Superstring theories are very promising theoretically, but the enormous landscape of string vacua and the (likely) very large underlying string scale imply that they may never be tested directly. Nevertheless, concrete constructions consistent with the observed world frequently lead to observable remnants, i.e., new particles or features that are apparently accidental consequences of the ultraviolet theory and that are typically not motivated by specific shortcomings of the standard models of particle physics or cosmology. For example, moduli, axions, large extended gauge sectors, additional $Z'$ gauge bosons, extended Higgs/Higgsino sectors, and quasi-chiral exotics are extremely common. They motivate alternative cosmological paradigms and could lead to observable signatures at the LHC. Similar features can emerge in other standard model extensions, but in the stringy case they are more likely to occur in isolation and not as part of a more complete TeV-scale structure. Conversely, some common aspects of the infinite “landscape” of field theories, such as large representations, are expected to be very rare in the string landscape, and observation of features definitively in the swampland could lead to falsification. In this article, common stringy remnants and their phenomenology are surveyed, and implications for indirectly supporting or casting doubt on string theory are discussed.
1. Introduction

Superstring theory\(^1\) yields a consistent and finite unification of quantum theory, gravity, and other interactions, at least at the perturbative level, and is therefore a promising candidate for an ultimate unified theory. However, it is not clear whether Nature actually takes advantage of string theory, or whether there is any way to confirm or falsify it.\(^2\) Part of the problem is that the fundamental string scale is most likely much larger than can ever be directly probed experimentally.\(^3\) Moreover, there are actually five superstring theories, each of which is consistent in 10 space-time dimensions. These are connected to one another by a number of dualities, and may all be different limits of a strongly coupled M-theory (which also contains 11-dimensional supergravity) that is only partially understood.

Further complications arise since there appears to be an enormous landscape of possible string vacua, each involving different low energy physics, with no known principle for the dynamical selection of vacua similar to ours. This situation is not unique to string theory: landscapes of vacua or metastable states naturally exist elsewhere in physics such as protein folding \cite{11} and spin glasses (e.g., \cite{12}). They also exist elsewhere in high energy physics in quantum field theories with multiple scalar fields, including many extensions of the standard model, or even within the standard model itself if one allows one or more of the space dimensions to be compact \cite{13}.

The situation can be argued to be even worse in the conventional field theory framework. Even putting aside the question of quantum gravity, restricting to four-dimensional renormalizable theories, and ignoring possible compactifications, there is apparently an infinite landscape of possible field theories, characterized by different particles, interactions, symmetries, and parameters. There is again no known selection principle to choose between them.

Furthermore, it has been argued that the landscape of string vacua may be surrounded by a much larger “swampland” of apparently-consistent effective field theories that are in fact inconsistent with quantum gravity or string theory \cite{14}.\(^4\) From this perspective, the low energy limit of string vacua yields a proper subset of the possible field theories, due to some additional consistency conditions and constraints on possible (or at least likely) representations and groups. The observation of new physics in violation of these constraints would go far towards falsifying the existence of an underlying string theory.

Conversely, string vacua that contain the standard model (SM) or MSSM and that are qualitatively consistent with what we know may contain apparently unmotivated remnant particles or features, with some types occurring frequently in the string landscape. Observation of some of these characteristic remnants, plus the enormous theoretical advantage of a finite quantum gravity, would lend considerable support to the notion of string theory.

String remnants are in sharp contrast with most bottom-up and other non-stringy investigations, where new physics is motivated by attempting to solve problems of the standard model, or

---

\(^1\)For a review, see, e.g., \cite{1}.

\(^2\)For general discussions of some possibilities, see \cite{2–10}.

\(^3\)This problem is likely to be shared by most alternative theories of quantum gravity.

\(^4\)Examples include theories with an exact continuous global symmetry \cite{15} or in which a gauge symmetry is weaker than gravity (the “weak gravity conjecture”) \cite{16}; see \cite{17–22} for recent weak gravity conjecture studies.
by notions of minimality, elegance, or naturalness.\textsuperscript{5} String remnants, on the other hand, need not solve any problems, might not appear to be minimal or elegant, and might or might not be consistent with notions of naturalness. Instead, they are motivated simply by their frequent occurrence in string theory, for example in surveys of semi-realistic vacua containing the standard model or MSSM. Compactifications involving any or certain classes of remnants are often discarded in many papers that search for the MSSM (and, possibly, common bottom-up extensions), despite the fact that they are often experimentally allowed.\textsuperscript{6}

Our attitude, on the other hand, is that such remnants could well exist, which would be a clue that there is an underlying string theory. Of course, such things as extended gauge, scalar, and fermion sectors are not unique to string vacua, and many possible standard model extensions could emerge either as remnants or from more conventional motivations. The best that we can do here is to emphasize (given current understanding) that some possible standard model extensions occur very frequently in the string landscape, while others are absent or very rare, and to contrast with bottom-up approaches.

For example, in this paper we are mainly concerned with those vacua involving a perturbative connection between the low energy and string scales, in contrast to the many bottom-up models that involve some form of strong coupling or compositeness at the (multi-)TeV scale, often with a fairly large number of interactions and elementary or composite particles.

By a perturbative connection\textsuperscript{7} between the low energy and string scales, we mean that the standard model and cosmological degrees of freedom are essentially the ones that emerge directly from the string compactification, rather than as composite objects associated with some new strong interactions. Unlike most conventionally-motivated bottom-up models, these constructions frequently include remnants that appear ad hoc and isolated, not part of a more complete low-energy theory, and not necessarily associated with a standard model problem.

In this paper we consider the common types of remnants that are likely to be found in semi-realistic string vacua with a perturbative standard model sector, and some of their phenomenological consequences. We will also consider cosmological implications of moduli and axions, and we are mainly but not exclusively concerned with the case of a large string scale. By semi-realistic we mean vacua that are not grossly inconsistent with observation, and in cases where particle (rather than cosmological) sectors are being considered we mean that the vacua contain at least the standard model or MSSM spectrum. We do not, however, insist that such details as the fermion masses and mixings are correctly predicted. There are few, if any, known examples of constructions that successfully predict such detailed features, due in part to the difficulty of moduli stabilization.

