and the index I runs over all prime toric divisors (1 to $h^{1,1} + 4$). The inequalities $H_i^a t^i \geq 1$ enforce that the volumes of all effective curves exceed unity, thereby converting the conical region into an affine polyhedron.

In the absence of explicit extremal generators, we take

$$\bar{K}_{X,s} = \{\bar{t}^i \mid \bar{t}^i > 0, B_i^a \bar{t}^i > 1\} \cap \{\tau_I > 1\}, \quad (\text{II.16})$$

where B is the equivalent of Eq. (II.12), but comprising *all* the non-basis hyperplane vectors. That is, the non-identity (and non-square) matrix contained in $H_j^a [\bar{\mathcal{H}}^{-1}]_i^j$.

The restriction of $\tau_I > 1$ in both parametrisations can be imposed as an extra condition on top of the WP measure (Eq. (II.9) and Eq. (II.14) respectively).

The above discussion extends directly to $K_{\cup}$, a cone contained in the Kähler cone of a toric hypersurface CY threefold and often used as an approximation to the true Kähler cone (which is generally difficult to compute). In particular, $K_{\cup}$ for a CY hypersurface X in a toric variety V is defined as the union of the toric Kähler

cones of all toric varieties V' related to V by birational operations which do not affect X [61]. In what follows, all CY threefolds we study are toric hypersurfaces, and we work exclusively with $K_{\cup}$ (in CYTOOLS, $K_{\cup}$ can be constructed from a CalabiYau object `cy` as `cy.mori_cone_cap(in_basis=True).dual()`).

How does the WP measure scale within the stretched Kähler cone of a CY compactification of type IIB string theory? Using $\mathcal{K} = -2\ln(\mathcal{V})$, the entries of the Kähler metric K^{ij} scale like $\sim \tau^{-2} \sim t^{-4}$. Using Eq. (II.5), the entries of the M matrix scale like $\sim t^{-2}$, giving $\sqrt{|\det \mathcal{M}|} \sim \sqrt{|\det M|} \sim t^{-n}$ where $n = h^{1,1}$. With this, we see that the WP measure scales with the CY volume in the following fashion:

$$D_{\text{WP}} \sim d\Omega_{n-1} \frac{d\mathcal{V}}{\mathcal{V}} \sim d\Omega_{n-1} d\mathcal{K}, \quad (\text{II.17})$$

where $d\Omega_{n-1}$ is some angular measure that we need not compute explicitly. An important point to note is that the WP measure enforces scale-invariance in the CY volume. While this means that the WP prior is not normalisable a priori, there are lower and upper cut-off CY volumes that can be imposed on practical grounds. The lower cut-off of $\mathcal{V}_{\min}$ can simply come from the tip of the stretched Kähler cone (TSKC), at which point the CY volume may even be larger than unity [15]. On the other hand, the upper cut-off of $\mathcal{V}_{\max}$ could be imposed by, e.g., an experimental lower bound on the SUSY scale, or upper bound on the size of extra dimensions. The scale and diffeomorphism invariance properties of the WP measure appear to be attractive from a Bayesian perspective due to the relation with the Jeffreys prior.

C. Topological and Geometric Prerequisites

Upon computing the parameter space for our Kähler moduli — the Kähler cone — and the axion EFT (see Section A), one might be tempted to set aside the geometric interpretation of the subvariety volumes which determine the kinetic matrix and potential and just treat them as free independent parameters. However, the topology and geometry of the CY furnishes important constraints and correlations on the free parameters in the EFT — namely, volumes of cycles. In particular, these will also be exploitable, enabling us to both engineer specific physical scenarios using topology, as well as constrain topology with phenomenology.

Three particular geometric phenomena will be of interest to us in this work: boundaries of the Kähler cone where the overall CY volume vanishes (which correspond to fibrations), boundaries where divisor volumes vanish (which include but are not restricted to boundaries of the former kind), and how the dichotomy between divisors with and without normal bundle deformations imprints itself on axion decay constants. We summarise

these three ideas in this section, but include a longer discussion in [Section C.3](#).

Let us begin with the two types of Kähler cone boundaries. The classical, geometric moduli space of a Calabi–Yau threefold decomposes as a union of Kähler cones, forming the extended Kähler cone. The faces of a Kähler cone where the CY volume or individual divisor volumes vanish are boundaries not only of that Kähler cone but of the entire extended Kähler cone. Even when our stretching condition is imposed, if one approaches these faces of the Kähler cone, there are imprints on the axion physics. In this way, the near-boundary regions of Kähler moduli space feature interesting axion phenomenology that can be dissimilar from the “bulk” phenomenology near a typical TSKC ray.

Boundaries where the CY volume vanishes are important to us first because they control the support of the WP measure. As highlighted earlier, the WP measure is normalised only upon imposing bounds on the overall CY volume $\mathcal{V}$. The support of the WP measure — i.e., the region where the CY volume $\mathcal{V}$ satisfies $\mathcal{V}_{\min} \leq \mathcal{V} \leq \mathcal{V}_{\max}$ — is compact if and only if $\mathcal{V}$ does not vanish on any face of the Kähler cone. In particular, the support of the WP measure asymptotically and non-compactly approaches any face where $\mathcal{V}$ vanishes. Such faces are in correspondence with fibrations of X , and correspond to limits where both the fiber volume and $\mathcal{V}$ vanish. While imposing stretching constraints will cut off this non-compactness, fibered geometries will still feature long, narrow regions in the Kähler cone near the faces where $\mathcal{V}$ vanishes, corresponding to the limit of small fiber volume and large base volume. Beyond giving rise to qualitatively distinct axion physics compared to the interior of moduli space, these regions are additionally difficult to sample and thus pose computational challenges.

Fibrations of CY manifolds must have generic fibers that are themselves CY, so the fiber can either be a genus-one curve (i.e., a torus T^2) if the fiber is one-dimensional, or either a four-torus T^4 or $K3$ surface if the fiber is two-dimensional [62]. When a one-dimensional fiber admits a section the fibration is elliptic; otherwise it is genus-one with a multisection. For brevity we will still say “elliptic” unless the distinction matters. Likewise, we will abuse notation and refer to the two-dimensional fiber case as a $K3$ fibration.

Now we turn our attention to Kähler cone faces where divisors D attain zero volume, $\tau = 0$. Inverse divisor volumes set important scales in string compactifications: for example, $1/\sqrt{\tau}$ sets the gauge coupling on worldvolume gauge theories arising from stacks of D7-branes wrapped on $\mathbb{R}^{1,3} \times D$. Similarly, the limit in which divisors shrink to points is also considered in constructions of Standard Model-like physics on D-branes located at the resulting singularities [33, 63–67]. Analogous to our discussion of the overall volume for fibrations, if D shrinks on a face of the Kähler cone, then the region in moduli space where D has volume less than a certain value — i.e., the region

where a phenomenologically realistic gauge theory can be realised — is non-compact, which is interesting and relevant for phenomenological purposes.

Additionally, while the scale of divisors volumes are often set by the overall volume, $\tau \sim \mathcal{V}^{2/3}$ [23, 25], some divisors can shrink along Kähler cone faces where the CY volume is finite, and thus their volume can be characteristically smaller than the overall volume. Such divisors are known in the literature as *blowup* divisors. We distinguish between divisors which shrink to points and to curves, distinguished by whether their quadratic volume polynomials vanish quadratically or linearly on a face. The decoupling with the overall volume is more prominent for the former, as we will see in [Section IV](#). By contrast, we will refer to divisors that shrink near faces where $\mathcal{V}$ does vanish (due to a shrinking fiber) as *fibration divisors*. Fibration divisors can also be small at large overall volume, but must be compensated by a large base volume, while blowup divisors do not require such a compensation. Divisors with volume τ decoupled from the overall volume are important, as then τ behaves more like a free parameter, yielding useful freedom when engineering the axion EFT. Shrinkability at large overall volume will prove important when implementing QCD axions within our constructions in [Section V.B](#).

In summary, then, we will find it useful to distinguish five categories of divisors, organised by their shrinking behaviour.

- **Non-shrinkable divisors**, which have finite volume everywhere in the Kähler cone
- **Fibration divisors**, which can shrink in the limit that the CY volume vanishes (i.e., the fiber shrinks)
- **$K3$ fiber divisors**, which (up to overall scaling) are the fibers of $K3$ fibrations.
- **Curve blowup divisors**, which can shrink to curves while the CY remains at finite volume (i.e., their volume vanishes linearly)
- **Point blowup divisors**, which can shrink to points while the CY remains at finite volume (i.e., their volume vanishes quadratically)

Divisors can belong to more than one of these categories (e.g., $K3$ divisors are a special case of fibration divisors, and divisors shrinking linearly along some face can shrink quadratically along another, etc.), so we will assign membership by iterating over this list from bottom to top and applying the first label whose condition is satisfied.

It is not an accident that this list resembles the taxonomy of facets of CY threefold Kähler cones (reviewed in [Section C.3](#)) but it’s worth stressing that the limits where shrinkable divisors (i.e., members of the first three categories) have vanishing volume need not be at codimension-one in the Kähler cone: higher codimension can and does happen, and such cases will be important for us.

Finally, we can discuss the relevance of normal bundle deformations for axion decay constants. If D is a divisor belonging to the basis of leading instanton contributions, then in the limit of hierarchical instanton scales (which arises generically in Kähler moduli space, and was studied in [17])⁶

$$f \gtrsim \mathcal{O}(1) \cdot \frac{1}{\sqrt{\tau^2 - \mathcal{V}C}}. \quad (\text{II.18})$$

for τ the volume of D and C the formal volume of the curve class $D \cap D$ (though such a class need not have holomorphic representatives).

The Hodge number $h^{2,0}(D)$ ⁷ counts normal bundle deformations of a divisor D in a CY threefold, and we say D is *geometrically rigid* if $h^{2,0}(D) = 0$, and *movable* otherwise.⁸ We explain in Section C.3 how, for geometrically rigid divisors D , $\tau^2 - \mathcal{V}C$ is often the sum of positive terms (as usually $C < 0$) and is essentially set by $\mathcal{V}$, while for movable D , $C > 0$, so $\tau^2 - \mathcal{V}C$ is smaller and bounded from above and below for fixed τ by positivity of C and of K^{ij} , respectively. Thus, decay constants for geometrically rigid divisors are characteristically smaller, and have larger variance compared to their movable counterparts.

An edge case worth noting is when D is a $K3$ fiber divisor, for which $C = 0$ and the decay constant is totally decoupled from the overall volume, being purely set by the divisor volume τ . Crucially, this can be kept $\mathcal{O}(1)$ even at large volume by taking the anisotropic limit of large base and small fiber. In this way, $K3$ divisors provide a mechanism for engineering anomalously large axion decay constants, in a manner independent of overall volume. This was exploited in [16] to engineer decay constant hierarchies, which in turn was inspired by [69, 70].

III. SAMPLING MODULI SPACE

In this section, we lay out two techniques for sampling the SKC: an MCMC method implemented via the Python-compatible package EMCEE [71], and an NF method implemented using the NORMFLOWS package [72].

A. Classical Markov Chain Monte Carlo

For the MCMC approach, we employ two distinct parameterisations of the SKC as outlined in the previous section. The first uses the extremal generators of the original Kähler cone, while the second uses the extremal hyperplanes of the modified Kähler cone.

In the two MCMC strategies, the former has the advantage of no rejections beyond the stretching conditions—by construction, all sampled points lie within the Kähler cone—but may require sampling in a space of dimension larger than $h^{1,1}$ (in the case of non-simplicial cones). The latter, in contrast, restricts sampling to an $h^{1,1}$ -dimensional space independent of the number of generators, but introduces additional rejections when the cone is non-simplicial (since then the modified Kähler cone is a strict subset of $\mathbb{R}_+^{h^{1,1}}$).

For both strategies, we sample according to the natural Weil-Petersson (WP) prior defined by the metric M_{ij} in Eq. (II.5). Using the measures derived in Eq. (II.9) and Eq. (II.14), the corresponding log-probabilities—hereafter referred to as the WP Prior—are given by

$$\log P_{\text{WP}}^{\mathcal{G}} = \frac{1}{2} \ln |\det M| + \sum_{\mu=1}^{N_g} \theta^\mu + \ln |\det \mathcal{G}|, \quad (\text{III.1})$$

$$\log P_{\text{WP}}^{\mathcal{H}} = \frac{1}{2} \ln |\det M| + \sum_{i=1}^{h^{1,1}} \bar{\theta}^i - \ln |\det \bar{\mathcal{H}}|. \quad (\text{III.2})$$

The above are defined up to additive constants. Points violating any facet or stretching constraint ($H_i^a t^i \geq 1$ or $\tau_p^{\text{pt}} > 1$) are assigned $-\infty$. These log-prior expressions form the probability landscape explored by the MCMC chains.

1. Implementation and convergence tests

For purposes of this section, each MCMC chain was run within the divisor-stretched Kähler cone (DSKC) of a given CY geometry, bounded between $\mathcal{V}_{\min} = \mathcal{V}_{\text{tip}}$ and $\mathcal{V}_{\max} = 10^{10} \mathcal{V}_{\text{tip}}$, where $\mathcal{V}_{\text{tip}}$ is the CY volume at the tip of the stretched Kähler cone⁹. For the generator-based sampling, the ensemble comprised $N_{\text{wlk}} = 10 N_g$ walkers, while the hyperplane-based sampling used $N_{\text{wlk}} = 10 h^{1,1}$. Convergence of each chain was determined through autocorrelation analysis of individual parameters: once the number of statistically uncorrelated samples in the slowest (most correlated) parameter exceeded 75, the run was terminated. This uniform criterion allows a geometry-independent comparison of convergence efficiency.

⁶ We are indebted to Sebastian Vander Ploeg Fallon for sharing Eq. (II.18) with us.

⁷ Convenient combinatorial formulae for Hodge numbers of certain divisors in toric hypersurface CY threefolds are presented in [68].

⁸ We say “geometrically rigid” rather than just “rigid” to avoid confusion with the divisors demonstrated in [40] to contribute to the ED3 non-perturbative superpotential in type IIB: those rigid divisors D are also geometrically rigid, but additionally satisfy $h^{0,0}(D) = 1$ and $h^{1,0}(D) = 0$.

⁹ The tip of the stretched Kähler cone is obtained by minimising the Euclidean norm $\|t\|_2$ subject to the constraint $H_i^a t^i \geq 1$.

The procedure was repeated across 100 randomly selected geometries for each $h^{1,1} \leq 11$. For each $h^{1,1}$, we uniformly sampled 100 polytopes from the KS list [73] and then choose a random geometry by employing the methods of [24, 54]. For each geometry, we recorded the total time required to reach convergence, and compiled the results both as a function of $h^{1,1}$ and of the number of extremal generators N_g . These results are summarised in Fig. 2.

In Fig. 3, we present the corner plot visualising the sampled region of the Kähler cone for an example geometry, while in Fig. 4 we show the associated distribution of CY volumes. The volume histogram exhibits logarithmic flatness, correctly reflecting the expected WP prior behaviour and providing an additional indication of convergence.

2. Comparison of sampling performance

The two strategies display complementary strengths. The generator-based method samples directly along the extremal directions of the cone, ensuring that all proposed points automatically belong to the Kähler cone. Its primary cost, however, is dimensional: for non-simplicial cones with $N_g > h^{1,1}$, the effective sampling space grows quickly, leading to longer convergence times. In contrast, the hyperplane-based method operates in a fixed $h^{1,1}$ -dimensional space and thus scales more mildly with $h^{1,1}$, but introduces additional rejections whenever the cone is a strict subset of the positive orthant.

The convergence-time distributions reflect this trade-off. Across the ensemble, convergence time correlates more strongly with the number of extremal generators N_g than with $h^{1,1}$ itself. At large $h^{1,1}$, the hyperplane-based approach exhibits superior numerical stability and faster convergence, whereas at low $h^{1,1}$ the generator-based sampling remains competitive and provides a more direct geometric interpretation of motion within the cone. For large-scale statistical scans of CY geometries, the hyperplane-based WP prior therefore constitutes the more efficient and robust choice. We will use the hyperplane-based sampling for cosmological purposes in later sections.

These observations motivate the use of more flexible, learnable priors for navigating highly non-simplicial cones, which we turn to next.

B. Normalising flows

Normalising flows (NFs) [74] are a method for constructing complex distributions by transforming a simple base distribution (e.g., a Gaussian) through a series of invertible mappings. Let $f : \mathbb{R}^D \rightarrow \mathbb{R}^D$ be an invertible and differentiable function whose inverse maps the input data $\mathbf{x} \in \mathbb{R}^D$ to the latent representation $\mathbf{z} \in \mathbb{R}^D$, i.e.,

$f^{-1}(\mathbf{x}) = \mathbf{z}$ or equivalently $f(\mathbf{z}) = \mathbf{x}$. The main idea is to express the distribution $p_x(\mathbf{x})$ in terms of the base distribution $p_z(\mathbf{z})$ and f . This is done using the change of variables formula:

$$p_x(\mathbf{x}) = p_z(\mathbf{z}) |\det J_f(\mathbf{z})|^{-1}, \quad (\text{III.3})$$

where J_f denotes the Jacobian of f . For more details on NF see Section B.

A base Gaussian distribution has full support, which means that $p_z(\mathbf{z}) > 0$ everywhere on $\mathbb{R}^D$. The invertible nature of f means that using such a base distribution, the target distribution $p_x(\mathbf{x})$ must also have full support. For this reason, we choose the extremal generators parametrisation of the original Kähler cone with target WP given in Eq. (III.1), as opposed to the extremal hyperplanes parametrisation of the modified Kähler cone with WP given in Eq. (III.2), since the latter does not have full support in the space of θ s. Unlike MCMC where the chains run inside the DSKC, for the NF we do not impose any stretching constraints during training. Instead, we train the NF to learn the WP measure on the full Kähler cone. Once trained, we then use the NF to sample points and only keep those that lie inside the DSKC and whose volume is between $\mathcal{V}_{\min} = \mathcal{V}_{\text{tip}}$ and $\mathcal{V}_{\max} = 10^{10} \mathcal{V}_{\text{tip}}$.

Specifically, we fit a real NVP model [75] $p_\theta(\theta; \Theta)$, where f is composed of a series of 16 masked affine flow layers f_i with parameters Θ , to the target distribution $p_\theta^*(\theta)$ defined by WP Eq. (III.1).¹⁰ Since we can easily evaluate $p_\theta^*(\theta)$ for a given point θ in Kähler moduli space, but cannot easily sample from this distribution, we train our NF model using reverse Kullback–Leibler (KL) divergence D_{KL} Eq. (B.3). In evaluating $p_\theta^*(\theta)$ we must compute the matrix M from Eq. (II.5). This involves transforming θ back into Kähler parameters $\mathbf{t}$, by first computing $\alpha^\mu = \exp(\theta^\mu)$ and then $t^i = \alpha^\mu \mathcal{G}_\mu^i$. In practice, if θ^μ is too large then we encounter overflow errors during training from $\exp(\theta^\mu)$, therefore we add an additional term to the loss function: $\text{ReLU}(\|\theta\| - R)^2$, which is zero if $\|\theta\| < R$ and $(\|\theta\| - R)^2$ otherwise. This term penalises θ larger than R which we set to be 15. The total loss function is thus

$$\mathcal{L}(\Theta) = D_{KL}[p_\theta(\theta; \Theta) | p_\theta^*(\theta)] + w \times \text{ReLU}(\|\theta\| - R)^2, \quad (\text{III.4})$$

where w is a tunable weight parameter that we set to 1.

In each of the 16 masked affine flow layers, we use an alternating binary mask and the translation and scale networks are MLPs consisting of 4 hidden layers each with 256 neurons. This gives a total of 6,373,472 learnable parameters. We train our NF model over 50,000 epochs using the Adam optimiser to find the optimal values for these parameters, which minimise the total loss Eq. (III.4) on a batch of size 1,024. The initial learning rate is set to 10^{-5} and we use PyTorch learning rate

¹⁰ For more details on the real NVP architecture, see Section B.