\textsuperscript{5}There, some form of uniqueness is often implicitly assumed, but in practice that is mostly a function of the gauge groups, representations, and other symmetries that are invoked. As already emphasized there is no known selection principle or consistency condition leading to a unique field theory. Furthermore, the standard model itself is very complicated, most viable extensions do not appear to be very minimal, and the notion of elegance is rather subjective. For further discussion, see, e.g., [9, 10].

\textsuperscript{6}Examples of papers that focus on general surveys of the landscape include [23–54]. Studies that are more focused on obtaining the MSSM or grand unified theories, often with remnants, include [55–87].

\textsuperscript{7}The restriction to a perturbative extension is motivated by practicality. If we were to observe a strong coupling extension of the standard model it would be extremely difficult to penetrate through the details well enough to connect to an underlying string theory, even though such sectors may arise in the landscape. Of course, other sectors, such as those that could be associated with supersymmetry breaking or dark matter, could be strongly interacting.
Moreover, some or all Yukawa couplings may well be essentially random features of the particular string vacuum, and it is not clear whether these would be strongly correlated with remnants. Even if string theory is correct it is unlikely that anyone will ever find the exact vacuum corresponding to our world. We therefore do not attempt to give fully realistic examples, but instead focus on general ideas and features suggestive of string theory.

Finally, if there really is an exponentially large number of string vacua\(^8\), and assuming something like eternal inflation, then one must seriously consider the possibility of environmental selection (e.g., as an explanation of apparent fine-tunings and fortuitous accidents) and the multiverse.\(^9\) We will briefly mention such notions, but the thrust of this paper is instead towards apparently ad hoc extensions of the standard model and the possibility that they could provide evidence for string theory. These remnants are likely concrete lessons from string theory regardless of the extent to which anthropic selection plays a role in landscape.

There is also the practical question of how to study such large datasets, which is discussed in section 2.2.

The plan of the paper is as follows. The main body is a non-technical description of some of the possible types of string theory remnants and their phenomenological implications. In particular, Section 2 contains a qualitative description of the various classes of string constructions, such as heterotic, Types IIA, IIB, and F-theory. Section 3 discusses the major classes of possible string remnants and their phenomenology. Included are moduli and axions; extended gauge sectors, especially \(U(1)\)'s; extended particle sectors, including Higgs, Higgsinos, quasi-chiral exotics, the absence of large representations, and (quasi-)hidden sectors; implications for couplings and hierarchies, such as leptoquarks, diquarks, family nonuniversality and flavor change, Yukawa interactions/hierarchies, and nonstandard mechanisms for neutrino mass; implications for grand unification and gauge unifications, a low string scale, and environmental selection; and more exotic possibilities. Section 4 is a brief discussion and summary. The appendices contain theoretical details of some of the possibilities, including dark sectors in F-theory, axions, moduli, and stringy consistency conditions with implications for exotics. Most of the content in the appendices was presented by J.H. at TASI 2017.

2. String Theory

2.1 General Features

In this section we provide a brief qualitative discussion of string theory and its compactifications. For each major topic introduced we cite critical original literature or reviews to aid the reader in further study.

Superstring theory is a quantum theory of gravity that is a leading candidate for unification. It naturally exhibits a massless graviton in its spectrum, and the amplitudes for two to two graviton

\(^8\)As opposed to an exponentially large number of “compactification possibilities” that may give rise to a much smaller number of vacua. It is certainly true that large numbers of possibilities exist, but concretely demonstrating the existence of metastable vacua has thus far been limited to regimes of control with low numbers of moduli.

\(^9\)Recent general discussions include [47, 88–91].
scattering are finite in perturbation theory. Its low energy effective action contains the Einstein-Hilbert action crucial to general relativity, as well as other fields. Superstring theory exists in ten dimensions, and there are five consistent possibilities that are known as heterotic $E_8 \times E_8$, heterotic $SO(32)$, type I, type IIA, and type IIB. Non-abelian gauge sectors may arise in each of these theories, due either to fundamental excitations of closed strings propagating in ten dimensions, or to open strings (or their generalizations) that end on solitonic objects collectively referred to as branes, or D-branes in more specialized cases.

These theories are related to one another by a number of dualities. For example, T-duality exchanges two string theories compactified on circles of radius $R$ and $1/R$; type IIA and type IIB are T-dual to one another, as are heterotic $E_8 \times E_8$ and heterotic $SO(32)$. In the second superstring revolution of the mid-1990’s dualities between weakly coupled and strongly coupled theories were also demonstrated, known as S-dualities. The type I and heterotic $SO(32)$ theory are S-dual to one another, and the type IIB theory is self-dual under S-duality. More generally, the type IIB theory allows for the string coupling to vary over the extra dimensions, a possibility known as F-theory [92]. The weakly coupled type IIA and heterotic $E_8 \times E_8$ theories are weakly coupled limits of an eleven-dimensional theory known as M-theory [93], which is a quantum ultraviolet completion of eleven-dimensional supergravity. Via this web of dualities, the different superstring theories in ten dimensions are different limits of one theory. For canonical textbooks on superstring theory and each duality, see [1, 94–98].

The universe that we observe around us has four apparent dimensions. However, as shown by Kaluza and Klein in the 1920’s, a higher-dimensional theory may appear to have fewer dimensions at low energies if the extra dimensions are small and compact. This general result may be applied to superstring theory: if some of its ten dimensions are along a compact $d$-manifold, the low energy effective physics would be $(10 - d)$-dimensional, in which case the choice $d = 6$ recovers a four-dimensional universe. In many cases this low energy effective physics is well approximated by a low energy effective Lagrangian. One of the most studied possibilities is to arrive at four-dimensional physics with $\mathcal{N} = 1$ supersymmetry\(^{10}\) via compactification on a Calabi-Yau threefold (CY3) [99], which has complex dimension three or real dimension six.

Given the framework of compactification, it is natural to consider the compactification of each of the superstring theories, of F-theory, or of M-theory. This has been the subject of a research program over the last thirty years that is sometimes referred to as string phenomenology, though often the research is formal in nature and is not immediately concerned with particle physics or cosmology. Given the choice of one of the higher-dimensional theories there are often many four-dimensional solutions, which are often referred to as the string landscape.

The remainder of this section contains a more technical treatment of some well-studied possibilities that preserve $\mathcal{N} = 1$ supersymmetry at leading order. The reader who is uninterested in these details may go directly to the phenomenological discussion in Section 3.

- **Heterotic on CY3.** Compactification of the heterotic string [100] on a Calabi-Yau threefold

\(^{10}\)Chiral fermions arise naturally from theories with $\mathcal{N} = 0$ and $\mathcal{N} = 1$ supersymmetry, which is part of the reason for the focus on $\mathcal{N} = 1$ supersymmetry.
$X$ is one of the first compactification classes that was studied. In addition to the choice of $X$, one must also make the choice of a gauge field background, as encoded in a holomorphic vector bundle $V$ on $X$. The four-dimensional gauge group $G$ is the commutant of the structure group of $V$ inside $E_8$. There are two central classes of models:

- **Standard Embedding.** In this case $V = TX$, the tangent bundle to $X$, in which case the structure group is $SU(3)$ and $G = E_6 \times E_8$. The $E_8$ gauge sector is a hidden sector and the $E_6$ factor may therefore give rise to grand unification, where the number of chiral particle generations is $\chi(X)/2$, where $\chi(X)$ is the Euler character of $X$.

- **Non-standard Embedding.** In this case $V \neq TX$, and there are a number of constructions for $V$, such as spectral covers or monads. The structure group, and therefore the four-dimensional gauge group $G$, varies depending on topological choices. $SU(5)$ or $SO(10)$ grand unification may arise from the choice of $SU(5)$ or $SU(4)$ structure group, respectively.

In both cases the low energy effective action contains scalar fields that are massless at leading order. The scalar fields that correspond to metric deformations of $X$ that preserve the Calabi-Yau condition are known as complex structure or Kähler moduli, and the scalar fields that correspond to holomorphic deformations of $V$ that preserve the Hermitian Yang-Mills equations are known as vector bundle moduli. Models with the standard embedding may in certain cases be specializations in the vector bundle moduli space of non-standard embedding models.

- **Type II Compactifications with Orientifolds, D-branes, and Fluxes.** Compactifications of the type II superstring theories on a Calabi-Yau threefold $X$ preserve $\mathcal{N} = 2$ supersymmetry. However, D-branes and negative tension objects called orientifold planes may be introduced into the compactification that break half of the supersymmetry, yielding a 4d, $\mathcal{N} = 1$ compactification. Organizing according to the origin of gauge sectors, subclasses of these compactifications include

  - **Type IIA Intersecting Branes.** Here an O6-plane is located on the fixed point locus of an antiholomorphic involution on $X$, and stacks of spacetime filling D6-branes exist to cancel the Ramond-Ramond charge of the O6-plane. Diverse gauge sectors arise from the D6-branes and matter charged under the gauge groups arises from intersections of D6-brane stacks in $X$, which occur at points.

  - **Type IIB Compactifications with D7-branes.** In these compactifications an O7-plane is located on the fixed point locus of a holomorphic involution on $X$ and spacetime filling D7-branes exist to cancel the Ramond-Ramond charge of the O7-plane. D7-branes may intersect, though they now do so along holomorphic curves, which have real dimension two, unlike points. Vector bundles or worldvolume fluxes on the branes give rise to chiral matter at their intersections, as can be understood in simple cases via T-duality, which exchanges the number of intersection points of D6-brane stacks with properties of worldvolume fluxes on seven-branes.
– **Type IIB Compactifications with Magnetized D9-branes.** Here gauge sectors arise from spacetime filling D9-branes and D5-branes, and again appropriate Gauss laws must be satisfied to ensure the net cancellation of Ramond-Ramond charge in the extra dimensions. The structure of chiral matter from the D9-branes depends on the choice of bundle or magnetic fluxes on the D9-branes. In simple cases, these can be understood from two T-dualities of configurations with D7-branes.

– **Type IIB with D3-branes at Singularities.** X may have point-like singularities, such as orbifold points. A D3-brane at or near such a singular point gives rise to a non-trivial quiver gauge theory, the structure of which is determined by the structure of the singularity. Such D3-brane sectors may also arise in compactifications with D7-branes.

For reviews of this class of models, see [72] and the very thorough [40]. For D3-branes at singularities, see [101, 102]. For generalized structures in type II compactifications that lead to \( \mathcal{N} = 1 \) vacua, see [103, 104].

Additionally, background fluxes may be turned [105, 106] on in type IIB compactifications that may stabilize the moduli of \( X \); see [30, 107] for studies of type IIA flux compactifications.

The best understood scenarios are type IIB flux compactifications with three-form background fluxes, which are used in the KKLT [108] and Large Volume Scenario [109] (LVS) moduli stabilization schemes. The number of choices of such fluxes depends on \( X \) but is exponentially large and may give rise to a small cosmological constant [110]. Given the large number of possibilities and the associated computational complexity, a statistical approach [111] has been advocated.

• **F-theory on elliptic CY4.** Weakly coupled type IIB orientifold compactifications are a subset of a larger set of constructions known as F-theory, which may have weakly coupled limits [112, 113]. However, in one of the two largest known ensembles it was shown [114] that such limits essentially never exist.