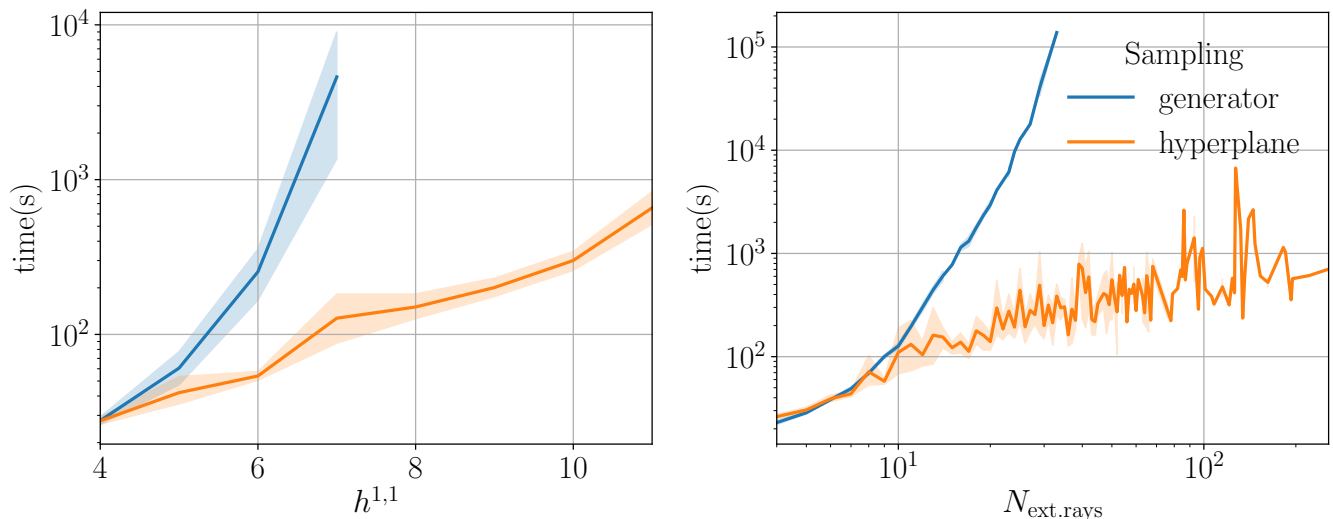


FIG. 2. Convergence times for various values of $h^{1,1}$ and different numbers of generators for hyperplane sampling using EMCEE. The full likelihood consists solely of the Weil–Petersson prior, together with an upper bound on the Calabi–Yau volume fixed at $10^{10} \times \mathcal{V}_{\text{tip}}$. We declare convergence once the total number of MCMC steps performed for the slowest converging parameter, exceeds 75 times its autocorrelation length.

scheduler REDUCELRONPLATEAU with factor 0.5 and patience 1000, which reduces the learning rate by one half if there is no improvement in the loss over 1,000 epochs. Training was performed on a 14-core CPU, 32-core GPU laptop equipped with 36 GB of unified memory, and required approximately 5 hours to complete 50,000 epochs. Due to the time taken to train the model, a limited amount of hyperparameter tuning was performed to determine these settings.

To compare the results of NF with MCMC, we generate samples (10,000) via both methods for the same geometry at $h^{1,1} = 3$, which has $N_g = 4$ generators, and compare their distributions. The corner plot of these samples with the associated distribution of CY volumes is shown in Fig. 5, which demonstrates reasonable agreement between the methods.

IV. GEOMETRIC RESULTS

Our primary application of the methods presented for sampling Kähler moduli space will be the phenomenology of C_4 axions in type IIB string theory, but first we will briefly study something simpler: distributions of volumes of divisors under the Weil–Petersson measure. Beyond the importance of understanding the distribution of Kähler moduli vevs — recalling Eq. (II.1) — under the natural moduli space metric, the statistics of divisor volumes τ is important as these geometric quantities set many scales in low energy string theory EFTs, including the gauge coupling in D7-brane worldvolume theories (as $1/\sqrt{\tau}$) and the C_4 axion masses (as $m \sim e^{-2\pi\tau}$). Indeed, efforts have been undertaken to model divisor volumes and apply such models to understand the universal be-

haviour of low-energy physics in string compactifications [23]. As discussed in Section II.C, it is especially worth understanding whether divisor volumes can be treated as free parameters, or if they are instead strongly correlated in a way that imprints on low energy phenomenology.¹¹

We will begin by presenting statistics of divisor volumes as a function of $h^{1,1}$ as well as the topological categories introduced in Section II.C, before exploring a specific example further illustrate the role of those topological categories. We will then interface with [23], expanding on their result regarding the role of the minface dimension of a prime toric divisor (defined shortly). Finally, we will present divisor volume spectra for a geometry with $h^{1,1} = 30$ with the important property (optimised using a genetic algorithm [24]) that its K_U is simplicial, making sampling feasible in spite of the large Hodge number.

In [15], distributions of divisor volumes on Kreuzer–Skarke CY threefolds were studied as a function of $h^{1,1}$, motivated by the desire to understand the string axiverse in this setting. This statistical study was undertaken by sampling points near the tip of the stretched Kähler cone: in particular, points minimising divisor and CY volumes subjects to unit stretching. We are now equipped to explore these distributions under the WP measure. In par-

¹¹ For divisor volumes in particular, we are essentially asking about the structure of the cone of dual coordinates of a CY threefold (in the sense of [76]), or the image of the extended Kähler cone under the map $J \mapsto J \wedge J$, as when expanded in a basis of four-forms, the coefficients of $J \wedge J$ are interpreted as the volumes of the dual basis of divisors classes. Strong correlations among divisor volumes are encoded in the narrowness of this cone of dual coordinates.

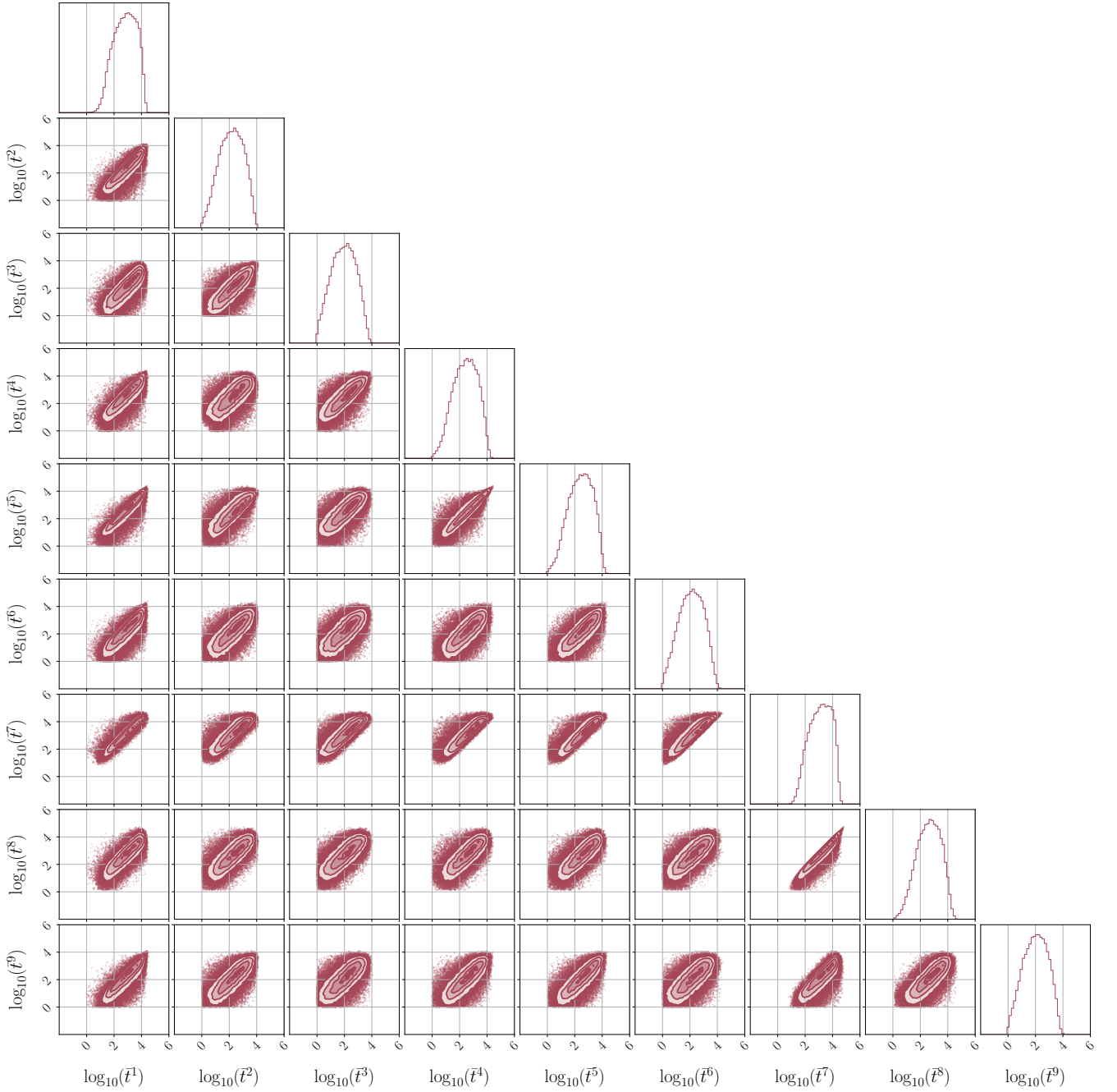


FIG. 3. Marginal one-dimensional and two-dimensional distributions of the modified positive semi-definite Kähler parameters, $\vec{t} \in \mathbb{R}_+$ basis as defined in Eq. (II.11), for a cone with $N_g = 28$ extremal generators, at $h^{1,1} = 9$ obtained using hyperplane sampling of the Weil–Petersson prior.

ticular, in Fig. 6 we present the distribution of prime toric divisor volumes as a function of $h^{1,1}$, marginalised over geometries. More specifically, for $h^{1,1} = 3, 4$ we marginalise over all FRST classes, while for larger $h^{1,1}$ we marginalise over the random geometries sampled in Section III A 1.

To illustrate the dependence of the divisor volume distribution on the topology of the divisor, in Fig. 7 we present a plot similar to the previous figure but now

marginalised over all geometries at $h^{1,1} = 4$, coloured by the type of the divisor, as introduced in Section II C. We find that $K3$ fibers and point blowup divisors, in particular, tend to be smaller, with the former featuring a smaller mean and the latter’s distribution featuring a longer tail toward unit volume (recall our prior imposes that it cannot be smaller).

To probe the extent to which divisor volumes behave as free independent parameters or are instead set by scales

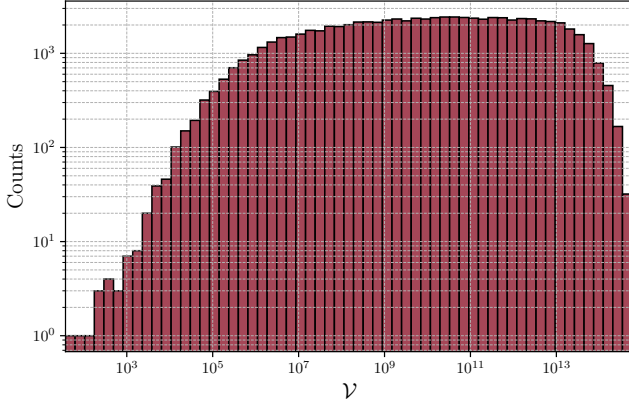


FIG. 4. Converged histograms of CY volume for example at $h^{1,1} = 9$ with a Kähler cone with $N_g = 28$ extremal generators. The likelihood is simply just the WP prior, with maximum CY volume set at $V_{\max} = 10^{10} \times V_o$ where V_o is the volume of the CY at the TSKC.

such as the overall volume, in Fig. 8 we plot the distribution of the dimensionless parameter $\tau/V^{2/3}$ as a function of V for all prime toric divisors for Kähler parameters marginalised over all geometries in KS at $h^{1,1} = 4$, coloured by the type of the divisor in the same way as the previous plot. We see that while divisor volume distributions peak near the locus $\tau/V^{2/3} = 1$ — i.e., the divisor volume scale is set by the overall volume — $K3$ and point blowup divisors in particular are able to be much smaller, even at large overall volume. Additionally, vertical and curve-blowup divisors are capable of smaller volumes than non-shrinkable divisors at each value of the overall volume. With the gray dashed line we have indicated our stretching constraint, requiring that all divisor volumes be at least one. With the black dashed line we emphasise an empirical finding that divisor volumes tend to obey $\tau \leq V$. For $K3$ fibered geometries, for example, limits where the overall volume factorises as $V \sim \tau \cdot C$ for τ and C the volumes of the $K3$ fiber and base curve, respectively, requiring $C \geq 1$ imposes $\tau \leq V$. However, we do not have a more general explanation for why this inequality holds for all divisor categories.

It is worth seeing in an example how divisor volume topology encodes correlations among volumes. Consider the 4D reflexive polytope from the Kreuzer–Skarke database whose vertices are the columns of the following matrix.

$$\begin{pmatrix} -3 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{IV.1})$$

A triangulation of this polytope yields a toric variety with a CY hypersurface with $h^{1,1} = 3$. Of the seven prime toric divisors, there are only four unique divisor classes, which we will denote D_1, D_4, D_5, D_7 according to the default ordering of prime toric divisors set in CY-

TOOLS. This geometry has a simplicial K_U : two facets correspond to flops, while one corresponds to a boundary of Kähler moduli space where D_7 shrinks to a point (i.e., D_7 is a point blowup divisor). The ray of K_U given by the intersection of the shrinking divisor facet and one of the flop facets corresponds to a $K3$ fibration over $\mathbb{P}^1$. The pullback of the hyperplane class on $\mathbb{P}^1$, which generates this ray and is the class of the $K3$ fiber, is D_5 . The ray given by the intersection of the two flop facets is also a boundary of moduli space, and the divisor D_1 shrinks to a curve along this ray. Finally, D_4 is non-shrinkable.

We can readily understand the corner plot of prime toric divisor and overall volumes featured in Fig. 9 given this analysis of the Kähler cone and divisor topologies. In particular, we see that the point blowup divisor D_7 and $K3$ fiber D_5 are less correlated with the overall volume and the other prime toric divisor volumes than the curve blowup divisor D_1 and non-shrinkable divisor D_4 (and D_4 is most correlated with the overall volume of them all). Additionally, we see that $\tau_5 > \tau_7$; one can verify this explicitly by comparing their volume polynomials, but intuitively this is because the limit of small $K3$ fiber is along a ray of K_U contained in the facet where D_7 shrinks, so τ_5 cannot be small without τ_7 being small as well. In this way, the face structure of K_U — i.e., its face lattice — controls correlations between divisor volumes.

Finally, we employ our WP sampling capabilities to interface with the Cheng–Gendler model proposed in [23] for distributions of prime toric divisor volumes on CY threefolds arising from a polytope Δ° . There it was observed that when performing a random walk in Kähler moduli space according to a particular scheme, volume distributions naturally organise into a trimodal distribution. In particular, the volume distribution of a prime toric divisor is characterised by its minface dimension [68], or the dimension of the smallest face of Δ° containing the point associated to the prime toric divisor. Vertices, points strictly interior to one-faces (edges), and points strictly interior to two-faces have minface dimension 0, 1, 2, respectively, and we call their associated divisors vertex, edge, and face divisors, respectively. For $h^{1,1} \lesssim 10$, the vast majority of prime toric divisors are of vertex type, but for larger $h^{1,1}$ more edge and face divisors begin to appear (and for $h^{1,1} \gg 50$, such divisors dominate). In [23], the normalised logarithmic volumes $\hat{\tau}_I$ were defined in terms of the vector τ_I of prime toric divisor volumes, which was treated with an overall prefactor such that the smallest entry in τ_I was 1. In particular,

$$\hat{\tau}_I = \frac{\log(\tau_I)}{\langle \log(\tau) \rangle}, \quad \langle \log(\tau) \rangle = \frac{\sum_{I=1}^{h^{1,1}+4} \log(\tau_I)}{h^{1,1} + 4}. \quad (\text{IV.2})$$

In Fig. 10, for the random geometries selected in Section III A 1 at $h^{1,1} = 10$ we aggregate each component of $\hat{\tau}$ across a WP chain (omitting the smallest entry in $\hat{\tau}$, following [23]), coloured by whether the associated divisor was of vertex, edge, or face type. We reproduce the

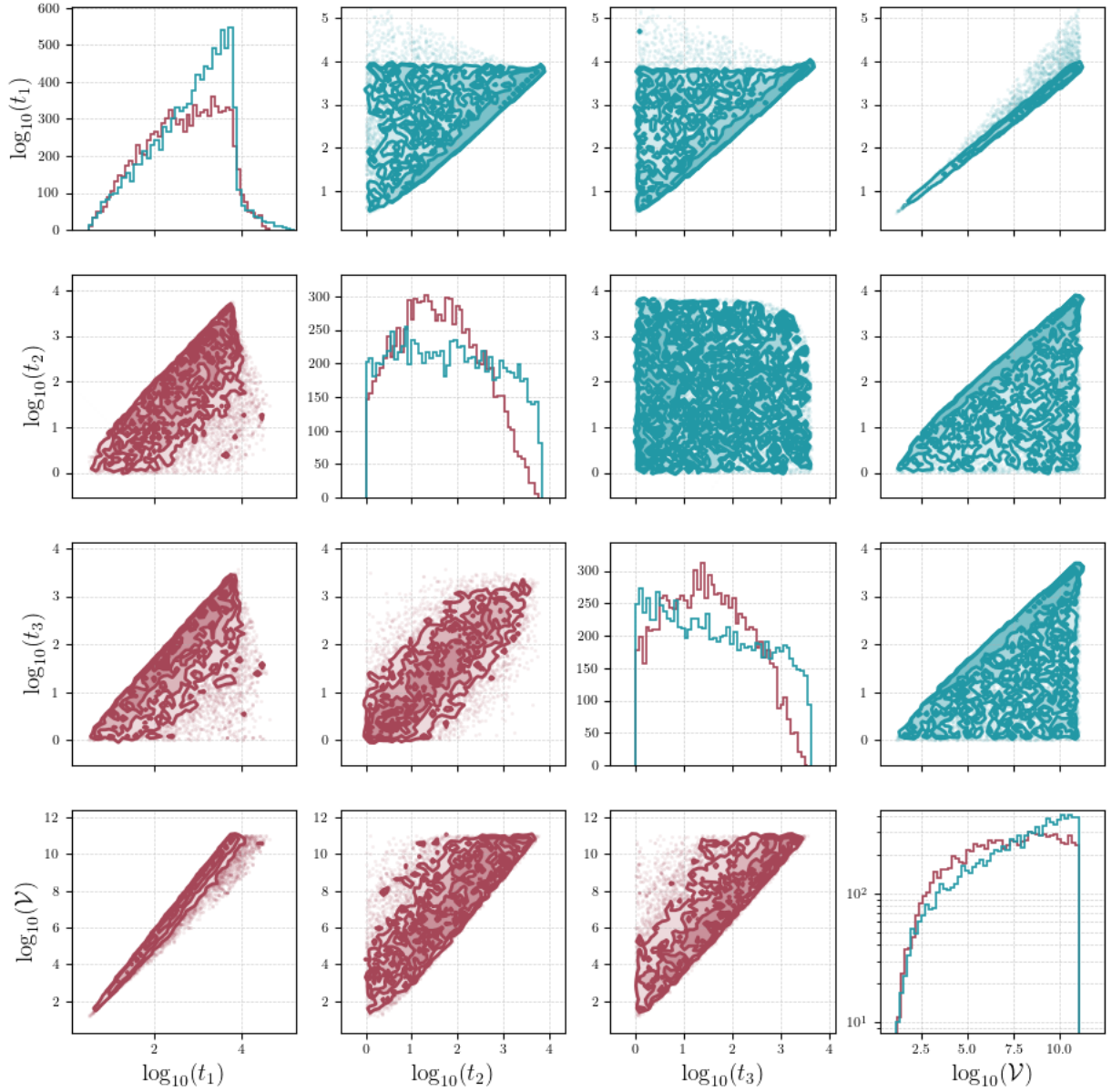


FIG. 5. Distributions of the Kähler parameters inside the DSKC and the associated distribution of Calabi-Yau volumes, for a cone with $N_g = 4$, at $h^{1,1} = 3$ obtained using MCMC (red) sampling with WP prior, and the trained NF model (cyan).

trimodality, verifying that the observations of [23] continue to hold upon implementation of a Bayesian sampling methodology using the natural metric on Kähler moduli space. We anticipate that the trimodality is related to the frequency with which vertex, edge, and face divisors belong to the various shrinkable divisor categories introduced in Section II.C, but we do not study this further here.