In F-theory the axiodilaton field of type IIB compactifications is treated as the complex structure of an auxiliary elliptic curve that varies (holomorphically) over the compact extra dimensions \( B \). Since the axiodilaton determines the string coupling \( g_s \), the latter also varies over the spacetime. The appropriate geometric structure is a CY4 \( X \) that is elliptically fibered over \( B \). The structure of four-dimensional gauge theories depends on the singular fibers of \( X \), which encodes information about seven-branes stacks, and also fluxes on the seven-branes that determine the chiral spectrum. Seven-branes in F-theory are more general than those of weakly coupled type IIB compactifications; for example, they may contain exceptional gauge groups.

The generalized seven-brane structure gives rise to interesting possibilities for grand unification [79, 80] that do not exist at weak coupling, such as a perturbative top-quark Yukawa coupling. See [83, 115] and references therein for additional discussion of grand unification in F-theory. See appendix A.1 for a more detailed introduction to F-theory, including non-Higgsable clusters, or [83, 115].
• **M-theory on $G_2$ manifolds.** Since M-theory is an eleven-dimensional theory, obtaining a four-dimensional compactification requires that the extra dimensions of space have seven dimensions. To obtain a 4d theory with $\mathcal{N} = 1$ supersymmetry, the seven-dimensional analog of a Calabi-Yau threefold is a seven-manifold $X$ with $G_2$ holonomy.

If $X$ is smooth and large, the 4d low energy effective theory has $\mathcal{N} = 1$ supersymmetry and has $b_2(X)$ massless abelian vector multiplets, but no charged matter, and $b_3(X)$ massless chiral multiplets arising from metric deformations that preserve the $G_2$ holonomy and the three-form $C_3$ of eleven-dimensional supergravity. If $X$ is fibered by a circle, it becomes a type IIA compactification with O6-planes and intersecting D6-branes in the limit that the circle becomes small.

Obtaining a non-abelian gauge sector with chiral matter therefore requires taking a limit in which $X$ becomes singular, in which case gauge sectors may arise along real codimension four ADE singularities, with vector pairs of charged chiral multiplets arising from codimension six singularities and charged chiral matter arising at codimension seven singularities. See the review [116] and references therein.

Each of these constructions has its own nuances and specialized considerations, but some low energy implications arise out of many of the constructions, which is the idea behind remnants and the subject of this review.

There are other well-studied constructions beyond those discussed here. For an introduction to those, as well as a thorough introduction to all of the compactification classes mentioned in this section, we refer the reader to the excellent textbook [1].

### 2.2 Remarks on Universality, Machine Learning, and the Landscape

There is also the practical question of how to study datasets that are as large as the ones that arise in string theory, where the number of data points is far larger than the $O(10^{78})$ protons that exist in the visible universe. Answering it is of great importance for understanding remnants.

As an analogy, consider large energy scales in particle physics. Successive generations of colliders probe successively higher energy scales, but there is a demarcation line for what constitutes a “large” energy scale, which is one so large that no conceivable collider will ever probe it directly. For many reasons, including uncertainties in the rate of progress in accelerator physics and in science funding, the value of this demarcation line is not clear. However, physicists know values above this line when they see it: we will almost certainly never directly probe the GUT scale or Planck scale, or the string scale if it is large.

Similarly, one good notion of what constitutes a large dataset, including in the context of the string landscape, is a dataset so large that no conceivable computer will ever be able to hold it in memory. Since each data point must be at least one bit of information, the geometric ensemble of [53] discussed in the appendix and current estimates for the number of flux vacua [45] give:

\[
\text{Memory required for string geometries} \geq 10^{745} \text{ GB}
\]

\[
\text{Memory required for flux vacua} \geq O(10^{272,000}) \text{ GB}.
\]
While one of these numbers is clearly much larger, we mention the smaller one because it arises from an exact lower bound on the number of topologically distinct geometries possible in string compactifications\textsuperscript{11}, i.e., it is an extremely large number from an exact combinatoric calculation. Just as physicists know that the GUT scale and Planck scale are “large” in the sense discussed, we know that these datasets are large: it is highly unlikely that any computer will ever be able to hold them in memory.

Another interesting possibility, a priori, would be to use so-called streaming algorithms to systematically read through the data and keep track of certain properties, without storing it in memory. However, under the assumption that reading each data point in the stream takes at least a Planck time, which is $O(10^{-61}) \ T_{\text{univ}}$, we obtain

$$\text{Time required for streaming string geometries} \geq 10^{684} \ T_{\text{univ}},$$

$$\text{Time required for streaming flux vacua} \geq O(10^{272,000}) \ T_{\text{univ}},$$

where $T_{\text{univ}} \simeq 10^{10}$ years is the age of the universe. Clearly both memory issues and time issues come to the fore, as of course one would have expected.

For these reasons a statistical approach \textsuperscript{[111]} to the landscape seems critical, and a natural question is how to study statistics (including possible vacuum and/or anthropic selection) of a set of vacua that cannot be read into memory or scanned over, particularly given that the difficulty of the problem is exacerbated by computational complexity \textsuperscript{[117–119]}.

One possibility is that some large ensembles may allow for a high degree of control, including potential demonstration of universality, due to the existence of a concrete construction algorithm from which statistical properties can be extracted; see \textsuperscript{[53]}. This technique is clearly useful when applicable, since analytic results are almost always easier to understand than numeric results. However, many problems may not admit such a concrete construction algorithm.

Using techniques from data science that have revolutionized other fields, such as supervised machine learning, genetic algorithms, or deep reinforcement learning, may also be a promising direction; see \textsuperscript{[120–124]} for initial studies in string theory. One might object that such techniques cannot be used on datasets that cannot fit into computer memory, but in fact this is not necessarily true, depending on the task at hand. For example, AlphaZero \textsuperscript{[125]} is the state of the art Go program that recently defeated AlphaGo Lee 100-0, the program that itself defeated the human world champion for the first time, in 2016. It did so with under 48 hours of training for its neural networks, with no human knowledge input into the training (i.e., training only with self-play), despite the fact that Go has $1.4 \times 10^{172}$ board positions, far more than can fit into memory. The situation is helped by the fact that the “dynamics” of playing the game well accesses a small subset of the board positions, but it may also be so with the dynamics of the landscape. Furthermore, in a concrete example in string theory (the $10^{755}$ ensemble) it is already the case \textsuperscript{[124]} that machine learning has led to a statistical result that had not been arrived at by analytic means. In that case, study of the machine learned model led to a conjecture that was turned into a rigorous analytic theorem.