In Fig. 11 we present the distributions of the volumes of the prime toric divisors under the WP measure for a geometry at $h^{1,1} = 30$. Because the computational complexity of sampling scales with both $h^{1,1}$ and the number of extremal rays of the Kähler cone (which itself tends to grow exponentially with $h^{1,1}$), the sampling of generic

geometries for $h^{1,1} \gtrsim 15$ is quite challenging without dedicated computational resources. However, if one can construct geometries at large $h^{1,1}$ with simplicial (or nearly simplicial) $K_\cup$, sampling can be performed on a laptop for even larger $h^{1,1}$. As a proof of concept, we employed the genetic algorithm of [24] to minimise the number of extremal rays of $K_\cup$, and found geometries at $h^{1,1} = 30$ with simplicial $K_\cup$ (while typical geometries had $\gtrsim 50$ extremal rays). We do not expect such geometries to be representative: we merely endeavoured to maximise the $h^{1,1}$ for which we could sample *some* geometry.

We style the curves in Fig. 11 according to the topological categories of the associated prime toric divisors introduced in Section II.C. In particular, this geometry

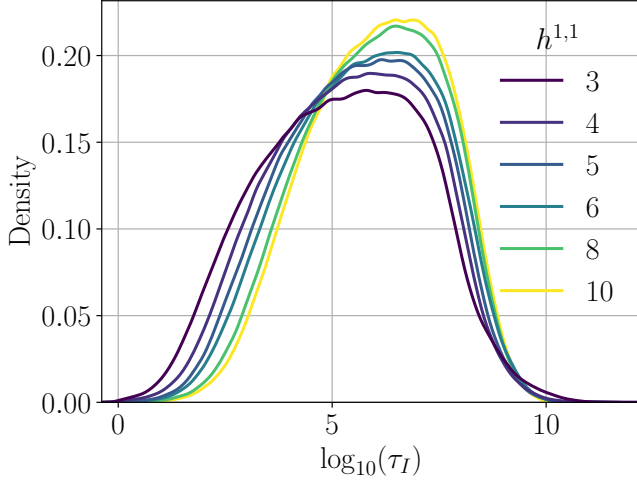


FIG. 6. Distributions of prime toric divisor volumes, marginalised over the WP prior, all prime toric divisors, and over topologies, labelled by $h^{1,1}$. Marginalisation over geometries is exhaustive for $h^{1,1} = 3, 4$, and otherwise is performed over the random geometries selected in [Section III A 1](#).

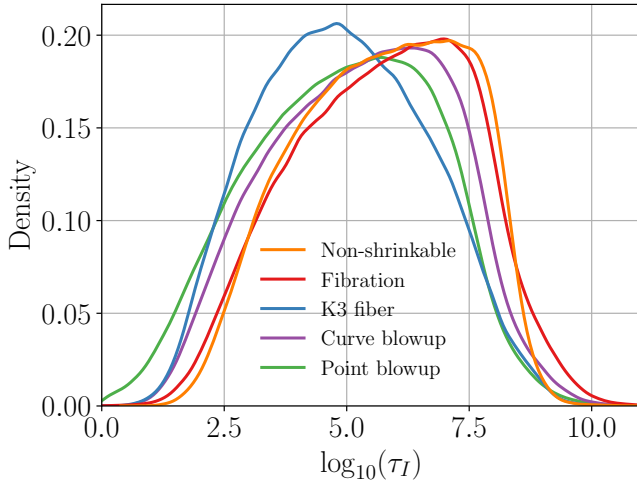


FIG. 7. Distributions of prime toric divisor volumes, marginalised over the WP prior and over all topologies in Kreuzer-Skarke at $h^{1,1} = 4$, labelled by the divisor topology categories specified in [Section II C](#).

has no fibrations (at least, none corresponding to faces of $K_\cup$) so that all shrinkable divisors are of blowup type, and we found it relevant to distinguish divisors based on whether they shrunk along facets (i.e., codimension one faces) or only along higher faces. It is unsurprising that divisors shrinking along facets have characteristically smaller volumes, as they are small in a greater fraction of the moduli space. More generally, the differences in distributions between point and curve blowup divisors vanishing on higher codimension faces can be accounted for by the exact codimension of those faces (the larger the codimension, the larger the volumes tend to be).

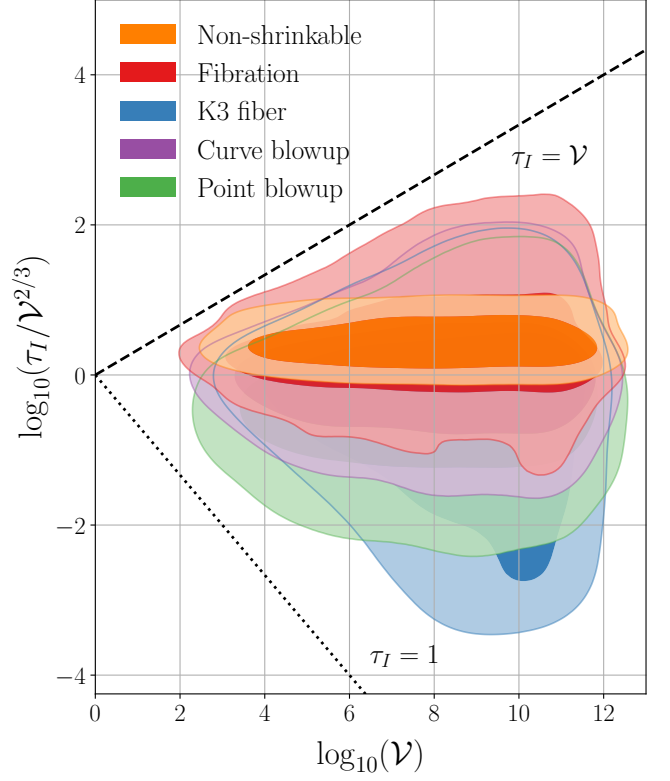


FIG. 8. Joint distribution of normalised prime toric divisor volumes and overall volume, marginalised over the WP prior and over all topologies in Kreuzer-Skarke at $h^{1,1} = 4$, labelled by the divisor topology categories specified in [Section II C](#).

V. APPLICATION TO AXIONS

As above, throughout this section we scan SKCs with the WP prior, while now imposing additional likelihoods that we will define. We run MCMC scans using hyperplane sampling technique, with number of walkers equal to $10 \times h^{1,1}$. All our walkers start within the geometrically allowed divisor stretched Kähler cone, randomly spread throughout the cone capped at the maximum volume $\mathcal{V}_{\max}$. To gauge convergence, we compute the autocorrelation times for each parameter, and run our MCMC until the total MCMC steps of the slowest converging parameter is equal to $\text{scale} \times \text{autocorrelation time}$, where we set the scale to be 75. This makes sure that for this slowest converging parameter, we have ~ 75 statistically independent samples, thereby guaranteeing that all other parameters have *at least* ~ 75 statistically independent samples. The minimum and maximum volume bounds that we impose (alongside the WP prior), are

$$\mathcal{V}_{\min} = 1$$

$$\mathcal{V}_{\max} = \left(\frac{g_s M_{\text{pl}}}{\sqrt{4\pi} M_{\text{KK}}} \right)^{3/2} \approx 6.2 \times 10^{12}, \quad (\text{V.1})$$

where we have taken $g_s = 0.5$, and $M_{\text{KK}} = 10^9 \text{ GeV}$ to incorporate inflationary scales as low as this (since

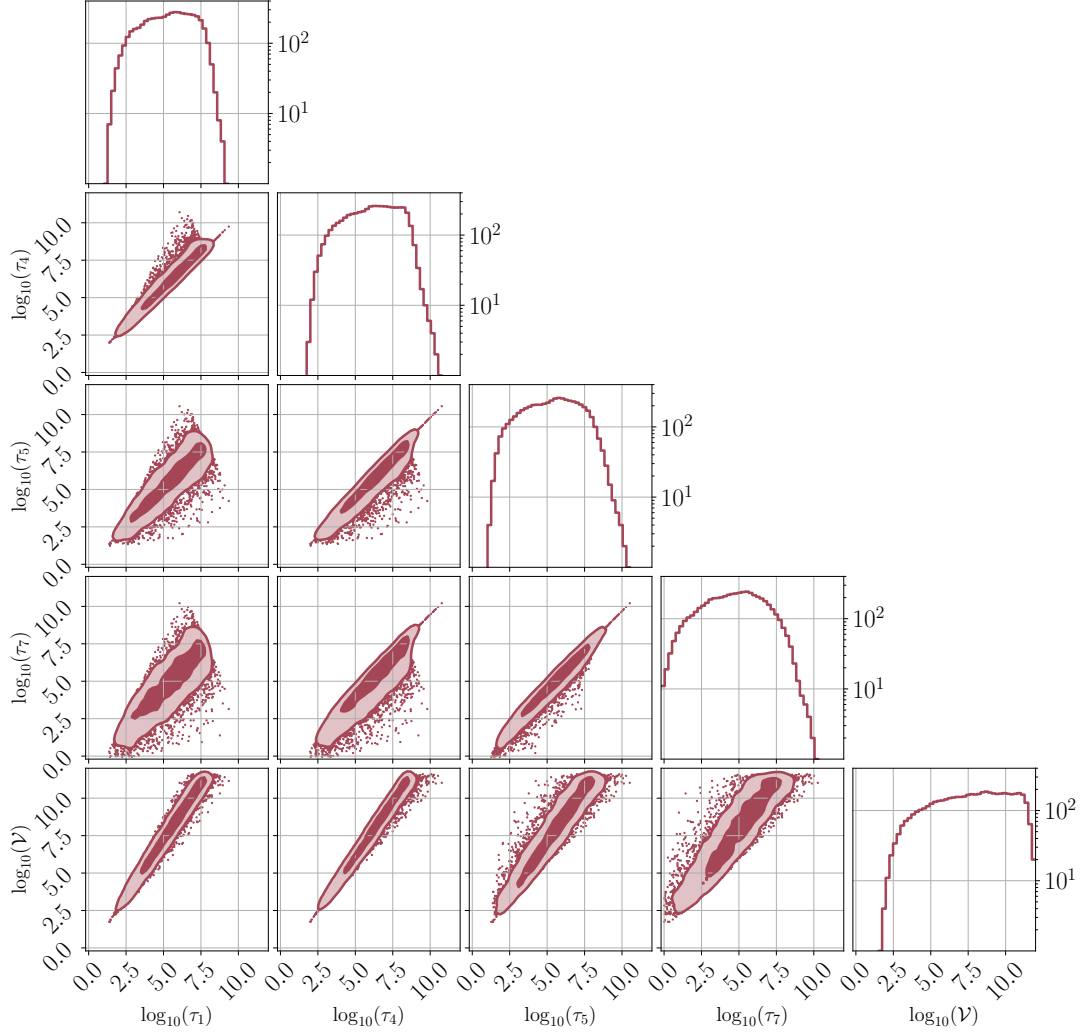


FIG. 9. Distributions of prime toric divisor volumes and overall volume under Weil–Petersson measure for an example geometry at $h^{1,1} = 3$ (for which there were 4 unique prime toric divisor classes).

$H_{\text{inf}} \geq M_{\text{KK}}$). The choice of $\mathcal{V}_{\text{max}}$ is a uniform and physically consistent upper bound across geometries (in contrast to the scaled value of $\mathcal{V}_{\text{tip}}$ used in the previous section to more fairly measure computational efficiency across geometries).¹²

A. Theory informed priors on (m_a, f_a)

Here, we present the collective distribution of axion masses and decay constants for all inequivalent FRSTs

with $h^{1,1} = 3$.¹³ On top of the WP prior on the Kähler moduli space, we implement a likelihood for a QCD-like axion. For each choice of Kähler parameters, we scan over all prime toric divisors and identify the one whose volume is closest to 40 (in string units). This is to achieve a realistic QCD gauge coupling in the UV.

The explicit divisor-volume likelihood that we impose on this closest-volume prime toric divisor is

$$\ln \mathcal{L}_{\text{vol}} = -\frac{1}{2 \times 5^2} (\tau_{\text{QCD}} - \bar{\tau}_{\text{QCD}})^2, \quad (\text{V.2})$$

with $\bar{\tau}_{\text{QCD}} = 40$. We then identify the instanton action associated to this divisor, and replace it with the QCD instanton action as discussed in Section A.

¹² We note that $\mathcal{V}_{\text{max}}$ in Eq. (V.1) inherently assumes an isotropic CY hypersurface, an assumption that needs to be modified for anisotropic CY hypersurfaces (such as a K3 fibered CY). We briefly comment further on this in Section C.3, but for our purposes in this paper, we use the simple estimate as the only purpose is to assign a well-defined and reproducible number to each geometry we study.

¹³ We provide distributions for $h^{1,1} = 4$ and 5 in the supplementary material .

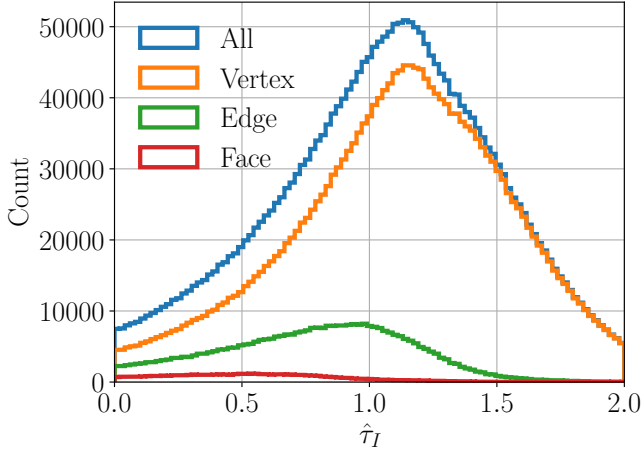


FIG. 10. Distribution of logarithmically normalised volumes $\hat{\tau}$ (see Eq. (IV.2)) for vertex, edge, and face prime toric divisors marginalised the random geometries of Section III A 1 at $h^{1,1} = 10$.

In Figs. 12 and 13 we display the collective distribution of masses and decay constants for *all* axions across all inequivalent FRSTs with $h^{1,1} = 3$, as obtained from our MCMC sampling. We adopt the hierarchical approximation (see Section A and also [17]) to compute the axion spectrum, under which leading instanton directions coincide with axion mass eigenstates. It is then useful to further split the spectrum according to the geometric rigidity of the divisor classes supporting the instantons, following the discussion of decay constants in Section II.C.

In Fig. 12, we do not impose any QCD-axion likelihood. The spectrum therefore contains all three axions for every sampled point across all Kähler cones. For both geometrically rigid and movable divisors, the distribution on axion masses has a logarithmically flat behaviour along with a slight blue tilt. The corresponding decay constants for both type of divisor classes, on the other hand, are localised and do not spread across many orders of magnitude. Still, we observe that geometrically rigid divisors ($h^{2,0}(D) = 0$) can support lower decay constants compared to movable divisors ($h^{2,0}(D) \geq 1$). We recall that these two trends were explained geometrically in Section II.C.

In contrast, in Fig. 13 we impose the divisor-volume Gaussian likelihood (on top of the WP prior), and we show the spectrum of both the QCD-like axion and the remaining stringy axions (ALPs). While many qualitative features of the string-axion spectrum remain unchanged (left panel), two effects stand out. First, there is a clear ‘dip’ in the ALP mass distribution around $\sim 10^{-55} M_{\text{Pl}}$. This occurs because the degree of freedom that would otherwise contribute an ALP at that mass has instead been modified to become the QCD-like axion, along with the mass distribution associated with movable divisors becoming noticeably flatter over-

all. Second, a QCD-like axion engineered on a movable divisor clusters in the neV mass range, whereas a significantly heavier QCD-like axion can arise when realised on a rigid divisor (right panel). This is because rigid divisors can support parametrically smaller decay constants. We return to this feature in more detail in Section V.B.

We now employ a NF model to learn the ALP spectra without imposing any QCD likelihood Eq. (V.2). This allows us to extract marginalised distributions for the ALP masses and decay constants, yielding physically meaningful priors that can be directly used in cosmological and astrophysical analyses.

As earlier, we employ a real NVP model. In this case our model is comprised of 8 masked affine flow layers where the translation and scale networks are MLPs consisting of 2 hidden layers each with 128 neurons, which gives a total of 274,496 learnable parameters. We train our model over 1,000 epochs using the same optimisation procedure as before to minimise the forward KL divergence Eq. (B.4) computed from a batch of 512.

In Fig. 14, we show the initial Gaussian base distribution $\mathcal{N}(0, I)$ alongside the learned distribution for 50,000 generated samples. Comparing this with the full spectra shown in Fig. 12 we can see that the NF model has successfully learned the (m_a, f_a) distribution even down to the tails at lower values in the decay constant.

While we do not attempt to learn the spectra shown in Fig. 13 here, we note that one could train a NF on the joint distribution of (1) the QCD-like axion mass and decay constant, (2) the volume of the QCD divisor, and (3) the ALP masses and decay constants, and then train a second NF model on the same variables but with the QCD-divisor volume likelihood imposed. In principle, the latter could then be used to infer the distribution appearing in Fig. 13. We leave this experiment to future work.

B. Inferring topology with haloscopes

In this section we examine more specifically how the mass of the QCD axion is related to the topology of the underlying CY manifold, with QCD realised on a stack of D7-branes (see Appendix A). While divisor volumes control the mass scales of generic axions receiving masses from ED3 instantons, the mass of the QCD axion, m_{QCD} , is determined only by the decay constant f_a once the D7-brane divisor volume has been fixed to give the desired ultraviolet gauge coupling.

There is a strong statistical correlation between Hodge numbers and axion decay constants, as seen in the existing literature: larger values of $h^{1,1}$ typically give smaller decay constants [15, 17, 19, 77]. The TSKC provides an indicator of how small the Kähler parameters can collectively become, and therefore of the largest possible decay constants, at fixed $h^{1,1}$. This implies that at each value of $h^{1,1}$ there is a maximum decay constant and a corresponding minimum m_{QCD} , both set roughly by the CY

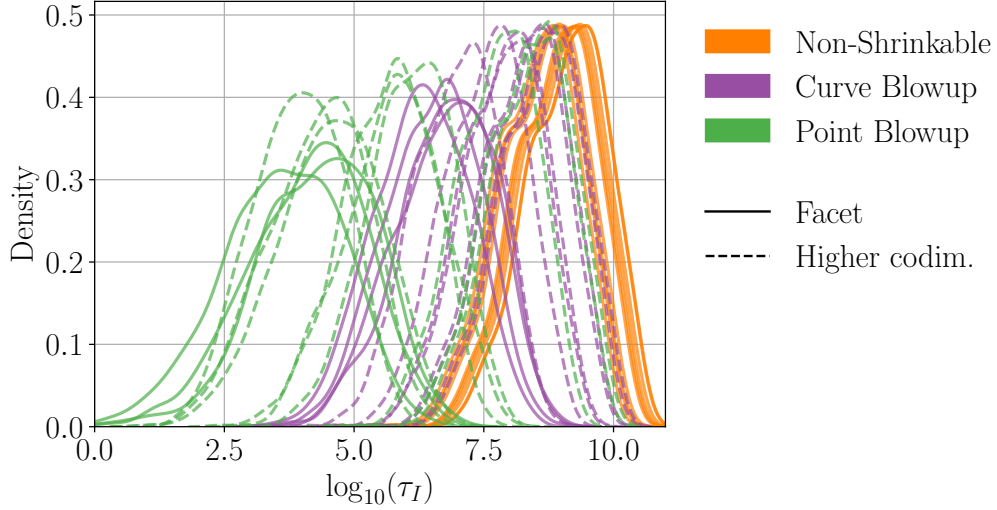


FIG. 11. Distribution of prime toric divisor volumes marginalised over the WP measure within the SKC for a geometry at $h^{1,1} = 30$ with a simplicial $\mathcal{K}_U$, categorised according to the topological category of the divisor.

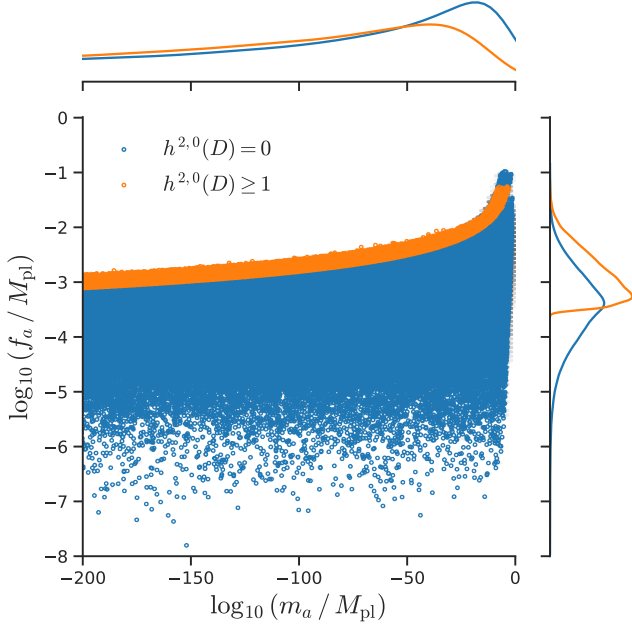


FIG. 12. Spectra of all axions (for $h^{1,1} = 3$) under the WP prior, without any imposition of a gauge-theory (QCD-like) axion. We also separate the distributions based on geometric rigidity of the three divisor classes that host the corresponding instantons.

volume at the tip, $\mathcal{V}_{\text{tip}}$. If one moves along a pure dilation of the TSKC, all decay constants decrease and all divisor volumes increase. Demanding that there exists a divisor with volume close to 40 in string units, in order to give the desired gauge coupling, and restricting attention to such dilations, leads to a minimum value of f_a and thus a maximum value of m_{QCD} . These facts underlie the correlation between QCD axion mass and topology discussed

in Refs. [18, 21].

However, as we constructed explicitly in Ref. [16], and informed by the geometric picture we further developed in Secs. II and IV, when exploring the full moduli space, one can find loci where the QCD divisor volume remains 40 but the overall volume is parametrically larger than it was on the TSKC ray, resulting in much smaller decay constants.¹⁴ In this case the naive topology–mass correlation can be evaded, provided the geometry contains divisor classes whose volumes are effectively decoupled from the overall CY volume. With this in mind, we revisit some aspects of Ref. [18] and ask to what extent axion haloscopes can make inferences about CY topology.

For CY threefolds with small $h^{1,1}$, the decay constants are typically large in regions that are not deep inside the Kähler cone, with representative values $f_a \gtrsim 10^{15}$ GeV. This leads, in general, to a light QCD-like axion in the neV range of masses. As an explicit example, for our ensemble with $h^{1,1} = 3$ this generic behaviour is visible in the right panel of Fig. 13. The orange spectra, which correspond to instantons on divisors with $h^{2,0}(D) \geq 1$, cluster in the neV range. These non-rigid divisors have volumes that closely track the overall Kähler scale and therefore do not permit sufficiently small decay constants for a heavier QCD-like axion at fixed gauge coupling.

In contrast, the rigid prime toric divisors with $h^{2,0}(D) = 0$ can support a high-mass QCD-like axion. This is the phenomenon anticipated at the end of the previous subsection: the blue spectra in the right panel

¹⁴ In a similar view, there could also *not* be a maximum far below the Planck scale. For instance K3 divisor classes may provide for large axion decay constants, in regions close to faces of the stretched Kähler cone.

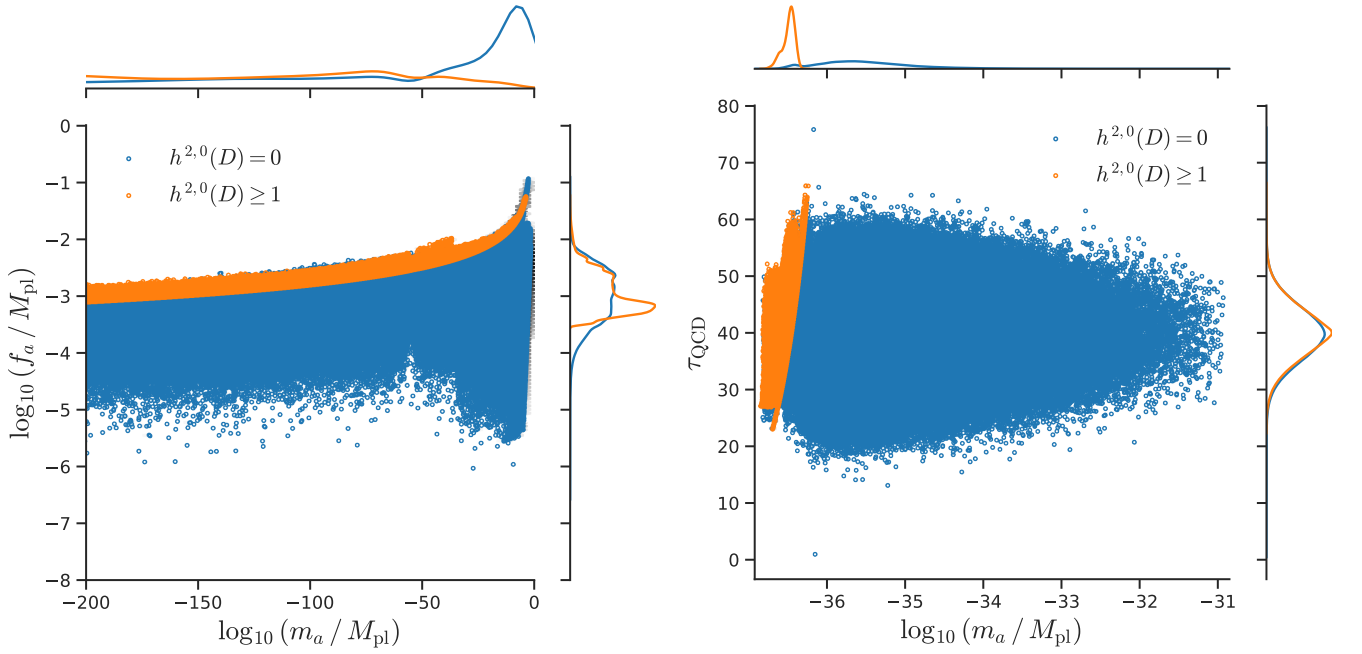


FIG. 13. Spectra of all axions (for $h^{1,1} = 3$) under the WP prior, with identification of a gauge-theory (QCD-like) axion according to the likelihood (V.2). Compare against Fig. 12 which shows the case with no such identification. The left panel shows the distributions of stringy axions (ALPs), while the right panel shows the spectrum of the QCD-like axion.

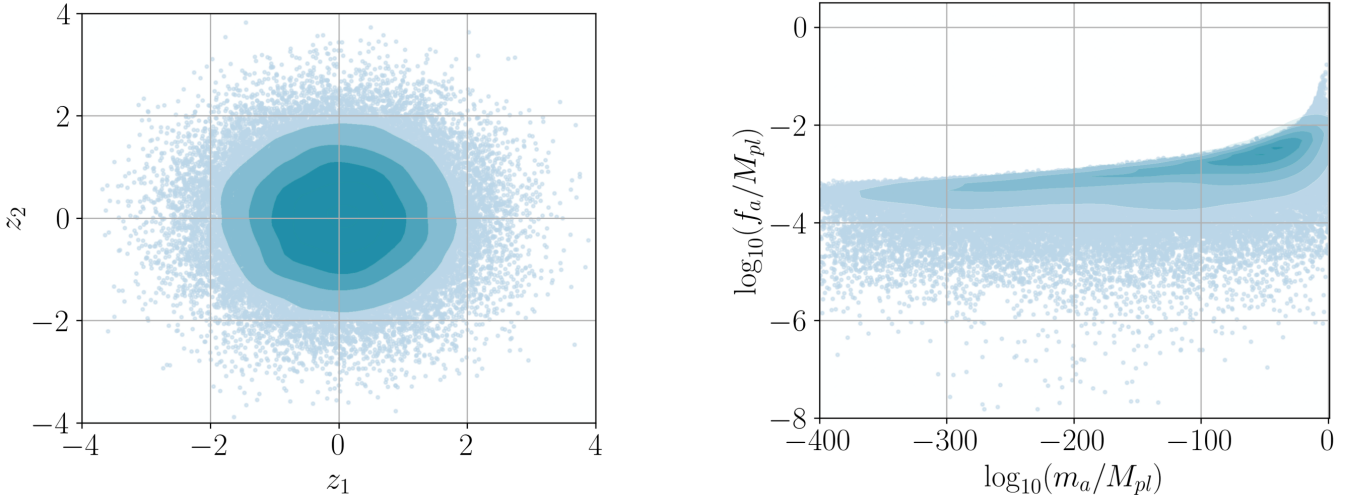


FIG. 14. Shown are 10,000 samples drawn from the base Gaussian $\mathcal{N}(0, I)$ (left) and the corresponding distribution obtained after transforming these same samples through the normalising flow trained on the MCMC sample set in Fig. 12 (right).

of Fig. 13 arise from instantons on rigid divisors and populate the high-mass tail of the QCD-axion distribution. For small $h^{1,1}$, such rigid divisors can remain at moderate volume even in vacua that lie deep inside the Kähler cone, so that some (or in special cases all) axionic directions acquire parametrically small decay constants. This may allow one axion to realise both a large mass and an appropriate gauge coupling. The MCMC analysis of Sec. IV shows that exactly these rigid divisors are responsible for the high-mass QCD-like axions in our sample.

In summary, achieving a heavy QCD-like axion requires simultaneously controlling the divisor volumes and accessing regions of moduli space where certain decay constants become parametrically small. These two ingredients do not always coexist, and only arise when the underlying geometry satisfies a pair of sharp geometric conditions. For clarity, we summarise them here:

- The vacuum must lie deep inside the (stretched) Kähler cone; i.e., the overall CY volume is large.

- The geometry must contain at least one prime toric divisor whose constant $\tau(\sim \bar{\tau}_{\text{QCD}})$ level sets extend towards this vacuum.

These two conditions together allow one axionic direction to realise both a parametrically small decay constant, yielding a heavy QCD-like mass, and a suitably small divisor volume that sets the gauge coupling.

Instantons on non-K3 shrinking divisors serve as possible candidates.¹⁵ This leaves us with two possible divisor types: (a) blow-up divisors, which can shrink while the CY has finite volume; (b) fibration divisors, which can only shrink if the whole CY also shrinks.

In general for $h^{1,1} = 3$, we find that $\sim 90\%$ of inequivalent FRSTs from the KS database either have a blow up prime-toric divisor or an elliptic fibration. One then needs to exploit this structure. That is, be able to host the QCD instanton on a special shrinking divisor and also take an appropriate limit in the moduli space. We illustrate this by examining two explicit geometries at $h^{1,1} = 3$, each exhibiting these types of special divisors: a non-fibered geometry with blow-up divisors that shrink on a facet and on a ray respectively, and an elliptically fibered geometry with non blow-up shrinking divisors.

For each of the two cases, we perform seven MCMC runs: one for every choice of the prime toric divisor for hosting QCD. We take a log-normal likelihood for the QCD axion mass centred at $\bar{m}_{\text{QCD}} = 6 \mu\text{eV}$ with sufficiently narrow width, and a Gaussian likelihood for the divisor volume centred at $\bar{\tau}_{\text{QCD}} = 40$ with a sufficiently narrow width. Explicitly:

$$\ln \mathcal{L} = \ln \mathcal{L}_{m_{\text{QCD}}} + \ln \mathcal{L}_{\text{vol}}, \quad (\text{V.3})$$

$$\text{where } \ln \mathcal{L}_{m_{\text{QCD}}} = -\frac{1}{2 \times (0.5)^2} \left[\log_{10} \left(\frac{m_{\text{QCD}}}{\bar{m}_{\text{QCD}}} \right) \right]^2, \\ \ln \mathcal{L}_{\text{vol}} = -\frac{1}{2 \times 5^2} (\tau_{\text{QCD}} - \bar{\tau}_{\text{QCD}})^2.$$

This likelihood is meant to model a putative detection of the axion by the haloscope ADMX [78].¹⁶ ADMX is sensitive in GHz frequencies, which count as “high” for $h^{1,1} = 3$ CYs, i.e., not in the range of $h^{1,1}$ found to be accessible to ADMX by dilation of TSKC in Ref. [18].

1. Blow-up divisors

We consider a 4D reflexive polytope Δ° with points given by the columns of

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{V.4})$$

along with the following gauged linear sigma model (GLSM) charge matrix (which we take as the instanton charge matrix following e.g. [15]):

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \quad (\text{V.5})$$

A choice of height vector induces a triangulation that defines a toric variety in which the generic anticanonical hypersurface is a smooth CY threefold with $h^{1,1} = 3$. The columns of the GLSM/instanton charge matrix correspond to prime toric divisor classes, and in particular our focus is on D_5 and D_7 . There are 3 extremal facets of the $K_\cup$ given by the rows of

$$\mathcal{H} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix}, \quad (\text{V.6})$$

and the two divisor volumes take the form

$$\tau_5 = \frac{1}{2} (t_1^2 + t_2 + 4t_2t_3 - 4t_3^2 + t_1(4t_3 - 6t_2)), \\ \tau_7 = (t_1 + t_2 - 2t_3)^2. \quad (\text{V.7})$$

Note that D_7 shrinks to a point along the cone facet whose normal is $(1, 1, -2)$, while D_5 shrinks to a point along the codimension-two ray $(1, 1, 1)$. Both divisors correspond to blow-up loci, in the sense that they can shrink while the CY volume remains finite.¹⁷ In Fig. 15 we show the derived parameters of QCD axion mass and divisor volumes, for both of the two choices of D_5 or D_7 hosting the QCD gauge group. To compare with the other five possibilities, in Fig. 16 we show the posterior probability density for all seven prime toric divisors. The posterior probability is significantly higher for QCD realised on D_5 or D_7 , indicating that an ADMX detection would prefer these discrete choices over the others available.

Both figures also illustrate that, between D_5 and D_7 , D_7 is the more suitable candidate. Its shrinking locus is a codimension-one surface (a facet), i.e. a higher-dimensional boundary of the Kähler cone, whereas D_5

¹⁵ The decay constants of axions that are associated with instantons on K3 fibers, depend solely on their volumes (i.e. $f \sim 1/\tau$). Requiring volume to be close to 40 then, necessarily enforces a large decay constant. See Section C3c for a discussion.

¹⁶ Computationally we were forced to impose a relatively wide likelihood on $\log_{10} m_{\text{QCD}}$, while in reality the axion mass will be measured with a relative accuracy around $\sigma_m/m_{\text{QCD}} \approx 1/Q \approx 10^{-5}$ by ADMX, where Q is the cavity quality factor.

¹⁷ This geometry is in fact of ‘Swiss-Cheese’ type, and the volume form has a simple decoupled structure $\mathcal{V} = \sqrt{8/5} \tau_6^{3/2} - \sqrt{1/2} (\tau_6 - \tau_4)^{3/2} - \tau_7^{3/2}$. Such a simplicity, however, is not essential for our purposes.

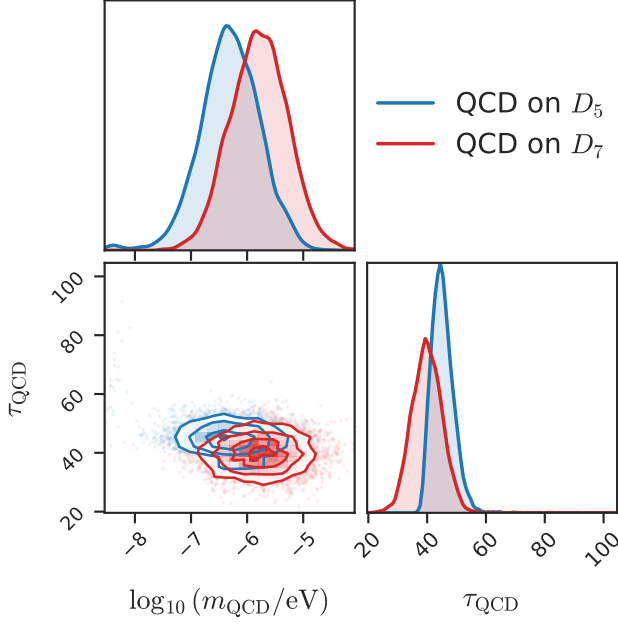


FIG. 15. Posterior distributions of the QCD-axion mass and the corresponding QCD divisor volume for the two blow-up divisor choices of the geometry defined in (V.4). The results are obtained using the likelihood (V.3) with fiducial values $\bar{\tau}_{\text{QCD}} = 40$ and $\bar{m}_{\text{QCD}} = 6 \mu\text{eV}$.

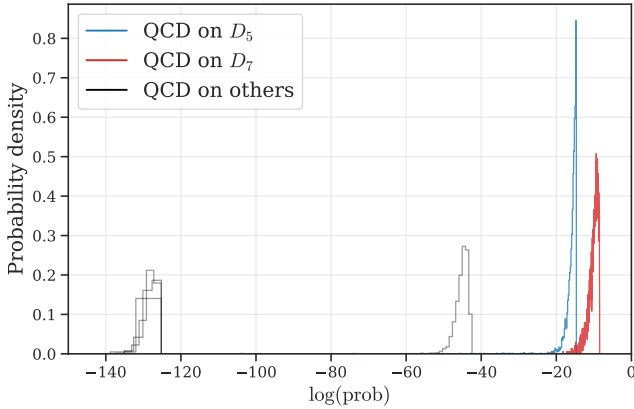


FIG. 16. Posterior probability density for all seven choices of prime toric divisors, given a likelihood on the QCD-axion mass peaked at $6 \mu\text{eV}$ and its divisor volume peaked at 40 (in string units). (Among the others, the $\log(\text{prob})$ values for D_6 are too small and are not shown here).

shrinks only along a codimension-two locus (an extremal ray), which is correspondingly smaller. The remaining prime toric divisors are not shrinkable.

2. Non-blow-up shrinking divisors

We consider a 4D reflexive polytope Δ° whose points are given by the columns of

$$\begin{pmatrix} 1 & -5 & 0 & 0 & 0 & -3 & -2 & -1 \\ 0 & -1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & -2 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}, \quad (\text{V.8})$$

along with the following instanton charge matrix

$$\begin{pmatrix} 5 & 1 & 1 & 2 & 1 & 0 & 0 \\ 3 & 0 & 1 & 1 & 0 & 1 & 0 \\ -3 & -1 & -1 & -2 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{V.9})$$

A choice of height vector induces a triangulation that defines a toric variety in which the generic anticanonical hypersurface is a smooth CY threefold with $h^{1,1} = 3$, whose volume is

$$\mathcal{V} = \frac{1}{6}(t_1 + t_3)(t_1(6t_2 + t_3) - t_1^2 - 6t_2^2 - 6t_2t_3 - t_3^2). \quad (\text{V.10})$$

The columns of the instanton charge matrix correspond to prime toric divisor classes, and in particular our focus is on D_2 and D_6 whose volumes take the form

$$\begin{aligned} \tau_2 &= \frac{1}{2}(t_1 + t_3)(-t_1 + 4t_2 + t_3), \\ \tau_6 &= (t_1 + t_3)(t_1 - 2t_2 - t_3). \end{aligned} \quad (\text{V.11})$$

There are three extremal facets of $K_\cup$, given by the rows of

$$\mathcal{H} = \begin{pmatrix} -1 & 2 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 1 \end{pmatrix}. \quad (\text{V.12})$$

From $\mathcal{V}$ and $\mathcal{H}$ it is clear that the facet of $K_\cup$ with normal vector $(1, 0, 1)$ corresponds to an elliptic fibration. Both D_2 and D_6 shrink to curves along this facet. Furthermore, they each shrink to points along codimension-two rays of $K_\cup$: D_2 shrinks along the intersection of this facet with the facet normal to $(-1, 2, 0)$, while D_6 shrinks along the intersection with the facet normal to $(0, -1, -1)$. These loci, where the two divisors shrink, are precisely the regions where a high-mass QCD instanton can be supported.

In Fig. 17 we show the derived QCD-axion mass and divisor volume for the two choices D_2 and D_6 hosting the QCD gauge group. To compare with the five remaining choices, in Fig. 18 we show the posterior probability density for all seven possibilities.

To connect with the previous example, it is not surprising that in this case both D_2 and D_6 are equally good candidates (and exhibit similar statistics), as each shrinks along a codimension-two locus.