\textsuperscript{11}Under the assumption that a consistent flux can be turned on in an $O(1)$ fraction of the geometries. In fact, each geometry likely admits many consistent fluxes, since it is essentially putting a Bousso-Polchinski \textsuperscript{[110]} type problem on top of each geometric choice. This is an interesting direction for further study.
It would be interesting to use any of these techniques to make further concrete statements about string remnants across such large (in the sense discussed) datasets in the landscape.

3. Typical Stringy Remnants

In this section we survey typical stringy remnants frequently encountered in semi-realistic constructions, commenting briefly on their theory, phenomenology, and implications.

3.1 Moduli and Axions

Scalar moduli and pseudoscalar axions are two of the most common remnants that arise in string compactifications. For additional discussion related to the content of this section, we refer the reader to appendices A.2 and A.3.

Moduli are often (but not always) associated with metric deformations of the extra dimensional manifold that preserve a property related to supersymmetry; for example, the Calabi-Yau condition of a Ricci-flat Kähler metric on a six-manifold. Axions typically arise from Kaluza-Klein reduction of higher form gauge fields, and the gauge symmetry of the higher form gauge field becomes a shift symmetry of the axion that survives to all orders in perturbation theory.

The relevance of moduli and axions arises from the fact that they couple to many fields in the low energy effective action, and therefore the process that cosmologically reheats the visible sector (such as inflationary reheating) likely also transfers some energy density into moduli or axions. Both possibilities may involve deviations from standard cosmological paradigms that were arrived at from the bottom-up.

The typicality of moduli and axions follows in part from concrete properties of compactification manifolds, with the implication that there are usually dozens or hundreds of moduli and / or axions arising in a string compactification.

3.1.1 Moduli Domination

If reheating transfers energy into the moduli sector, on long time-scales the modulus will begin to oscillate around its minimum, in which case it behaves like non-relativistic matter. Then the energy density of the modulus \( \rho_\phi \) scales as

\[
\rho_\phi \propto \frac{1}{a(t)^3},
\]

where \( a(t) \) is the scale factor of a Friedmann-Robertson-Walker (FRW) cosmology. In the standard cosmological paradigm reheating gives rise to a phase of radiation domination that carries through big bang nucleosynthesis (BBN) and eventually gives rise to matter domination at the time of matter-radiation equality (MRE). However, a reheated modulus will often come to dominate the energy density of the universe prior to MRE due to its scaling relative to the scaling \( \rho_{rad} \propto 1/a^4 \) of radiation in the visible sector. This early phase of matter domination due to moduli domination spoils the successful predictions of BBN if the universe is matter dominated at the would-be BBN time; this is known as the cosmological moduli problem [126–128]. This is avoided in string-motivated scenarios where the modulus decays via a dimension-five Planck-suppressed operator if \( m_\phi \gtrsim 50 \) TeV.
There are relatively few constraints on such a phase of matter domination prior to nucleosynthesis, and if it arises due to modulus domination new cosmological paradigms arise, such as the production of dark matter via modulus decay (e.g., [129–133]) rather than thermal freeze-out. Observables of inflationary models can be modified depending on whether they have a standard or moduli dominated cosmology prior to BBN [134].

See the review [135] and appendix A.3 for further discussion. See also [136] for recent work on localized, long-lived, non-linear excitations of moduli known as oscillons.

3.1.2 Axion Inflation, Dark Matter, and Dark Radiation

Broadly, axions or axion-like particles (ALPS) are to be expected in string compactifications, forming a so-called axiverse [137]. Their basic phenomenology is determined by their couplings, decay constants, and masses. Distributions of axion decay constants tend to be centered an order of magnitude or two below $M_p$, while their masses may take a variety of values; see, e.g., [138, 139]. Accordingly, they may play a number of qualitatively different roles, and we categorize the possibilities according to three such roles.

Models of inflation where the inflaton is an axion have been extensively studied. Axions are one of the few particle types that may successfully give rise to large-field inflation in string theory, due to the extremely sensitive dependence on large-field inflation on Planck-suppressed operators. These operators are forbidden in perturbation theory by the axion shift symmetry, which plays a critical role in the viability of the models. Two basic categories of models are referred to as axion inflation (see, e.g., [140–142]), where the inflationary potential is a standard cosine in the axion field, and axion monodromy [143–145], which includes extra terms such as a linear term in the axion potential.

When axions oscillate around their minimum they behave as non-relativistic matter, in which case they may be dark matter candidates. They may behave as standard QCD axion dark matter, in which case the axion couples strongly to the gluon field strengths and is produced by misalignment. However, the upper bound on the QCD axion decay constant $f < 10^{12}$ GeV$^{12}$ is several orders of magnitude smaller than typical decay constant scales in string theory. Alternatively, many of the axions in string theory do not behave as the QCD axion, and they may comprise some or all of the dark matter. See [147] or appendix A.2 for a discussion of the possibility that string axions comprise so-called fuzzy dark matter, which may account for some astrophysical anomalies such as the core vs. cusp problem.

Axions may also behave as dark radiation, in which case it is constrained by bounds on $\Delta N_{\text{eff}}$. See [148–150] for studies of axionic dark radiation in LARGE volume compactifications. Axion-photon scattering during BBN can affect light element abundances [151], and axion dark radiation can constrain neutralino dark matter when it is produced by modulus decay, due to correlations between the physics of axions and moduli [152]. Considerations from the string axiverse imply that a cosmic axion background exists today; see [151]. However, axions arising from the decay of moduli that solve the cosmological moduli problem may still be significantly constrained [153].

---

12This constraint can be weakened in some cases if the initial value of the axion field is fortuitously close to the minimum. See, e.g., [146].
We refer the reader to [147] for a review of axion cosmology, to [154] for a review on inflation in string theory that includes axions, and also to references therein.

3.2 Extended Gauge Sectors

3.2.1 TeV-Scale $Z'$s and other Extended Gauge Symmetries

Many types of new physics lead to the existence of new gauge bosons at the TeV scale or higher [155]. $Z'$s are especially common because it is often harder to break $U(1)'s$ than non-abelian symmetries, i.e., they tend to “slip through” when a larger non-abelian symmetry is broken. They frequently occur not only as string remnants, but also in alternative models of electroweak breaking such as dynamical symmetry breaking or composite Higgs models, and in the breaking of grand unified theories larger than $SU(5)$.

In the string context (e.g., [34, 44, 55, 59, 64, 72, 156, 157]) additional non-anomalous $U(1)'s$ often emerge as unbroken subgroups of non-abelian groups such as $SO(10)$ [158], $E_6$ [159], or the $SU(4) \times SU(2)_L \times SU(2)_R$ Pati-Salam group [160] (any of which can be broken in the extra dimensions). The $U(1)'$ charges are then typically linear combinations of $T_3$, $B - L$, or the $E_6$ charges. Additional $U(1)'s$ may instead be “random”, i.e., not resembling familiar grand unification or other extended gauge symmetries. In the latter case the associated $Z'$s may couple to particles in a hidden or quasi (partially)-hidden sector, to some or all of the standard model particles, or both, and they can sometimes mediate transitions between the sectors by direct couplings or by kinetic mixing. The couplings are often family-nonuniversal, leading to flavor-changing effects. In many cases anomalies are canceled by new quasi-chiral fermions (i.e., fermions that are vector-like under the standard model gauge group).

$U(1)'s$ may also be “anomalous” (in the traditional field theory sense [161]), although the apparent anomalies are actually canceled by Chern-Simons terms to yield consistent theories [1, 162, 163]. For example, heterotic constructions often include one anomalous $U(1)'$. They are especially prevalent in Type II string theories, which involve $U(N)$ gauge groups rather than $SU(N)$ [$Sp(2N)$ and $SO(2N)$ are also possible], with the trace $U(1)$ usually anomalous. The gauge boson associated with an anomalous $U(1)$ generally acquires a string-scale mass by the Stückelberg mechanism [164] and is therefore irrelevant except in the special case of a low string scale. However, the $U(1)'$ symmetry survives as a global symmetry on the perturbative couplings of the low energy theory, possibly broken by non-perturbative string instanton effects [165–167]. Moreover, one or more linear combinations of the anomalous $U(1)'s$ may be non-anomalous due to (accidental) cancellations, with the associated $Z'$s emerging as ordinary gauge bosons that can acquire mass by the Higgs or other mechanisms. The standard model hypercharge often emerges in this way, and other cancellations can lead to remnant $Z'$s. The latter are often family-nonuniversal.

Experimental signatures of a $Z'$

A $Z'$ can be searched for at the LHC [155, 168], e.g., through Drell-Yan $pp \rightarrow Z' \rightarrow \mu^+\mu^-, e^+e^-$, $q\bar{q}, W^+W^-$, or $ZH$. Limits from early LHC running exclude masses below 3-4 TeV for electroweak-strength couplings to ordinary fermions, and should eventually reach 5-6 TeV, with the specific

\[^{13}\text{Such four-dimensional field theories are sometimes referred to as augmented field theories.}\]
value dependent on the $Z'$ gauge coupling, charges, and the effects of decays into exotics and superpartners. Diagnostic studies of the charges, $Z - Z'$ mixing, etc., could in principle be carried out by measurements of branching ratios, asymmetries, polarizations, rapidity distributions, associated production, and rare decays, but in practice this would be difficult above 3-4 TeV. Masses up to $O(30 \text{ TeV})$ could be probed at a future 100 TeV $pp$ collider [169]. $Z'$s could also be searched for (and their couplings probed) at a future $e^+ e^-$ collider, where interference effects with $s$-channel $Z$ and $\gamma$ exchange should be sensitive to masses well above $\sqrt{s}$ [170, 171]. Existing $Z$-pole and weak neutral current results typically limit any $Z - Z'$ mixing angles $|\theta_{Z-Z'}|$ to less than a few $\times 10^{-3}$.

Lighter $Z'$s, e.g., in the 10-1000 GeV range, with couplings much smaller than electroweak or with special couplings (e.g., leptophobic) are also possible [172–174]. These could be searched for in high precision weak neutral current experiments, rare decays, or high-luminosity LHC running. Very light (e.g., $< 10$ GeV or even much smaller) $Z'$s are sometimes invoked in connection with hidden sector dark matter models, often serving as a portal between the dark and ordinary sectors. Their couplings to the ordinary particles would have to be extremely weak, e.g., generated only by kinetic mixing. Very light $Z'$s can also be searched for in beam dump experiments and in astrophysics/cosmology, and could be relevant to the muon magnetic moment anomaly.

Other implications of a TeV-scale $Z'$

The existence of a new $Z'$ (whether or not associated with string theory) would have many other implications. Here we mainly consider a $U(1)'$ with electroweak couplings broken at the (multi-) TeV scale. For more detail see [155].