We finally note that in this example there are two additional divisors (D_3 and D_4) that shrink to curves

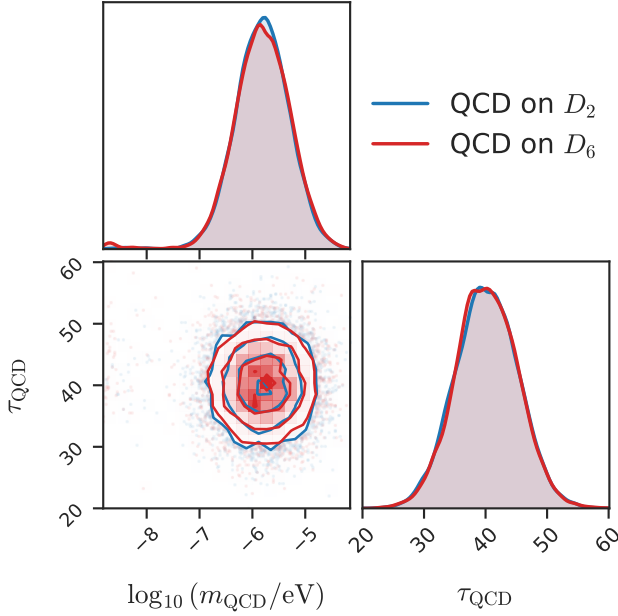


FIG. 17. Posterior distributions of the QCD-axion mass and the corresponding QCD divisor volume for the two non-blow-up shrinking divisor choices of the geometry defined in (V.8). The results are obtained using the likelihood (V.3) with fiducial values $\bar{\tau}_{\text{QCD}} = 40$ and $\bar{m}_{\text{QCD}} = 6 \mu\text{eV}$.

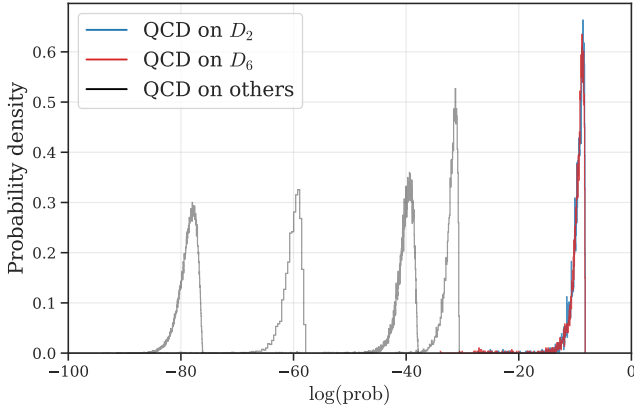


FIG. 18. Posterior probability density for all seven choices of prime toric divisors, given a likelihood on the QCD-axion mass peaked at $6 \mu\text{eV}$ and its divisor volume peaked at 40 (in string units). (Among the others, the $\log(\text{prob})$ values for D_1 are too small and are not shown here).

along the fibration facet. However, there is no locus where they can shrink to a point, and it turns out that they cannot support a high-mass QCD-like instanton. This aligns with Fig. 8, where we see that point blowup divisors are capable of smaller volumes at larger overall volume than curve blowup divisors.

We also performed a subset of the analysis in this subsection with $\bar{m}_{\text{QCD}} = 20 \text{ neV}$ to model a putative detec-

tion of the QCD axion by DMRadio [79]. Since masses in the neV range are “typical” for dilations of the TSKC at $h^{1,1} = 3$, the DMRadio analysis yielded trivial results in which QCD could be realised on any type of divisor with high posterior probability. We have thus shown that, for CYs with low $h^{1,1}$, a detection by DMRadio would give little to no specific topological information; by contrast, a detection by ADMX has greater constraining power, statistically preferring certain CY topologies and divisor choices to host QCD.

C. Fuzzy dark matter and the Lyman-alpha forest

Ultralight axions with $m_a \lesssim 10^{-17} \text{ eV}$ differ from standard cold DM (CDM) in their effects on cosmological observables. Ultralight axions suppress matter clustering below the de Broglie wavelength (which sets an effective Jeans scale), leading to a reduced amplitude of the linear matter power spectrum $P(k)$ on small scales [9, 80–84], often dubbed “fuzzy” DM (see reviews in Refs. [14, 85, 86]). Cosmological observables are each individually consistent with the predictions of the CDM model. Thus, the absence of such suppression in $P(k)$ (and other effects) lead to upper limits on the allowed ultralight axion density $\rho_a = \Omega_a h^2 \times 8 \times 10^{-11} \text{ eV}^4$ [80, 82, 84, 87–92]. When two observables, however, yield inconsistent inferences within the ΛCDM model, this implies either unmodeled astrophysical or instrumental systematics, or physics beyond CDM [e.g., 57, 91]. Ref. [57] finds that a 5σ discrepancy in the inference of the small-scale (wavenumber $k \sim 1 h \text{ Mpc}^{-1}$), high-redshift ($z \sim 3$) $P(k)$ given *Planck* CMB and eBOSS Lyman-alpha forest (Ly α F) data can be resolved by a contribution of ultralight axions (although see Refs. [93, 94]).

It is somewhat difficult to find CYs in type IIB that give rise to a large fuzzy DM relic density. Only a sub-percent fraction of the total number of CYs with $h^{1,1} \leq 7$ yield axion models within reach of current and upcoming cosmological observations [16, 95]. In Ref. [16], we constructed a specific example of fuzzy DM at $h^{1,1} = 7$ that has the largest misalignment abundance of the lightest axion in the model, while also containing a divisor of volume near 40 to host QCD.¹⁸ Upon dilating the CY volume, this model furnishes a combination of (m_a, Ω_a) near the maximum posterior probability of these parameters inferred from the joint analysis [57] of *Planck* CMB

¹⁸ The fuzzy axion with a large decay constant in this example corresponds to a $K3$ fibration. Interestingly, though, the associated divisor D is not quite the $K3$ fiber, but rather a dP_9 that is “half” of the $K3$ fiber (see, e.g., footnote 19 in [96]). The canonical class of dP_9 is pure torsion, so $D \cap D$ still vanishes numerically and the discussion of $K3$ divisors and the Kähler metric of Section II.C applies to D , explaining the large decay constant. Usefully, dP_9 is also rigid, so it does not carry neutral fermionic zero modes that would obstruct a nonperturbative contribution to the superpotential.

and eBOSS Lyman- α forest data (alternative new physics resolutions to this parameter discrepancy include running of the primordial power spectrum spectral index, e.g. Refs. [57, 97]). Thus, this particular CY model is a possible candidate to explain the discrepancy between these datasets.

We now employ the tools we have developed in the previous sections to sample the moduli space of this example CY, with the goal of inferring the region of moduli space preferred by a solution to the CMB-Ly α F discrepancy. We sample the CY moduli space using the Kähler parameters and the WP prior, and compute the axion masses m_a and decay constants f_a that arise at each point. We impose a prior that there exists a divisor with volume close to 40 [Eq. (V.2)] and place QCD there in order to have the correct effect of the QCD axion mass on the overall spectrum of masses and decay constants.

We simultaneously sample cosmological parameters ($\omega_{\text{DM}}, \log \frac{\Omega_a}{\Omega_{\text{DM}}}, \omega_{\text{baryon}}, h, A_s, n_s, \tau$). These parameters are respectively: the total DM physical density, the logarithm of the fraction of the DM composed of ultralight axions, the physical baryon density, the amplitude and spectral index of the primordial power spectrum and the optical depth to reionization. Cosmological parameter priors are uniform other than a Gaussian prior on the optical depth to reionization used as a compression of the *Planck* *SRoll2* low-multipole polarization data [98]. The only axion whose physics we treat precisely is the lightest axion in the spectrum, which in the parameter space of interest has $m_a \in [10^{-27}, 10^{-23}] \text{ eV}$. For this axion, we use the Einstein-Boltzmann solver AXICLASS [99–102] to solve the Klein-Gordon equation for the relic density and the evolution of perturbations. We thus compute the CMB temperature T and polarisation E angular power spectra ($C_\ell^{TT}, C_\ell^{TE}, C_\ell^{EE}$) and the linear matter power spectrum $P(k)$.¹⁹ The effect of all other axions is taken into account only through the CDM density Ω_c . As discussed in detail in Ref. [16], avoiding overproduction of CDM from heavier axions including the QCD axion implies a fine tuning of the initial misalignment for these axions and/or modified cosmology (low reheating scale), which we do not address here. We set the isocurvature amplitude (CDM and ultralight axion) to zero (for details on ultralight axion isocurvature, see Refs. [106, 107]). Inflation is modeled only via the effective parameters A_s and n_s . This simplified model with minimal tuning of the heavy axion misalignment angles and vanishing isocurvature is equivalent to an assumption of a low inflationary Hubble scale [85, 108]. Going beyond these minimal assumptions is beyond the scope of the present work but can be addressed in future.

We jointly sample the posterior distribution of the Kähler and cosmological parameters given *Planck* CMB and eBOSS Ly α F data using the Markov chain Monte Carlo sampler EMCEE [71]. For the CMB temperature and polarization data, we use the compressed version PLANCK-LITE-PY [109] of the *Planck* 2018 low-multipole TT and high-multipole TT, TE, EE data [110] (as discussed above, the low-multipole polarization data are compressed to the prior on τ). For the eBOSS Ly α F 1D flux power spectrum data [111], we use the compressed version [57, 112] of a Gaussian likelihood on the amplitude and spectral index of the linear matter power spectrum $P(k = 0.009 \text{ s km}^{-1}, z = 3)$. Both these compressions are lossless for the axion cosmology we consider here. For comparison, we also perform a second analysis where we do not sample the Kähler parameters and instead sample $(\log m_a, \log f_a)$ with an upper limit on f_a at the reduced Planck scale. This second analysis matches previous fuzzy DM studies and we dub the log-uniform prior on axion mass and decay constant an effective field theory (EFT) prior.

Fig. 19 (left) shows the marginalized posterior distribution of the axion mass and DM fraction. The posteriors of the EFT and CY analyses are consistent. The EFT analysis (purple contours) reproduces the results of Ref. [57], which finds a preference for $\sim (1 - 3)\%$ DM abundance in an axion with $10^{-26} \text{ eV} \lesssim m_a \lesssim 10^{-23} \text{ eV}$. The CY analysis similarly finds a posterior peak in this region. However, the CY analysis (orange contours) only allows for $10^{-25} \text{ eV} \lesssim m_a \lesssim 10^{-23} \text{ eV}$. The reason for the CY result differing from the EFT result is shown in Fig. 19 (right), which shows the marginalized posterior of the axion decay constant. The EFT analysis allows decay constants up to the Planck scale, which allows for larger axion DM abundance at lower mass. In contrast, the CY model does not allow for such large decay constants and the posterior is peaked near $f_a \sim 10^{-2} M_{\text{Pl}}$. This is the largest f_a in the moduli space of the CY that we consider (given the prior and likelihood above). This result is consistent with the indications in Refs. [16, 113] concerning the maximum DM abundance of ultralight axions.

The power of this Bayesian analysis is demonstrated in Fig. 1, which shows the marginalized posterior of the CY volume. Compared to the prior alone, the cosmological data lead to a sharply peaked posterior on the CY volume. In fact, the posterior of all seven Kähler parameters is sharply peaked in comparison to the prior. This analysis demonstrates that it is possible to perform a rigorous statistical analysis using modern precision cosmological data to find the maximum posterior probability in CY moduli space, while jointly sampling cosmological model parameters.

VI. CONCLUSIONS

In this work we have presented steps towards a long held ideal in string phenomenology: the ability to make

¹⁹ With m_a and f_a as input, AXICLASS uses a shooting method to find the initial misalignment angle $\langle \theta_i \rangle$ necessary to obtain the desired relic abundance Ω_a . For a given f_a , this method only gives a solution for some minimum m_a since arbitrary fine-tunings to $\langle \theta_i \rangle = \pi$ (“extreme axion” models, e.g., Refs. [103–105]) are disallowed. This sets an effective prior $\langle \theta_i \rangle \sim \mathcal{O}(1)$.

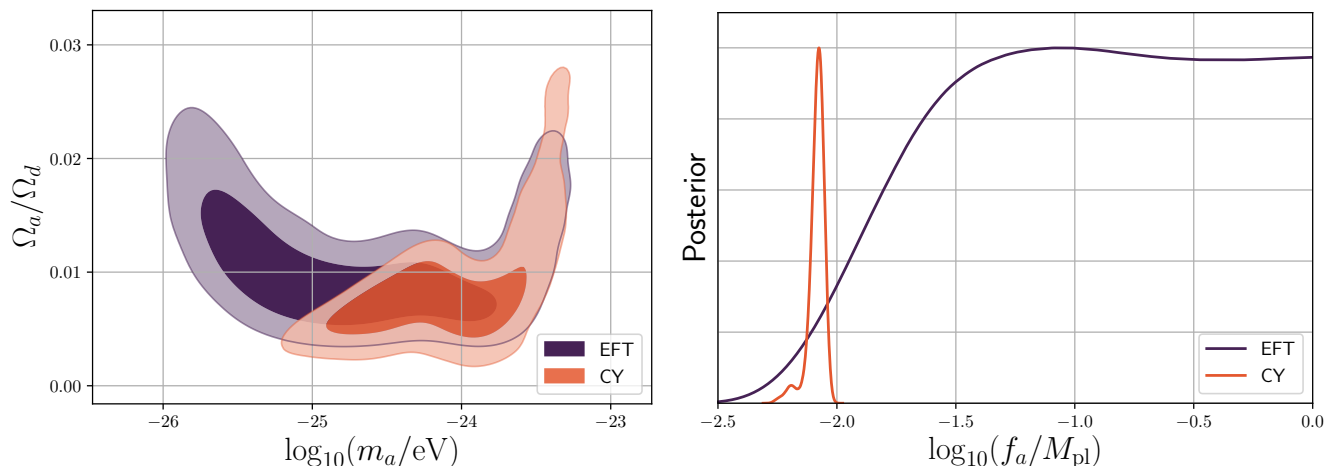


FIG. 19. *Left*: the marginalized posterior distribution of axion mass m_a and axion DM fraction $\frac{\Omega_a}{\Omega_{\text{DM}}}$. We indicate by the darker and lighter contours respectively the 68% and 95% credible regions. *Right*: the marginalized posterior distribution of axion decay constant f_a . The likelihood is a product of the *Planck* 2018 cosmic microwave background temperature and polarization likelihood and the eBOSS Lyman-alpha forest 1D flux power spectrum compression. We show both an analysis that samples the CY moduli space for a model with $h^{1,1} = 7$ with the WP prior and the QCD divisor volume prior, and an EFT analysis where the mass and decay constant of a single ultralight axion are independent free parameters with a log-uniform prior. In a pure CDM model, there is a parameter discrepancy between the CMB and Ly α F data, which can be resolved by an ultralight axion causing a suppression in the linear matter power spectrum, leading to a posterior peak with non-zero Ω_a . We find that sampling the CY moduli space for a chosen axion model prefers a lower- f_a part of the posterior.

robust statistical inference about string theory using observations. Steps towards this lofty goal required us to begin with many simplifications, the first of which was to take a “modular approach” to phenomenology, treating the vevs of the Kähler moduli as continuous free parameters. With this assumption, we developed robust methodology to sample from a special diffeomorphism and scale invariant (and thus uninformative in the Bayesian sense) prior on moduli space: the Weil–Petersson (WP) prior. We used the classical Kähler potential to construct the WP prior: we leave a study of quantum corrections (see, e.g., Ref. [31]) to future work.

We employed two sampling methods: affine invariant MCMC, and normalising flows. With a laptop GPU, MCMC can achieve statistically reliable, converged samples from WP in geometries with up to 30 Kähler moduli in a modest amount of time. Sampling was also fast enough that we could exhaustively sample the WP measure in all CYs derivable from the Kreuzer–Skarke database with $h^{1,1} \leq 5$ with very modest (laptop time) computing resources. Using normalising flows, we can achieve sampling consistent with MCMC. Computation is more intensive, but once learned, sampling from the distribution can be applied again and again “for free”, and such methods may therefore be of use in the future.

With a statistically robust method of moduli space sampling, we first explored distributions of geometric data of CY manifolds: namely, volumes of divisors and correlations between volumes of divisors and of the CY. In particular, we illustrate how the topology of a divisor — especially how it shrinks along boundaries of

the Kähler cone — controls the statistics of its volume distribution. This includes ensembles constructed by marginalising over geometries as well as over the prior within specific geometries, including a geometry at $h^{1,1} = 30$ optimised by a genetic algorithm [24] to have simplicial $K_{\square}$. Additionally, we verify the expectation (stated, for example, in Ref. [15]) that divisor volumes in the SKC tend to become larger as $h^{1,1}$ increases, and we show that the observations of Ref. [23] regarding distributions of divisor volumes under a particular normalisation persist for samples from the WP prior.

We applied our sampling methodology to the physics of axions. We first presented the theory informed prior on the joint distribution of axion masses and decay constants, marginalised over the WP measure for all FRST classes of CYs constructable from the Kreuzer–Skarke database at $h^{1,1} = 3, 4$, and demonstrated that this distribution can be learned using normalising flows.

We next employed Bayesian inference with mock likelihoods designed to simulate the possible future detection of the QCD axion with a DM haloscope, extending the results of Ref. [18] across moduli space. We showed, in explicit examples at $h^{1,1} = 3$, that detection of a relatively high-mass axion by ADMX is compatible only with certain CY topologies. Although these topologies comprise $\sim 90\%$ of all FRST classes at $h^{1,1} = 3$, they nonetheless require the engineering of QCD-like instantons on special divisors. Within these topologies, such a detection would therefore constrain which divisor can host QCD on D7-branes and would localize the vacuum to a region of moduli space close to a wall of the Kähler cone. The

conclusion is powerful: detection of a high-mass QCD axion can provide detailed topological information about CYs with low Hodge number. It would be interesting to study the converse phenomenon at large $h^{1,1}$: for topologically complex CY manifolds, can one find decay constants parametrically larger than their typical TSKC values, thereby giving rise to light QCD axions at high $h^{1,1}$?

In the final part of this paper, we performed Bayesian inference on moduli space with observationally derived cosmological likelihoods. Specifically, we used the *Planck* CMB anisotropies and eBOSS Ly- α forest flux power spectrum. We sampled the cosmological parameters in conjunction with the Kähler parameters of a CY with $h^{1,1} = 7$. The maximum posterior probability was localised in moduli space in a region favouring the existence of an ultralight axion contributing $\mathcal{O}(1\%)$ to the DM density. We interpret this result as being due to a tension between these data in a pure CDM model that is alleviated by the non-zero ultralight axion DM density: the axion Jeans scale causes a suppression of the matter power spectrum relative to CDM [57]. Whether the tension in these data continues to persist in future with a solution based on fundamental physics (rather than e.g. systematics) is not central to our conclusions: we have demonstrated the *principle* of performing precise Bayesian analysis with a string axiverse model and cosmological data as a blueprint for future analyses. The next stage in such an analysis would be to compute the Bayesian evidence, in which case one could perform a model comparison between the CY and naive EFT models we have studied, and between these models and Λ CDM or other extensions of it. Such a computation could be carried out using nested sampling, and we leave this to a future work.

Taken as a whole, the methods and results we have presented open the door to future uses of observational data to perform Bayesian inference on CY moduli spaces, thus making tentative progress towards statistical model comparison both within string theory and relative to other theories.

ACKNOWLEDGMENTS

We are grateful to Naomi Gendler, Jim Halverson, Liam McAllister, Jakob Moritz, and Sebastian Vander Ploeg Fallon for helpful conversations. MJ and DJEM are supported by a Leverhulme Trust Research Project (Grant No. RPG-2022-145). DJEM is supported by an Ernest Rutherford Fellowship (Grant No. ST/T004037/1) and a consolidator grant (Grant No. ST/X000753/1) from the Science and Technologies Facilities Council. The research of ES and AS is supported by NSF grant PHY-2309456. KKR is supported by an Ernest Rutherford Fellowship from the UKRI Science and Technology Facilities Council (grant number ST/Z510191/1). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2139899. This work

was initiated at the Banff International Research Station during the workshop “Progress in the String Axiverse”. This work made use of the PYTHON-compatible packages EMCEE [71], GETDIST [114], NORMFLOWS [72], MATPLOTLIB [115], NUMPY [116], and SCIPY [117].

Appendix A: Axion effective field theory

Here we briefly recapitulate the relevant structure of the four-dimensional axion EFT in type IIB Calabi–Yau orientifolds, following the conventions of [16, 17]. Our aim is only to summarise definitions and scaling relations already implicit in Section II, where the Kähler moduli space and its Weil–Petersson measure were described.