- The $U(1)'$ charges could give clues as to the embedding into the underlying theory.
- A $U(1)'$ could give a natural solution [175, 176] to the $\mu$ problem [177] of the MSSM.\footnote{Other solutions include the Giudice-Masiero mechanism [178] and string instantons [165, 166].} That is, if the two Higgs doublets $H_u, H_d$ of the MSSM do not form a vector pair then an elementary $\mu H_u H_d$ term in the superpotential would be forbidden, but an effective $\mu_{\text{eff}} = \lambda_S \langle S \rangle$ would be generated if $\lambda_S S H_u H_d$ is allowed, where $S$ is a standard model singlet\footnote{We use the same symbol for both the chiral superfield and its scalar component.} with nonzero charge. $\langle S \rangle$ also generates or contributes to the $Z'$ mass and may generate masses for exotics. This can be viewed as a stringy version of the NMSSM [179, 180], except there is no discrete $Z_3$ symmetry or attendant cosmological domain wall problems. Also, the MSSM bound of around 135 GeV on the lightest Higgs scalar mass is increased to $\sim 150$ GeV by new $F$-term (also present in the NMSSM) and $D$-term contributions to the scalar potential [181].
- More generally, but assuming that the $\mu$ problem is somehow solved, the electroweak and $U(1)'$ breaking scales are typically both set by the soft supersymmetry breaking and $\mu$ scales. Thus, the $Z'$ and related exotic masses are usually not too much larger than the superpartner masses, or to the electroweak scale, up to an order of magnitude or so. This conclusion can be evaded, however, if the $U(1)'$ is broken along a flat or almost flat direction [182] or by fine-tuning.
\item \(U(1)'\) models typically involve one or more standard model singlet fields \(S\) whose expectation values break the symmetry, implying at least one additional Higgs scalar. The extra scalars are often at the TeV scale, but in principle could be lighter. They could mix with the standard model or MSSM Higgs bosons, modifying couplings and their experimentally-allowed ranges [181].

\item Supersymmetric \(U(1)'\) models have an extended neutralino sector [183], including at least the \(Z'\) gaugino (\(\tilde{Z}'\)) and the singlino (\(\tilde{S}\)), modifying couplings or allowing new possible scenarios for collider physics (e.g., extended cascades ending in the production of a light invisible singlino and a \(Z\) or \(H\)), cold dark matter, etc.

\item Non-anomalous \(U(1)'\)'s often involve quasi-chiral exotic fermions that cancel anomalies. Familiar examples are

\[
D_L + D_R, \quad \left( \begin{array}{c} E^0 \\ E^- \end{array} \right)_L + \left( \begin{array}{c} E^0 \\ E^- \end{array} \right)_R,
\]

where \(D_{L,R}\) are an \(SU(2)\)-singlet pair of quarks with electric charge \(-1/3\), and \((E^0E^-)^{L,R}_{L,R}\) are an \(SU(2)\)-doublet pair of leptons or Higgsinos. These occur in the 27 of \(E_6\) along with a standard model family, a right-handed neutrino \(v_R\), and a singlet \(s\) whose scalar partner \(\langle S\rangle\) can break the \(U(1)'\), and are also frequently encountered as string remnants. Their properties are further discussed below.

\item The decays of a heavy \(Z'\) could be an efficient means of producing superpartners, exotics, heavy right-handed neutrinos, etc. (e.g., [184]).

\item Stringy \(Z'\) couplings are often family-nonuniversal, e.g., because the standard model families have different string constructions, because of Kaluza-Klein excitations coupling to fermions with different spatial distributions in large or warped extra dimensions [185]. Nonuniversal couplings usually lead to flavor-changing vertices when fermion mixing is turned on [186]. The limits from rare decays and \(K_L - K_S\) mixing are sufficiently strong that such nonuniversality is probably restricted to the third family. It could possibly be associated with the observed hints of nonuniversality in \(\Gamma(B \rightarrow K \mu^+\mu^-)/\Gamma(B \rightarrow Ke^+e^-)\) (see, e.g., [187–189]), could lead to new tree-level effects that could compete with standard model loops (e.g., \(B_d\) penguins or in \(B_s - \bar{B}_s\) mixing) [190], or to effects in \(\tau\) or \(t\) decays.

\item Additional anomalous or non-anomalous \(U(1)'\)'s can sometimes distinguish between lepton and down-Higgs doublets, which have the same gauge quantum numbers in the (\(R_P\)-violating) MSSM. They can also sometimes distinguish SM-singlet Higgs fields from sterile neutrinos.

\item A \(Z'\) could serve as a portal to a (quasi-)hidden sector, either by direct coupling or kinetic mixing [191]. Similarly, a \(Z' - \tilde{Z}'\) mass difference generated in a supersymmetry breaking sector could mediate supersymmetry breaking to the standard model sector [192, 193].

\item An additional \(U(1)'\) can affect the possibilities for generating small neutrino masses [194]. Depending on the couplings, conventional models such as the Type I seesaw may be allowed
or forbidden, while novel mechanisms such as small Dirac masses by higher-dimensional operators or non-holomorphic soft terms may be allowed. The various possibilities are described in a more general stringy context below.

- There could also be significant implications for cosmology [155]. For example, trilinear $SH^*H$ or $SH_dH_u$ Lagrangian terms could allow a strong first-order electroweak phase transition (as in the NMSSM), necessary for electroweak baryogenesis. Cold dark matter and big bang nucleosynthesis could be modified in a number of ways by the extended neutralino/Higgs/$Z'$ sectors.

**Non-abelian gauge symmetries at the TeV scale**

Additional non-abelian gauge symmetries are also possible, both in string theories and in other types of standard model extension. These can easily act in hidden or quasi-hidden sectors, and could be either weakly or strongly coupled. One well-motivated stringy extension in the standard model sector involves an $SU(2)$ that couples to right-chiral fermions, as in $SU(2)_L \times SU(2)_R \times U(1)$ models (see, e.g., [195–197]). (There need not be any discrete left-right symmetry.) The phenomenological implications depend on whether the $\nu_R$ are light (as in Dirac or pseudo-Dirac neutrino masses) or heavy (as in a seesaw model), and also on the right-handed quark mixing matrix. These include precision measurements of $\beta$ and $\mu$ decay, weak universality, the $K_L - K_S$ mass difference, $M_W$, and possible heavy $W'$ and $\nu_R$ production and decay.