1. Moduli and Kähler potential

As recalled in Section II, the complexified Kähler moduli are

$$T_i = \tau_i + i\phi_i = \frac{1}{2} \int_{D_i} J \wedge J + i \int_{D_i} C_4, \quad (\text{A.1})$$

where τ_i are the divisor volumes and ϕ_i are the axions descending from the Ramond–Ramond four-form C_4 . The Calabi–Yau volume and Kähler potential are

$$\mathcal{V} = \frac{1}{6} \kappa_{ijk} t^i t^j t^k \quad \mathcal{K} = -2 \log \mathcal{V}. \quad (\text{A.2})$$

The metric on moduli space, $K^{ij} = \partial_{T_i} \partial_{\bar{T}_j} \mathcal{K}$, then leads to the kinetic term for the complex moduli,

$$\mathcal{L}_{\text{kin}} = K^{ij} \partial_\mu T_i \partial^\mu \bar{T}_j = K^{ij} (\partial_\mu \tau_i \partial^\mu \tau_j + \partial_\mu \phi_i \partial^\mu \phi_j). \quad (\text{A.3})$$

The axion sub-block of this metric therefore governs the kinetic mixing and canonical normalisation of the ϕ_i .

The square root of the eigenvalues of K are the periodicities of the axions (modulo 2π), and therefore reflect the rough scaling of the axion decay constants f_i .

$$f_i^2 \sim \frac{1}{\lambda_i}, \quad \lambda_i \propto (\tau_i)^2 - \mathcal{V} \kappa_{iik} t^k. \quad (\text{A.4})$$

In the large-volume regime, the typical scaling (for deformable divisors) is [16]

$$f_i \sim \frac{M_{\text{Pl}}}{\sqrt{\mathcal{V}}}, \quad (\text{A.5})$$

up to order-one factors determined by the Kähler metric eigenvalues.

2. Nonperturbative potential and axion masses

Instanton and gaugino-condensation effects generate the nonperturbative axion potential

$$V(\phi) = \sum_\alpha \Lambda_\alpha^4 [1 - \cos(q_\alpha^i \phi_i)], \quad \Lambda_\alpha^4 \sim e^{-2\pi\tau_\alpha}. \quad (\text{A.6})$$

Here the ‘ α ’ summation runs over deformable divisors (wrapping ED3 branes), and q_α^i are the *reduced* instanton charges. That is, $q = Q \cdot S$ where Q agrees with the original integer-valued GLSM charge matrix, and S is the matrix that canonicalises the kinetic term (i.e., it brings the Kähler metric to the identity by a rotation and rescaling). The reduced charge matrix is lower triangular and therefore upon using a hierarchical approximation $\Lambda_\alpha \ll \Lambda_{\alpha+1}$, we have the following (for axions associated to deformable divisors)

$$m_i^2 \approx \frac{\Lambda_i^4}{f_i^2} \approx \frac{M_{\text{Pl}}^2}{\mathcal{V}} e^{-2\pi\tau_i}. \quad (\text{A.7})$$

3. The QCD axion

A stack of D7-branes wrapping a prime toric divisor supports a gauge group, and the corresponding closed-string axion becomes the gauge-theory axion associated with that sector.²⁰ For a generic gauge theory, the mass scales as $m_a \sim \Lambda_c^2/f_a$, where Λ_c is the confinement scale of the gauge group, and the gauge coupling satisfies $g^2 \sim 1/\text{vol}(D)$ at the compactification scale.

We model the QCD axion by replacing the corresponding instanton scale with the QCD scale inferred from chiral perturbation theory at zero temperature [118] giving:

$$m_{\text{QCD}} = 5.7 \mu\text{eV} \left(\frac{10^{12} \text{ GeV}}{f_a} \right). \quad (\text{A.8})$$

We demand that $\text{vol}(D)$ be close to 40, i.e., giving roughly the value of the QCD gauge coupling after running to high scales with the Standard Model particle content. The ‘‘axion quality’’ [20] quantifies how closely the induced θ angle of QCD approaches zero—equivalently, how nearly the gauge-theory axion aligns with a mass eigenstate—and we impose a benchmark quality criterion of 10^{-10} [119].

Appendix B: Normalising Flows

Normalising flows (NF) learn a complex distribution $p_x(\mathbf{x})$ by flowing a simple distribution (e.g., a Gaussian) $p_z(\mathbf{z})$ through a series of transformations. Let $f: \mathbb{R}^D \rightarrow \mathbb{R}^D$ be an invertible and differentiable function such that $f(\mathbf{z}) = \mathbf{x}$. Then one can express the distribution $p_x(\mathbf{x})$

in terms of the base distribution $p_z(\mathbf{z})$ and f using the change of variables formula:

$$p_x(\mathbf{x}) = p_z(\mathbf{z}) |\det J_f(\mathbf{z})|^{-1}, \quad (\text{B.1})$$

where J_f denotes the Jacobian of f , from which we get the following log likelihood:

$$\log p_x(\mathbf{x}) = \log p_z(\mathbf{z}) - \log |\det J_f(\mathbf{z})|. \quad (\text{B.2})$$

Given two invertible and differentiable transformations, f_1 and f_2 , their composition $f_2 \circ f_1$ is also invertible and differentiable. Therefore, we can build complex transformations by composing multiple simpler transformations $f = f_K \circ \dots \circ f_1$, where f_k transforms $\mathbf{z}_{k-1}$ into $\mathbf{z}_k$, where $\mathbf{z}_0 = \mathbf{z}$ and $\mathbf{z}_K = \mathbf{x}$.

Fitting a flow-based model $p_x(\mathbf{x}; \Theta)$ to a target distribution $p_x^*(\mathbf{x})$ can be done by minimising the Kullback-Leibler (KL) divergence between them, with respect to the model’s parameters Θ . Here we have two options; we can take the KL divergence of $p_x(\mathbf{x}; \Theta)$ with respect to $p_x^*(\mathbf{x})$ or vice versa. The former is called the reverse KL divergence, and the latter is called the forward KL divergence. By a change of variables we can write the reverse KL divergence loss as:

$$\begin{aligned} \mathcal{L}(\Theta) &= D_{KL}[p_x(\mathbf{x}; \Theta) || p_x^*(\mathbf{x})] \\ &= \mathbb{E}_{p_x(\mathbf{x}; \Theta)} [\log p_x(\mathbf{x}; \Theta) - \log p_x^*(\mathbf{x})] \\ &= \mathbb{E}_{p_z(\mathbf{z}; \psi)} [\log p_z(\mathbf{z}) - \log |\det J_f(\mathbf{z}; \Theta)| \\ &\quad - \log p_x^*(f(\mathbf{z}; \Theta))] \end{aligned} \quad (\text{B.3})$$

Similarly, we can write the forward KL divergence loss as

$$\begin{aligned} \mathcal{L}(\Theta) &= D_{KL}[p_x^*(\mathbf{x}) || p_x(\mathbf{x}; \Theta)] \\ &= -\mathbb{E}_{p_x^*(\mathbf{x})} [\log p_x(\mathbf{x}; \Theta)] + \text{const.} \\ &= -\mathbb{E}_{p_x^*(\mathbf{x})} [\log p_z(f^{-1}(\mathbf{x}; \Theta)) \\ &\quad + \log |\det J_{f^{-1}}(\mathbf{x}; \Theta)|] + \text{const.} \end{aligned} \quad (\text{B.4})$$

The forward KL divergence is often used when we can generate samples from the target distribution but we cannot evaluate the target density. On the other hand, the reverse KL divergence is used when we have the ability to evaluate the target density but we cannot easily sample from it.

1. Real NVP

In our experiments, we use a Real NVP (real-valued non volume preserving) flow model which consists of a series of affine coupling layers. The idea here is to partition the latent-variable $\mathbf{z}$ into two parts $\mathbf{z} = (\mathbf{z}_A, \mathbf{z}_B)$, where $\mathbf{z}_A$ has dimension d and $\mathbf{z}_B$ has dimension $D - d$. The output vector $f(\mathbf{z}) = \mathbf{x} = (\mathbf{x}_A, \mathbf{x}_B)$ is then given as

$$f(\mathbf{z}) = \begin{cases} \mathbf{x}_A = \mathbf{z}_A \\ \mathbf{x}_B = \mathbf{z}_B \odot \exp(S(\mathbf{z}_A)) + T(\mathbf{z}_A) \end{cases}, \quad (\text{B.5})$$

²⁰ For concreteness, one may envision a stack of n D7-branes wrapping a prime toric divisor, which in general supports a $U(n)$ gauge group. If the stack coincides with an O7-plane—corresponding to local tadpole cancellation—the group is instead projected to $SO(2n)$ or $Sp(n)$ depending on the orientifold charge of the O7-plane. In this work we remain agnostic about the precise realisation, as it would require specifying the O7-plane configuration and global tadpole cancellation conditions, which lie beyond the present scope.

where $\odot$ is the element-wise product and S and T are called the scale and translation functions respectively, which are implemented using neural networks.

The inverse function is given by

$$f^{-1}(\mathbf{x}) = \begin{cases} \mathbf{z}_A = \mathbf{x}_A \\ \mathbf{z}_B = (\mathbf{x}_B - T(\mathbf{x}_A)) \odot \exp S(\mathbf{x}_A) \end{cases}, \quad (\text{B.6})$$

where $\odot$ denotes element-wise division, and the Jacobian J_f is given by

$$J_f = \begin{pmatrix} \mathbf{I}_d & \mathbf{0}_{d \times (D-d)} \\ \frac{\partial \mathbf{x}_B}{\partial \mathbf{z}_A} & \text{diag}(\exp(S(\mathbf{z}_A))) \end{pmatrix} \quad (\text{B.7})$$

Partitioning into sets A and B can be achieved using a binary mask function B . In our implementation, we employ an alternating binary mask b where $b_i = 1$ for even indices and $b_i = 0$ for odd indices.

Using reverse KL divergence loss, the overall training procedure involves creating mini-batches of data points $\{\mathbf{z}_i\}$ by sampling the base distribution, flowing samples through the transformation layers to get $\{\mathbf{x}_i\}$ and computing Eq. (B.3). Similarly, using forward KL divergence loss, training involves sampling batches $\{\mathbf{x}_i\}$ from the target distribution and flowing samples through the inverse transformation layers to get $\{\mathbf{z}_i\}$ and computing Eq. (B.4). Gradients of the loss are evaluated using automatic differentiation and the network parameters are updated by stochastic gradient descent.

Appendix C: Geometry of Kähler Moduli of Calabi–Yau Threefolds

In this appendix we summarise the classical description of the Kähler moduli of a Calabi–Yau threefold. At tree level, the Kähler class determines the volumes of holomorphic cycles and provides the natural expansion parameters of the low-energy effective theory. Beyond tree level, these geometric quantities receive perturbative and non-perturbative corrections which play an essential role in realistic moduli stabilisation scenarios; see [37] for a recent overview.

1. Ricci-flat deformations and Kähler moduli

Ricci-flat metrics on Calabi–Yau threefolds occur in continuous families: for a fixed topology, the space of solutions to $R_{mn} = 0$ is parametrised by a *moduli space* of deformations. Let $g_{\alpha\bar{\beta}}$ be a Hermitian Ricci-flat metric on X , expressed in complex coordinates z^α and $\bar{z}^{\bar{\alpha}}$ so that

$$g_{\alpha\beta} = g_{\bar{\alpha}\bar{\beta}} = 0, \quad R_{\alpha\bar{\beta}}(g) = 0. \quad (\text{C.1})$$

Consider an infinitesimal deformation $g \rightarrow g + \delta g$ preserving Hermiticity,

$$(g + \delta g)_{\alpha\beta} = (g + \delta g)_{\bar{\alpha}\bar{\beta}} = 0, \quad (\text{C.2})$$

and keeping the metric Ricci-flat to first order,

$$R_{\alpha\bar{\beta}}(g + \delta g) = 0. \quad (\text{C.3})$$

In the de Donder gauge,

$$\nabla^m \delta g_{mn} - \frac{1}{2} \nabla_n \delta g^m_m = 0, \quad (\text{C.4})$$

this condition becomes the Lichnerowicz equation

$$\nabla^m \nabla_m \delta g_{pq} + 2 R_p^r q^s \delta g_{rs} = 0. \quad (\text{C.5})$$

The fluctuation decomposes as

$$\delta g = \delta g_{\alpha\bar{\beta}} dz^\alpha d\bar{z}^{\bar{\beta}} + \delta g_{\alpha\beta} dz^\alpha dz^\beta + \text{c.c.} \quad (\text{C.6})$$

Since on a Calabi–Yau the only non-vanishing Riemann components are $R_{\gamma\bar{\alpha}\delta\bar{\beta}}$, the $(\alpha\bar{\beta})$ and $(\alpha\beta)$ sectors of the Lichnerowicz operator decouple. Thus metric deformations fall into two independent classes.

In this appendix we focus exclusively on the mixed variations $\delta g_{\alpha\bar{\beta}}$, which correspond to deformations of the Kähler class. Such a deformation defines a harmonic $(1,1)$ -form,

$$\delta J := \delta g_{\alpha\bar{\beta}} dz^\alpha \wedge d\bar{z}^{\bar{\beta}}. \quad (\text{C.7})$$

By Hodge theory, harmonic $(1,1)$ -forms represent cohomology classes in $H^{1,1}(X)$, whose dimension is the Hodge number $h^{1,1}$. Hence these deformations correspond precisely to the *Kähler moduli*.

2. Weil–Petersson metric on Kähler moduli space

It is a general principle that if a theory features scalar fields φ^i , $1 \leq i \leq N$, valued in an N -dimensional target space, that the metric on that target space is the kinetic matrix K_{ij} featuring in the Lagrangian as $K_{ij} \partial^\mu \varphi^i \partial_\mu \varphi^j$. In the special case these scalar fields take values in a mathematical moduli space, including (but by no means limited to) either the Kähler or complex structure moduli space of Ricci-flat metrics on a Calabi–Yau threefold, this is historically known as the Weil–Petersson metric.

Thus, as our focus in this work is on Kähler moduli space, the natural metric is the Weil–Petersson metric on that moduli space. In particular, by studying the variations of the Ricci-flat metric induced by deformations of the Kähler class, we obtain the kinetic matrix of the Kähler moduli and thus the Weil–Petersson metric on Kähler moduli space. For the metric deformations $\delta g_{\alpha\bar{\beta}}$, this leads to

$$2G_{ij} \delta t^i \delta t^j = \frac{1}{2\mathcal{V}} \int_X g^{\alpha\bar{\beta}} g^{\gamma\bar{\delta}} \delta g_{\alpha\bar{\delta}} \delta g_{\bar{\beta}\gamma} \sqrt{g} d^6 y, \quad (\text{C.8})$$

We choose a basis $\omega_i \in H^{1,1}(X, \mathbb{R})$ so that

$$\delta g_{\alpha\bar{\beta}} = (\omega_i)_{\alpha\bar{\beta}} \delta t^i \quad (\text{C.9})$$

which implies that (C.8) can be written as

$$G_{ij} = \frac{1}{4\mathcal{V}} \int_X \omega_i \wedge \star_6 \omega_j, \quad (\text{C.10})$$

which agrees with [120], see also [30, 58]. One can then show that the action of the Hodge-star $\star_6$ on $\omega_j \in H^{1,1}(X, \mathbb{R})$ is given by [30]

$$\star_6 \omega_j = -J \wedge \omega_j + \frac{\int_X J \wedge J \wedge \omega_j}{4\mathcal{V}} J \wedge J. \quad (\text{C.11})$$

Let us introduce the quantities

$$\tau_j = \frac{1}{2} \int_X J \wedge J \wedge \omega_j = \frac{1}{2} \kappa_{jkl} t^k t^l, \quad (\text{C.12})$$

corresponding to the (classical) string frame divisor volumes. We can then write (C.10) as [30]

$$\begin{aligned} G_{ij} &= \frac{1}{4\mathcal{V}} \left[-\int_X J \wedge \omega_i \wedge \omega_j + \frac{\tau_i}{2\mathcal{V}} \int_X \omega_i \wedge J \wedge J \right] \\ &= \frac{1}{4\mathcal{V}} \left[-\kappa_{ijk} t^k + \frac{\tau_i \tau_j}{\mathcal{V}} \right]. \end{aligned} \quad (\text{C.13})$$

The Weil–Petersson metric G_{ij} on the Kähler moduli space thus obtained is entirely determined by the intersection data of X , a choice of Kähler parameters t^i and the classical divisor volumes τ_i . This classical form of G_{ij} will provide the reference point for incorporating perturbative and non-perturbative corrections in the next sections.

Having expressed the moduli space metric entirely in terms of the classical geometric data (t^i, κ_{ijk}) , we now identify the corresponding holomorphic coordinates appropriate for the four-dimensional $\mathcal{N} = 1$ theory. This requires promoting the real Kähler parameters to complex fields by incorporating the R-R axions. We define the complex coordinate

$$T_i := \frac{1}{2} \int_X J \wedge J \wedge \omega_i + i \int_X C_4 \wedge \omega_i \equiv \tau_i + i\theta_i \quad (\text{C.14})$$

where the classical string frame divisor volumes τ_i are given by (C.12). The T_i defined by (C.14) are good Kähler coordinates on the Kähler moduli space, at leading order in g_s and α' . The corresponding Kähler potential is

$$\mathcal{K} = -2 \log(\mathcal{V}), \quad \mathcal{V} := \frac{1}{6} \int_X J^3 = \frac{1}{6} \kappa_{ijk} t^i t^j t^k. \quad (\text{C.15})$$

Let us show that the tree-level Kähler potential (C.15) leads to the appropriate field space metric $\mathcal{K}_{i\bar{j}}$ which is directly related to the Weil–Petersson metric G_{ij} in (C.13) and corresponds a Kähler metric defined as $\mathcal{K}_{i\bar{j}} = \partial_{T_i} \partial_{\bar{T}_j} \mathcal{K}$. Noting that at the level of the tree level theory

$$\frac{\partial T_i}{\partial t^j} = \kappa_{ijk} t^k = \kappa_{ij}, \quad \partial_{t^i} \partial_{t^j} \mathcal{K} = 8G_{ij} \quad (\text{C.16})$$

we obtain the expression

$$\partial_{T_i} \partial_{\bar{T}_j} \mathcal{K} = \frac{1}{8\mathcal{V}^2} \left[-2\mathcal{V} \kappa^{ij} + t^i t^j \right], \quad (\text{C.17})$$

where κ^{ij} is the inverse of κ_{ij} .

3. Detailed Geometric Prerequisites

In this appendix, we take some time to more thoroughly introduce the geometric ideas summarised in [Section II.C](#). In particular, we will discuss fibrations, shrinkable divisors, and geometric aspects of Kähler metric eigenvalues and axion decay constants.

To organise the former two phenomena, it is useful to recall the classification of facets (codimension-one faces) of Kähler cones of CY threefolds, based on the four types of contractions they give rise to [121–125]

1. The CY threefold shrinks to a variety of smaller dimension (in particular, the CY is fibered, and both the fiber and CY shrink to zero volume),
2. A divisor shrinks to a curve,
3. A divisor shrinks to a point,
4. A curve shrinks to a point.

Facets in the first three categories correspond to those which bound the geometric moduli space of the CY, while the final category corresponds to a flop transition, in which the adjoining Kähler cone belongs to a birational Calabi–Yau threefold obtained by flopping a curve. Facets where the overall volume or individual divisor volumes vanish have been studied closely [76], but the same behaviour can occur along higher-codimension faces of the Kähler cone, so we will not restrict to the codimension-one case.

a. Fibrations

Fibered CY threefolds play a crucial role in string dualities [126, 127], furnish canonical examples of infinite distance limits in moduli spaces studied by the Swampland program [128, 129], and have historically played an important role in string phenomenology: e.g., $K3$ fibrations in the Large Volume Scenario (LVS) [69, 70, 130]. We mentioned in [Section II.C](#) that because WP measure is normalised by fixing minimum and maximum values of the overall Calabi–Yau volume, compactness of the support of the WP measure is related to faces where $\mathcal{V}$ vanishes. Because $\mathcal{V}$ is a monotonic under overall rescaling of the Kähler parameters, along any ray where $\mathcal{V}$ doesn’t vanish identically, the region satisfying $\mathcal{V}_{\min} \leq \mathcal{V} \leq \mathcal{V}_{\max}$ is bounded. As a homogeneous polynomial in the Kähler parameters, for any given ray, $\mathcal{V}$ either vanishes only at the origin or vanishes identically, with the latter only possible on faces of the Kähler cone. The region $\mathcal{V}_{\min} \leq \mathcal{V} \leq \mathcal{V}_{\max}$ supporting the WP measure near such faces extend to infinity, as the bounding level sets of $\mathcal{V}$ becomes asymptotically parallel to the faces. Compactness of the WP support, then, is determined by whether there are Kähler cone faces where $\mathcal{V}$ vanishes. We now explain why faces where $\mathcal{V}$ vanishes correspond to with fibrations of X .

If $\mathcal{V} = 0$ for some choice of Kähler form J , then J lies on a boundary of the Kähler cone, corresponding to a Poincaré-dual nef divisor D whose triple self-intersection vanishes, $D^3 = 0$.²¹ A nef divisor D is “big” if and only if $D^3 > 0$ [131]. It is a standard fact in algebraic geometry that ample divisors A define embeddings of X into projective space by mapping $x \in X$ to $[s_0(x) : \cdots : s_n(x)]$ for s_i the global sections of mA for m a sufficiently large integer (indeed, this is often taken as the definition of an ample divisor) [131]. Intuitively, relevance of the “big” condition is that the multiples of a big divisor D carry enough global sections to map X to a three-dimensional image in projective space — without ampleness, this may not be an honest embedding of X , but the big property ensure that the powers of D probe all directions of the threefold. When a nef D is not big (i.e., $D^3 = 0$), its sections fail to vary along some directions, collapsing those dimensions to form a lower-dimensional base B . Consequently, $D^3 = 0$ signals a fibration rather than an embedding (sometimes known as a semi-ample or Iitaka fibration [131]). The process of approaching the Kähler form J for which $\mathcal{V}$ vanishes implements the limit where the generic fiber of this fibration shrinks (while the base volume is free to take on any value).

If $\mathcal{V}$ vanishes on a k -dimensional face of the Kähler cone, then the integral Kähler forms J on this face are Poincaré dual to divisors D which are pullbacks f^*D' of nef divisors D' on B , which satisfies $h^{1,1}(B) = k$.²² That is, the D' are curves if B is a surface, or isolated points if B is a curve. Because these divisors are pullbacks from the base, each $D = f^*D'$ contains the entire fiber above the corresponding divisor $D' \subset B$. Consequently, the intersection of ℓ such divisors in X corresponds to the union of fibers lying above the intersection of the ℓ divisors D'_i on the base. This explains why $D^3 = 0$: the base is at most two-dimensional, and an intersection of $\ell = 3$ nef divisors on a space of dimension at most two, is empty.

We reviewed the classification of fibrations of Calabi–Yau threefolds in Section II.C: the fiber can have complex dimension one, and be an elliptic fibration, or complex dimension two, and be a $K3$ fibration, where here we continue to follow the convention set in Section II.C where we ignore distinctions between genus-one and elliptic fibrations and also between $K3$ and $T4$ fibrations for notational simplicity. We can diagnose the fiber dimension (and hence the type of fibration) directly from intersections of pullbacks from the base. Recall our logic from the previous paragraph for why $D^3 = 0$: if B is a surface ($\dim_{\mathbb{C}} B = 2$, the elliptic case), the intersection of two base divisors is a finite set of points, so intersecting two pullback divisors in X yields a union of fibers (a curve class). If B is a curve ($\dim_{\mathbb{C}} B = 1$, the $K3$ case),

two distinct base divisors do not meet, so the square of a pullback class vanishes. We conclude that a fibration is $K3$ if $D^2 = 0$ and elliptic otherwise. Indeed, we have recovered the famous Kollár condition for the existence of an elliptic fibration on a CY threefold [133]: namely, the existence of a so-called Kollár divisor, or a nef divisor satisfying $D^3 = 0$ but $D^2 \neq 0$. This discussion has an additional corollary: for an elliptic fibration, the fiber class is proportional to $f^*D' \cdot f^*D''$ for any two base divisors D', D'' falling on the Kähler cone face. For a $K3$ fibration, the fiber class is proportional to f^*D' itself. Two fibrations are equivalent if and only if the classes of their fibers coincide. Computing the class of the base is much more difficult: this amounts to constructing a section of the fibration, which is a challenging problem. It is worth noting that the base of an elliptic fibration has unconstrained $h^{1,1}$, while a $K3$ fibration must have base $\mathbb{P}^1$. So, elliptic fibrations are free to correspond to Kähler cone faces of any dimension while $K3$ fibrations can only correspond to one-faces. In this work, $h^{1,1}$ is sufficiently small that we can compute extremal faces of $K_{\cup}$, our approximation to the Kähler cone, and test for fibrations in the above fashion, but we comment in passing that there additionally exist useful methods for computing the fibrations of toric hypersurface Calabi–Yau varieties inherited from fibrations of the ambient toric variety [73].

We conclude our discussion of fibrations with a more physical comment on Kaluza–Klein scales. In practice, we will bound our overall volume (in order to bound the WP measure) by the physical requirement that the overall KK scale $\sim g_s M_{\text{Pl}}/\mathcal{V}^{2/3}$ is sufficiently large. In the limit of small fiber, however, the KK scale of the base B can become dangerously small even when the KK scale of the Calabi–Yau is phenomenologically viable. In particular, imposing our stretching condition that the fiber must have volume ≥ 1 , the base KK scale for unit fiber volume is $g_s M_{\text{Pl}}/\mathcal{V}^{1/2+1/\dim B}$, which is smaller by a factor of $\mathcal{V}^{-1/3}$ for elliptic fibrations and $\mathcal{V}^{-1/12}$ for $K3$ fibrations. Because it is generally difficult to compute the class of the base of a fibration, we defer the problem of incorporating relevant divisor and curve KK scales to future work, but acknowledge that constraints on these scales should in general be imposed.

b. Shrinking Divisors

Now we turn our attention to divisors that can shrink to zero volume at faces of the Kähler cone. Divisor volumes are quadratic forms in the Kähler parameters and have signature $(+, -, \dots, -)$ (this follows from the Hodge index theorem, at least for smooth divisors). Thus, a (possibly complex) basis of Kähler parameters can always be chosen such that their volume is $\tau = (t^1)^2 - \sum_{i=2}^k (t^i)^2$ for k an invariant of the divisor that counts the dimension of the subspace of $H^{1,1}(D)$ inherited from the ambient Calabi–Yau (equal to the number of independent ambient two-forms restricting non-trivially to D).

²¹ A nef divisor D satisfies $D \cdot C \geq 0$ for all effective curves C .

²² For the case of elliptic fibrations, these divisors are known in the F-theory literature as “vertical divisors” [132].

Let us first consider divisors that shrink at facets of the Kähler cone: i.e. codimension-one faces. If the divisor shrinks to a point along a facet, then its volume vanishes quadratically at that facet, so $k = 1$ and $\tau = t_1^2$. Likewise, if it shrinks to a curve along a facet, then its volume vanishes linearly, so $\tau = fg$ for f, g two linear functions of the Kähler parameters which we can formally write as $\tau = ([f + ig]/2e^{i\pi/4})^2 - ([f - ig]/2e^{i\pi/4})^2 = (t^1)^2 - (t^2)^2$, so $k = 2$. Divisors which can shrink at finite overall volume have been referred to as blowup divisors — indeed, they are the exceptional divisors achieved by blowing up a point or curve — while divisors whose volumes are a perfect square (i.e., $k = 1$) have been called diagonal divisors [70]. We note that satisfying $k = 1$ ($k = 2$) is necessary but not sufficient for a divisor to shrink to a point (curve) along some facet: sometimes the locus where the shrinking would take place is strictly outside of the Kähler cone.

Because all boundaries of the CY Kähler moduli space must either correspond to fibrations or shrinking divisors, divisors that shrink at facets are actually relatively common. However, for the purposes of axion physics, it isn't strictly necessary that the shrinking happens at codimension one. The volume polynomials for divisors which shrink at higher codimension faces do not need to factor: it suffices for them to factor only in the limit where one approaches the relevant face while remaining in the Kähler cone. We do not endeavour to provide a comprehensive classification of divisors shrinking at higher codimension faces: however, we present a numerical approach for computing which divisors shrink along each face, and we can also make the following remark. Recall from Section IV that a divisor D is geometrically rigid (i.e., has no normal bundle deformations) if $h^{2,0}(D)$ (which agrees with $h^0(X, \mathcal{O}_X(D_i)) - 1$ for Calabi–Yau threefolds X [68]) equals zero, and movable otherwise. If D is movable, then the divisors in its class sweep out a dense subset of the Calabi–Yau threefold. Such divisors cannot shrink without the entire CY shrinking: that is, they cannot shrink independently from the overall volume. Geometrically rigid divisors, on the other hand, are only represented by a single holomorphic divisor. The geometric rigidity property is a necessary but insufficient condition for a divisor to be able to shrink independent of the CY.

An important class of shrinking divisors actually arise from fibrations. If F is a face of the Kähler cone corresponding to a fibration, then any divisor D which lies on the linear subspace generated by that face will have its volume vanish at least linearly along that face. Intuitively, these divisors are the pullback divisors from the base of the fibration, so they are themselves fibered, so their volume also shrinks to zero as the fiber shrinks. These correspond to the “fibration divisor” category defined in Section II.C.

We now take some time to describe how to identify and taxonomise divisors which shrink along a given face F of the Kähler cone generated by Kähler forms $J_1, \dots, J_n$. A divisor D with volume τ vanishes along

F if $\kappa_{ijk} J^i J^j D^k = 0$ for J any linear combination of $J_1, \dots, J_n$. One can form a matrix K whose columns are the J_i : then we require that $\kappa_{ijk} K^{ia} K^{jb} D^k = 0$ identically, for all a, b . If we flatten a, b indices of the tensor $\kappa_{ijk} K^{ia} K^{jb}$ into a single index c , leaving a matrix L_i^c , then we merely have the condition that D lies in the kernel of L . We can refine this approach to identify the divisors which contract to points, or have a volume that approaches zero quadratically near F . In particular, we now require that D lie in the kernel of M_k^c given as $\kappa_{ijk} K^{ka}$ with the a, j indices flattened into c .

With this in hand, we elaborate on how to assign one of the five categories named in Section II.C to a divisor D . If D shrinks to a point along some face F that doesn't correspond to a fibration (i.e., $\mathcal{V}$ doesn't vanish identically on F) then it is a point blowup divisor. Otherwise, if it shrinks at all along some non-fibration face F , it is a curve blowup divisor. Otherwise, if D is proportional to the class of a $K3$ fiber, it will shrink in the limit of shrinking fiber and we say it is a $K3$ fiber divisor. Otherwise, if D shrinks along a face F that does correspond to a fibration, we say D is a fibration divisor. Otherwise, D does not shrink and is a non-shrinkable divisor.

The challenge here is knowing the faces of the Kähler cone: we take an inner approximation to the Kähler cone given by $K_\cup$, and computing all of the faces of this approximation is computationally hard for $h^{1,1} \gtrsim 10$.

c. Kähler metric and decay constants

Consider the inverse Kähler metric

$$[K^{-1}]_{ij} = 2(\tau_i \tau_j - \mathcal{V} \kappa_{ijk} t^k). \quad (\text{C.18})$$

This furnishes a metric and thus a norm $\|\cdot\|_K$ on divisors D .²³ In particular,

$$\|D\|_K^2 \propto \tau^2 - \mathcal{V}C, \quad (\text{C.19})$$

for τ the volume of D , and C the formal volume of the curve class of the self-intersection $D \cap D$. But, this class may have no holomorphic representatives: i.e., it may not be an effective curve class. In the limit of hierarchical instanton scales, one can consider the axion decay constant f associated to a divisor D belonging to the basis of leading instantons, and compute it perturbatively using [17]. This decay constant satisfies

$$f \geq \frac{1}{2\pi\|D\|_K}. \quad (\text{C.20})$$

²³ This is a mildly subtle point: the Kähler metric, as defined, acts on divisor volumes τ_i (and more generally the Kähler moduli T_i) as well as curve classes C_i , as these both belong to $H_2(X, \mathbb{R})$, while the inverse Kähler metric acts on curve volumes (parametrised by Kähler parameters t^i) and divisor classes D^i , members of $H^2(X, \mathbb{R}) \cong H_4(X, \mathbb{R})$.

Briefly, this is because $1/2\pi f$ is a component of the divisor D expressed in a basis where $K^{-1} = 1$ (i.e., the kinetic matrix is canonical), so that $1/(2\pi f)^2 + \text{positive} = \|D\|_K^2$. In this way, $\|D\|_K$ provides an important probe of axion decay constants.

There are two qualitatively different cases to consider. First, the case that D is geometrically rigid. In this scenario there is no effective curve arising from self-intersection of representatives of D , because D has a unique representative, so the intersection of two representatives is a divisor, not a curve. Because D^2 isn't an effective curve class, it is nearly always negative (as a numerical coincidence, it can sometimes be positive for certain choices of Kähler forms, but it is empirically almost always negative). In this case the Kähler metric is manifestly positive — the sum of two positive numbers — and either term may dominate. Second, consider the case that D is movable: then the self-intersection curve associated to C is effective, and in particular $C \geq 0$. Combining this with the constraint that $\|D\|_K$ be positive, we find $0 < \mathcal{V}C < (\tau_i)^2$, which is quite restrictive. Empirically, we find that $(\tau_i)^2/\mathcal{V}C$ never comes too close to unity, so in the limit of large volume $(\tau_i)^2 - \mathcal{V}C$ tends to be large.

Let us consider the consequences of this discussion for

axion physics. Firstly if some axions correspond to divisors D with normal bundle deformations, then the associated decay constants will generally tend to be larger and have less spread than those of geometrically rigid divisors. This is because the $\|D\|_K$ are smaller — a difference, rather than a sum — and bounded from above and below. In the large-volume limit, the overall Calabi–Yau volume still sets the common scale of the decay constants, though the mechanism differs between the two classes of divisors. Geometrically rigid divisors can shrink without significantly affecting the total volume — examples include divisors that contract to points or curves at finite $\mathcal{V}$. In this case, the two contributions in Eq. (C.19) to $\|D\|_K$ are somewhat decoupled, and the $\mathcal{V}$ term dominates in the large volume regime, as was noted in [17] and exploited in [23]. By contrast, deformable divisors sweep out a dense subset of the Calabi–Yau and cannot shrink independently of the overall volume $\mathcal{V}$, implying a strong correlation between their volume and $\mathcal{V}$; in this sense, $\mathcal{V}$ acts as an overall factor in $\|D\|_K$.

We conclude by commenting briefly on the case where D is the fiber of a $K3$ fibration. In this case, from the adjunction formula it follows that $D \cap D$ is the canonical class on D , which must vanish for $K3$, so $C = 0$. From this it follows that $\|D\|$ — and thus, a lower bound on f — depends only on the volume of the $K3$ fiber.

-
- [1] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, *Vacuum configurations for superstrings*, *Nucl. Phys. B* **258** (1985) 46–74.
 - [2] M. Demirtas, A. Rios-Tascon and L. McAllister, *CYTools: A Software Package for Analyzing Calabi–Yau Manifolds*, [2211.03823](#).
 - [3] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.
 - [4] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.
 - [5] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.
 - [6] E. Witten, *Some Properties of $O(32)$ Superstrings*, *Phys. Lett. B* **149** (1984) 351–356.
 - [7] P. Svrcek and E. Witten, *Axions In String Theory*, *JHEP* **06** (2006) 051, [[hep-th/0605206](#)].
 - [8] J. P. Conlon, *The QCD axion and moduli stabilisation*, *JHEP* **05** (2006) 078, [[hep-th/0602233](#)].
 - [9] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper and J. March-Russell, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530, [[0905.4720](#)].
 - [10] B. S. Acharya, K. Bobkov and P. Kumar, *An M Theory Solution to the Strong CP Problem and Constraints on the Axiverse*, *JHEP* **11** (2010) 105, [[1004.5138](#)].
 - [11] M. Cicoli, M. Goodsell and A. Ringwald, *The type IIB string axiverse and its low-energy phenomenology*, *JHEP* **10** (2012) 146, [[1206.0819](#)].
 - [12] F. Chadha-Day, J. Ellis and D. J. E. Marsh, *Axion dark matter: What is it and why now?*, *Sci. Adv.* **8** (2022) abj3618, [[2105.01406](#)].
 - [13] C. B. Adams et al., *Axion Dark Matter*, in *Snowmass 2021*, 3, 2022, [2203.14923](#).
 - [14] C. A. J. O’Hare, *Cosmology of axion dark matter*, *PoS COSMICWISPers* (2024) 040, [[2403.17697](#)].
 - [15] M. Demirtas, C. Long, L. McAllister and M. Stillman, *The Kreuzer-Skarke Axiverse*, *JHEP* **04** (2020) 138, [[1808.01282](#)].
 - [16] E. Sheridan, F. Carta, N. Gendler, M. Jain, D. J. E. Marsh, L. McAllister et al., *Fuzzy axions and associated relics*, *JHEP* **09** (2025) 016, [[2412.12012](#)].
 - [17] N. Gendler, D. J. E. Marsh, L. McAllister and J. Moritz, *Glimmers from the axiverse*, *JCAP* **09** (2024) 071, [[2309.13145](#)].
 - [18] N. Gendler and D. J. E. Marsh, *Possible Implications of QCD Axion Dark Matter Constraints from Helioscopes and Haloscopes for the String Theory Landscape*, *Phys. Rev. Lett.* **134** (2025) 081602, [[2407.07143](#)].
 - [19] V. M. Mehta, M. Demirtas, C. Long, D. J. E. Marsh, L. McAllister and M. J. Stott, *Superradiance in string theory*, *JCAP* **07** (2021) 033, [[2103.06812](#)].
 - [20] M. Demirtas, N. Gendler, C. Long, L. McAllister and J. Moritz, *PQ axiverse*, *JHEP* **06** (2023) 092, [[2112.04503](#)].
 - [21] J. N. Benabou, K. Fraser, M. Reig and B. R. Safdi, *String theory and grand unification suggest a submicroelectronvolt QCD axion*, *Phys. Rev. D* **112** (2025) 066003, [[2505.15884](#)].

- [22] Z. Yin, H. Cheng, E. Di Valentino, N. Gendler, D. J. E. Marsh and L. Visinelli, *Constraining the axiverse with reionization*, [2507.03535](#).
- [23] J. Cheng and N. Gendler, *Universality in the Axiverse*, [2507.12516](#).
- [24] N. MacFadden, A. Schachner and E. Sheridan, *The DNA of Calabi-Yau Hypersurfaces*, [2405.08871](#).
- [25] S. V. P. Fallon, J. Halverson, L. McAllister and Y. Zhu, *F-theory Axiverse*, [2511.20458](#).
- [26] R. Petrossian-Byrne and G. Villadoro, *Open string axiverse*, *JHEP* **07** (2025) 049, [[2503.16387](#)].
- [27] M. Reig and T. Weigand, *Testing the Heterotic String with the Axion-Photon Coupling*, [2509.08042](#).
- [28] V. Loladze, A. Platschorre and M. Reig, *Higher axion strings*, *JHEP* **08** (2025) 182, [[2503.18707](#)].
- [29] J. M. Leedom, M. Putti and A. Westphal, *Towards a Heterotic Axiverse*, [2509.03578](#).
- [30] P. Candelas and X. de la Ossa, *Moduli Space of Calabi-Yau Manifolds*, *Nucl. Phys. B* **355** (1991) 455–481.
- [31] L. McAllister and F. Quevedo, *Moduli Stabilization in String Theory*, [2310.20559](#).
- [32] J. Louis, M. Rummel, R. Valandro and A. Westphal, *Building an explicit de Sitter*, *JHEP* **10** (2012) 163, [[1208.3208](#)].
- [33] M. Cicoli, D. Klevers, S. Krippendorff, C. Mayrhofer, F. Quevedo and R. Valandro, *Explicit de Sitter Flux Vacua for Global String Models with Chiral Matter*, *JHEP* **05** (2014) 001, [[1312.0014](#)].
- [34] C. Crinò, F. Quevedo and R. Valandro, *On de Sitter String Vacua from Anti-D3-Branes in the Large Volume Scenario*, *JHEP* **03** (2021) 258, [[2010.15903](#)].
- [35] S. AbdusSalam, S. Abel, M. Cicoli, F. Quevedo and P. Shukla, *A systematic approach to Kähler moduli stabilisation*, *JHEP* **08** (2020) 047, [[2005.11329](#)].
- [36] M. Demirtas, M. Kim, L. McAllister, J. Moritz and A. Rios-Tascon, *Small cosmological constants in string theory*, *JHEP* **12** (2021) 136, [[2107.09064](#)].
- [37] L. McAllister, J. Moritz, R. Nally and A. Schachner, *Candidate de Sitter Vacua*, [2406.13751](#).
- [38] S. AbdusSalam, C. Hughes, F. Quevedo and A. Schachner, *Coexisting Flux String Vacua from Numerical Kähler Moduli Stabilisation*, [2507.00615](#).
- [39] S. B. Giddings, S. Kachru and J. Polchinski, *Hierarchies from fluxes in string compactifications*, *Phys. Rev. D* **66** (2002) 106006, [[hep-th/0105097](#)].
- [40] E. Witten, *Nonperturbative superpotentials in string theory*, *Nucl. Phys. B* **474** (1996) 343–360, [[hep-th/9604030](#)].
- [41] R. Bousso and J. Polchinski, *Quantization of four form fluxes and dynamical neutralization of the cosmological constant*, *JHEP* **06** (2000) 006, [[hep-th/0004134](#)].
- [42] F. Denef and M. R. Douglas, *Distributions of flux vacua*, *JHEP* **05** (2004) 072, [[hep-th/0404116](#)].
- [43] M. Berg, M. Haack and B. Kors, *Loop corrections to volume moduli and inflation in string theory*, *Phys. Rev. D* **71** (2005) 026005, [[hep-th/0404087](#)].
- [44] M. Berg, M. Haack and B. Kors, *String loop corrections to Kahler potentials in orientifolds*, *JHEP* **11** (2005) 030, [[hep-th/0508043](#)].
- [45] A. Dubey, S. Krippendorff and A. Schachner, *JAXVacua — a framework for sampling string vacua*, *JHEP* **12** (2023) 146, [[2306.06160](#)].
- [46] J. Ebelt, S. Krippendorff and A. Schachner, *W0.sample* = *np.random.normal(0,1)?*, *Phys. Lett. B* **855** (2024) 138786, [[2307.15749](#)].
- [47] S. Krippendorff and A. Schachner, *New non-supersymmetric flux vacua in string theory*, *JHEP* **12** (2023) 145, [[2308.15525](#)].
- [48] A. Chauhan, M. Cicoli, S. Krippendorff, A. Maharana, P. Piantadosi and A. Schachner, *Deep observations of the Type IIB flux landscape*, *JHEP* **07** (2025) 271, [[2501.03984](#)].
- [49] PARTICLE DATA GROUP collaboration, S. Navas et al., *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.
- [50] T. W. Grimm, *Non-Perturbative Corrections and Modularity in N=1 Type IIB Compactifications*, *JHEP* **10** (2007) 004, [[0705.3253](#)].
- [51] T. W. Grimm, M. Kerstan, E. Palti and T. Weigand, *On Fluxed Instantons and Moduli Stabilisation in IIB Orientifolds and F-theory*, *Phys. Rev. D* **84** (2011) 066001, [[1105.3193](#)].
- [52] M. Cicoli, A. Schachner and P. Shukla, *Systematics of type IIB moduli stabilisation with odd axions*, *JHEP* **04** (2022) 003, [[2109.14624](#)].
- [53] M. Jain, F. Carta and D. J. E. Marsh, *Systematics of Axion Physics on Calabi-Yau Orientifolds* (forthcoming), .
- [54] N. MacFadden, *Efficient Algorithm for Generating Homotopy Inequivalent Calabi-Yaus*, [2309.10855](#).
- [55] N. Gendler, N. MacFadden, L. McAllister, J. Moritz, R. Nally, A. Schachner et al., *Counting Calabi-Yau Threefolds*, [2310.06820](#).
- [56] A. Chandra, A. Constantin, C. S. Fraser-Taliente, T. R. Harvey and A. Lukas, *Enumerating Calabi-Yau Manifolds: Placing Bounds on the Number of Diffeomorphism Classes in the Kreuzer-Skarke List*, *Fortsch. Phys.* **72** (2024) 2300264, [[2310.05909](#)].
- [57] K. K. Rogers and V. Poulin, *5 σ tension between Planck cosmic microwave background and eBOSS Lyman-alpha forest and constraints on physics beyond Λ CDM*, *Phys. Rev. Res.* **7** (2025) L012018, [[2311.16377](#)].
- [58] T. W. Grimm and J. Louis, *The Effective action of N = 1 Calabi-Yau orientifolds*, *Nucl. Phys. B* **699** (2004) 387–426, [[hep-th/0403067](#)].
- [59] P. Candelas, T. Hubsch and R. Schimmrigk, *Relation Between the Weil-petersson and Zamolodchikov Metrics*, *Nucl. Phys. B* **329** (1990) 583–590.
- [60] J. Stout, *Infinite Distances and Factorization*, [2208.08444](#).
- [61] D. A. Cox and S. Katz, *Mirror Symmetry and Algebraic Geometry*, vol. 68 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1999.
- [62] K. OGUISO, *On algebraic fiber space structures on a calabi-yau 3-fold*, *International Journal of Mathematics* **04** (1993) 439–465, [[https://doi.org/10.1142/S0129167X93000248](#)].
- [63] M. Cicoli, S. Krippendorff, C. Mayrhofer, F. Quevedo and R. Valandro, *D-Branes at del Pezzo Singularities: Global Embedding and Moduli Stabilisation*, *JHEP* **09** (2012) 019, [[1206.5237](#)].
- [64] M. Cicoli, S. Krippendorff, C. Mayrhofer, F. Quevedo and R. Valandro, *D3/D7 Branes at Singularities: Constraints from Global Embedding and Moduli Stabilisation*, *JHEP* **07** (2013) 150, [[1304.0022](#)].

- [65] M. Cicoli, S. Krippendorff, C. Mayrhofer, F. Quevedo and R. Valandro, *The Web of D-branes at Singularities in Compact Calabi-Yau Manifolds*, *JHEP* **05** (2013) 114, [[1304.2771](#)].
- [66] M. Cicoli, I. Garcia-Etxebarria, C. Mayrhofer, F. Quevedo, P. Shukla and R. Valandro, *Global Orientifolded Quivers with Inflation*, *JHEP* **11** (2017) 134, [[1706.06128](#)].
- [67] M. Cicoli, I. G. Etxebarria, F. Quevedo, A. Schachner, P. Shukla and R. Valandro, *The Standard Model quiver in de Sitter string compactifications*, *JHEP* **08** (2021) 109, [[2106.11964](#)].
- [68] A. P. Braun, C. Long, L. McAllister, M. Stillman and B. Sung, *The Hodge Numbers of Divisors of Calabi-Yau Threefold Hypersurfaces*, **1712.04946**.
- [69] M. Cicoli, J. P. Conlon and F. Quevedo, *General Analysis of LARGE Volume Scenarios with String Loop Moduli Stabilisation*, *JHEP* **10** (2008) 105, [[0805.1029](#)].
- [70] M. Cicoli, M. Kreuzer and C. Mayrhofer, *Toric K3-Fibred Calabi-Yau Manifolds with del Pezzo Divisors for String Compactifications*, *JHEP* **02** (2012) 002, [[1107.0383](#)].
- [71] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, *emcee: The MCMC Hammer*, *Publ. Astron. Soc. Pac.* **125** (2013) 306–312, [[1202.3665](#)].
- [72] V. Stimper, D. Liu, A. Campbell, V. Berenz, L. Ryll, B. Schölkopf et al., *normflows: A pytorch package for normalizing flows*, *Journal of Open Source Software* **8** (2023) 5361.
- [73] M. Kreuzer and H. Skarke, *Reflexive polyhedra, weights and toric Calabi-Yau fibrations*, *Rev. Math. Phys.* **14** (2002) 343–374, [[math/0001106](#)].
- [74] D. Rezende and S. Mohamed, *Variational inference with normalizing flows*, in *Proceedings of the 32nd International Conference on Machine Learning*, pp. 1530–1538, PMLR, 2015, <https://proceedings.mlr.press/v37/rezende15.html>.
- [75] L. Dinh, J. Sohl-Dickstein and S. Bengio, *Density estimation using real NVP*, *arXiv preprint arXiv:1605.08803* (2016), [[1605.08803](#)].
- [76] N. Gendler, B. Heidenreich, L. McAllister, J. Moritz and T. Rudelius, *Moduli space reconstruction and Weak Gravity*, *JHEP* **12** (2023) 134, [[2212.10573](#)].
- [77] J. Halverson, C. Long, B. Nelson and G. Salinas, *Towards string theory expectations for photon couplings to axionlike particles*, *Phys. Rev. D* **100** (2019) 106010, [[1909.05257](#)].
- [78] ADMX collaboration, N. Du et al., *A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment*, *Phys. Rev. Lett.* **120** (2018) 151301, [[1804.05750](#)].
- [79] DMRADIO collaboration, L. Brouwer et al., *Projected sensitivity of DMRadio-m3: A search for the QCD axion below 1 μ eV*, *Phys. Rev. D* **106** (2022) 103008, [[2204.13781](#)].
- [80] J. A. Frieman, C. T. Hill, A. Stebbins and I. Waga, *Cosmology with ultralight pseudo Nambu-Goldstone bosons*, *Phys. Rev. Lett.* **75** (1995) 2077–2080, [[astro-ph/9505060](#)].
- [81] W. Hu, R. Barkana and A. Gruzinov, *Cold and fuzzy dark matter*, *Phys. Rev. Lett.* **85** (2000) 1158–1161, [[astro-ph/0003365](#)].
- [82] L. Amendola and R. Barbieri, *Dark matter from an ultra-light pseudo-Goldstone-boson*, *Phys. Lett. B* **642** (2006) 192–196, [[hep-ph/0509257](#)].
- [83] D. J. E. Marsh and P. G. Ferreira, *Ultra-Light Scalar Fields and the Growth of Structure in the Universe*, *Phys. Rev. D* **82** (2010) 103528, [[1009.3501](#)].
- [84] A. Laguë, J. R. Bond, R. Hložek, K. K. Rogers, D. J. E. Marsh and D. Grin, *Constraining ultralight axions with galaxy surveys*, *JCAP* **01** (2022) 049, [[2104.07802](#)].
- [85] D. J. E. Marsh, *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1–79, [[1510.07633](#)].
- [86] A. Eberhardt and E. G. M. Ferreira, *Ultralight fuzzy dark matter review*, **2507.00705**.
- [87] R. Hložek, D. Grin, D. J. E. Marsh and P. G. Ferreira, *A search for ultralight axions using precision cosmological data*, *Phys. Rev. D* **91** (2015) 103512, [[1410.2896](#)].
- [88] T. Kobayashi, R. Murgia, A. De Simone, V. Iršič and M. Viel, *Lyman- α constraints on ultralight scalar dark matter: Implications for the early and late universe*, *Phys. Rev. D* **96** (2017) 123514, [[1708.00015](#)].
- [89] K. K. Rogers and H. V. Peiris, *Strong Bound on Canonical Ultralight Axion Dark Matter from the Lyman-Alpha Forest*, *Phys. Rev. Lett.* **126** (2021) 071302, [[2007.12705](#)].
- [90] K. K. Rogers and H. V. Peiris, *General framework for cosmological dark matter bounds using N-body simulations*, *Phys. Rev. D* **103** (2021) 043526, [[2007.13751](#)].
- [91] K. K. Rogers, R. Hložek, A. Laguë, M. M. Ivanov, O. H. E. Philcox, G. Cabass et al., *Ultra-light axions and the S_8 tension: joint constraints from the cosmic microwave background and galaxy clustering*, *JCAP* **06** (2023) 023, [[2301.08361](#)].
- [92] H. Winch, K. K. Rogers, R. Hložek and D. J. E. Marsh, *High-redshift, Small-scale Tests of Ultralight Axion Dark Matter Using Hubble and Webb Galaxy UV Luminosities*, *Astrophys. J.* **976** (2024) 40, [[2404.11071](#)].
- [93] M. A. Fernandez, S. Bird and M.-F. Ho, *Cosmological constraints from the eBOSS Lyman- α forest using the PRIYA simulations*, *JCAP* **07** (2024) 029, [[2309.03943](#)].
- [94] M. Walther et al., *Emulating the Lyman-Alpha forest 1D power spectrum from cosmological simulations: new models and constraints from the eBOSS measurement*, *JCAP* **05** (2025) 099, [[2412.05372](#)].
- [95] C. Preston, K. K. Rogers, A. Amon and G. Efstathiou, *Prospects for disentangling dark matter with weak lensing*, *Mon. Not. Roy. Astron. Soc.* **2698** (2025) 2713, [[2505.02233](#)].
- [96] R. Blumenhagen, V. Braun, T. W. Grimm and T. Weigand, *GUTs in Type IIB Orientifold Compactifications*, *Nucl. Phys. B* **815** (2009) 1–94, [[0811.2936](#)].
- [97] M. Fairbairn, L. Heurtier and M. O. Olea-Romacho, *Is Λ CDM on the run? Reconciling the CMB with the Lyman- α Forest*, **2511.01612**.
- [98] J. M. Delouis, L. Pagano, S. Mottet, J. L. Puget and L. Vibert, *SRoll2: an improved mapmaking approach to reduce large-scale systematic effects in the Planck High Frequency Instrument legacy maps*, *Astron. Astrophys.* **629** (2019) A38, [[1901.11386](#)].
- [99] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic*

- Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **07** (2011) 034, [[1104.2933](#)].
- [100] V. Poulin, T. L. Smith, D. Grin, T. Karwal and M. Kamionkowski, *Cosmological implications of ultralight axionlike fields*, *Phys. Rev. D* **98** (2018) 083525, [[1806.10608](#)].
- [101] T. L. Smith, V. Poulin and M. A. Amin, *Oscillating scalar fields and the Hubble tension: a resolution with novel signatures*, *Phys. Rev. D* **101** (2020) 063523, [[1908.06995](#)].
- [102] N. Lee and Y. Ali-Haïmoud, *HYREC-2: a highly accurate sub-millisecond recombination code*, *Phys. Rev. D* **102** (2020) 083517, [[2007.14114](#)].
- [103] U.-H. Zhang and T. Chiueh, *Cosmological Perturbations of Extreme Axion in the Radiation Era*, *Phys. Rev. D* **96** (2017) 063522, [[1705.01439](#)].
- [104] A. Arvanitaki, S. Dimopoulos, M. Galanis, L. Lehner, J. O. Thompson and K. Van Tilburg, *Large-misalignment mechanism for the formation of compact axion structures: Signatures from the QCD axion to fuzzy dark matter*, *Phys. Rev. D* **101** (2020) 083014, [[1909.11665](#)].
- [105] H. Winch, R. Hlozek, D. J. E. Marsh, D. Grin and K. K. Rogers, *Extreme axions unveiled: A novel fluid approach to cosmological modeling*, *Phys. Rev. D* **110** (2024) 043517, [[2311.02052](#)].
- [106] D. J. E. Marsh, D. Grin, R. Hlozek and P. G. Ferreira, *Axiverse cosmology and the energy scale of inflation*, *Phys. Rev. D* **87** (2013) 121701, [[1303.3008](#)].
- [107] R. Hlozek, D. J. E. Marsh and D. Grin, *Using the Full Power of the Cosmic Microwave Background to Probe Axion Dark Matter*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 3063–3085, [[1708.05681](#)].
- [108] M. P. Hertzberg, M. Tegmark and F. Wilczek, *Axion Cosmology and the Energy Scale of Inflation*, *Phys. Rev. D* **78** (2008) 083507, [[0807.1726](#)].
- [109] H. Prince and J. Dunkley, *Data compression in cosmology: A compressed likelihood for Planck data*, *Phys. Rev. D* **100** (2019) 083502, [[1909.05869](#)].
- [110] PLANCK collaboration, N. Aghanim et al., *Planck 2018 results. V. CMB power spectra and likelihoods*, *Astron. Astrophys.* **641** (2020) A5, [[1907.12875](#)].
- [111] EBOSS collaboration, S. Chabanier et al., *The one-dimensional power spectrum from the SDSS DR14 Ly α forests*, *JCAP* **07** (2019) 017, [[1812.03554](#)].
- [112] S. Goldstein, J. C. Hill, V. Iršič and B. D. Sherwin, *Canonical Hubble-Tension-Resolving Early Dark Energy Cosmologies Are Inconsistent with the Lyman- α Forest*, *Phys. Rev. Lett.* **131** (2023) 201001, [[2303.00746](#)].
- [113] M. Cicoli, V. Guidetti, N. Righi and A. Westphal, *Fuzzy Dark Matter candidates from string theory*, *JHEP* **05** (2022) 107, [[2110.02964](#)].
- [114] A. Lewis, *GetDist: a Python package for analysing Monte Carlo samples*, *JCAP* **08** (2025) 025, [[1910.13970](#)].
- [115] J. D. Hunter, *Matplotlib: A 2D Graphics Environment*, *Computing in Science and Engineering* **9** (May, 2007) 90–95.
- [116] C. R. Harris et al., *Array programming with NumPy*, *Nature (London)* **585** (Sept., 2020) 357–362, [[2006.10256](#)].
- [117] P. Virtanen et al., *SciPy 1.0: fundamental algorithms for scientific computing in Python*, *Nature Methods* **17** (Feb., 2020) 261–272, [[1907.10121](#)].
- [118] G. Grilli di Cortona, E. Hardy, J. Pardo Vega and G. Villadoro, *The QCD axion, precisely*, *JHEP* **01** (2016) 034, [[1511.02867](#)].
- [119] C. Abel et al., *Measurement of the Permanent Electric Dipole Moment of the Neutron*, *Phys. Rev. Lett.* **124** (2020) 081803, [[2001.11966](#)].
- [120] A. Strominger, *Yukawa Couplings in Superstring Compactification*, *Phys. Rev. Lett.* **55** (1985) 2547.
- [121] P. M. H. Wilson, *The kähler cone on calabi-yau threefolds*, *Inventiones mathematicae* **107** (1992) 561–583.
- [122] Y. Hayakawa, *Degeneration of calabi-yau manifold with w-p metric*, 1995.
- [123] D. R. Morrison and C. Vafa, *Compactifications of F theory on Calabi-Yau threefolds. 2.*, *Nucl. Phys. B* **476** (1996) 437–469, [[hep-th/9603161](#)].
- [124] T.-m. Chiang, B. R. Greene, M. Gross and Y. Kanter, *Black hole condensation and the web of Calabi-Yau manifolds*, *Nucl. Phys. B Proc. Suppl.* **46** (1996) 82–95, [[hep-th/9511204](#)].
- [125] C. Cordova, *Decoupling Gravity in F-Theory*, *Adv. Theor. Math. Phys.* **15** (2011) 689–740, [[0910.2955](#)].
- [126] P. S. Aspinwall and J. Louis, *On the ubiquity of K3 fibrations in string duality*, *Phys. Lett. B* **369** (1996) 233–242, [[hep-th/9510234](#)].
- [127] G. Curio and R. Y. Donagi, *Moduli in N=1 heterotic / F theory duality*, *Nucl. Phys. B* **518** (1998) 603–631, [[hep-th/9801057](#)].
- [128] P. Corvillain, T. W. Grimm and I. Valenzuela, *The Swampland Distance Conjecture for Kähler moduli*, *JHEP* **08** (2019) 075, [[1812.07548](#)].
- [129] S.-J. Lee, W. Lerche and T. Weigand, *Emergent strings from infinite distance limits*, *JHEP* **02** (2022) 190, [[1910.01135](#)].
- [130] M. Cicoli, C. P. Burgess and F. Quevedo, *Fibre Inflation: Observable Gravity Waves from IIB String Compactifications*, *JCAP* **03** (2009) 013, [[0808.0691](#)].
- [131] R. K. Lazarsfeld, *Positivity in algebraic geometry I: Classical setting: line bundles and linear series*, vol. 48. Springer, 2017.
- [132] T. Weigand, *F-theory*, *PoS TASI2017* (2018) 016, [[1806.01854](#)].
- [133] J. Kollár, *Deformations of elliptic calabi-yau manifolds*, *Recent advances in algebraic geometry* **417** (2015) 254–290.