**3.2.2 Extended Non-abelian Gauge Sectors and Hidden Sector Dark Matter**

Many constructions involve one or more (quasi)-hidden sectors of gauge-charged particles that are coupled to the standard model sector via gravity and some combination of moduli, axion, and gauge-charged mediator couplings (e.g., [198–200] and Appendix A.1). We are restricting attention here to a weakly coupled standard model sector, but such hidden or quasi-hidden sectors may well involve strong coupling. In such a vacuum, the standard model may represent a small fraction of the gauge-charged matter that could be excited. There are two basic possibilities:

a) Extended gauge sectors that couple to the standard model only via gravity, axions, and $l$ or moduli. Such interactions with the standard model are very weak due to the fact that the Planck scale or the string scale appears as the effective cutoff in the associated effective operators.

b) Extended gauge sectors that interact with the standard model via a gauge-charged mediator, or a sequence of them if a gauge factor is separated from the standard model by multiple gauge factors. The strength of the interactions with the standard model depend on a number of factors, including the number of mediators and their masses.

These ideas are well-exemplified in a quiver diagram, such as
which is also discussed in appendix A.1. Nodes are gauge factors that are labeled according to their group, and edges represent jointly charged matter. If the compactification realizes the standard model on the $SU(3) \times SU(2)$ factor that is disconnected from the rest of the gauge nodes, the other gauge sectors interact with the standard model only via gravity, axions, and moduli, and are well-hidden. On the other hand, if the standard model is realized on one of the $SU(3) \times SU(2)$ pairs in the large connected component, then adjacent nodes may interact with it via a single mediator, while faraway nodes are more hidden according to the number of intermediate mediators. From this we see that some extended non-abelian gauge sectors are much more hidden than others. In principle one can therefore distinguish between truly hidden sectors and quasi-hidden sectors.

There have been many phenomenological studies of hidden and quasi-hidden sectors. The most familiar are supersymmetry breaking sectors, with supersymmetry breaking mediated to the standard model sector by gravity, anomaly, gauge, $Z'$, or other mechanisms [201].

More recently, there has been considerable interest in possible hidden dark matter sectors [174], often connected by Higgs, $Z'$, neutrino, or axion portals, but sometimes with very weak interactions with the standard model. These and other possibilities may arise from remnants of string theory; see, e.g., [202, 203] for discussions. For example, from extended non-abelian gauge sectors that confine, dark hadrons and / or glueballs may arise that comprise the dark matter. Dark glueball dark matter [204–214] from strongly-coupled Yang-Mills sectors are possible but often problematic [208] if the sector is hidden, rather than quasi-hidden. Dark glueballs may also arise in warped throats [215], which occur frequently (see, e.g., [216]) in the landscape. Another possibility is that dark matter is dynamical [202, 217], existing in multiple components with a variety of lifetimes and abundances that are balanced against one another.

Hidden or quasi-hidden sectors (e.g., [59, 65]) could also be random and not connected to either supersymmetry breaking or dark matter. These would not necessarily have any significant observational consequences, but in some opportune cases there could be “hidden valley” signatures such as lepton jets, displaced vertices, and missing energy [218].
3.3 Particles Beyond the Standard Model

3.3.1 Extended Higgs/Higgsino sectors

Extended Higgs/Higgsino sectors are quite common, with implications for collider physics, CP violation, flavor physics, and cosmology.

Standard model singlets $S$ have already been mentioned in connection with additional $Z'$s, but occur much more generally [34, 44, 219]. Often they can acquire vacuum expectation values and possibly can mix with Higgs doublets. These can break $U(1)'$ and other symmetries, give masses to quasi-chiral exotics (including the $H_{u,d}$ pair, solving the $\mu$ problem and allowing a strong first-order electroweak transition), and can occur in higher-dimensional operators (e.g., to generate suppressed Yukawa couplings). Other singlets can have couplings and quantum numbers that allow them to play the role of right-handed or other sterile neutrinos, or to be dark matter candidates (e.g., [220]). Yet others can be unrelated to either neutrinos or symmetry breaking.

Also very common are extra Higgs doublets, especially vector-like pairs in addition to the single $H_{u,d}$ pair of the MSSM. These and the associated Higgsinos obviously lead to a much richer spectrum and collider phenomenology. They may also be needed for or expand the possibilities for describing quark and lepton masses and mixings. For example, stringy symmetries that differentiate between the fermion families would restrict the perturbative couplings of a single Higgs doublet, so that different elements of the fermion mass matrix could be associated with different doublets. This would allow some of all of the fermion mass hierarchies to be due to hierarchies in the expectation values of the multiple Higgs fields rather than in the Yukawa couplings. However, such models also open up the possibility of flavor changing neutral current effects mediated by neutral Higgs fields [221]. Another difficulty is that extra Higgs doublets, e.g., an extra $H_{u,d}$ pair added to the MSSM, would significantly modify gauge coupling unification unless the doublets are accompanied by other new particles, e.g., to form complete 5-plets of $SU(5)$.

Higgs triplets, especially those with hypercharge 0 [i.e., $(\phi^+, \phi^0, \phi^-)^T$] can also occur.

See the Higgs Boson review in [168] for a detailed discussion of extended Higgs sectors in a general context.

3.3.2 Quasi-Chiral Exotics

One frequently encounters exotics that are non-chiral with respect to the standard model gauge group but chiral under extended gauge, (perturbative) global, or discrete symmetries [34, 44, 222]. Such exotics are often required to cancel anomalies for non-anomalous extended gauge symmetries, or due to string consistency conditions that do not have a simple analog in quantum field theory (see appendix A.4 for a discussion of the latter). They may couple directly to the standard model sector through gauge or Yukawa interactions, but could also be restricted to a hidden sector. If they are charged under the standard model gauge group they generally modify gauge unification unless they occur in full $SU(5)$ multiplets (similar to the comment on additional Higgs doublets).

Common examples of quasi-chiral exotics include the isosinglet down-type quark and isodoublet lepton/Higgsino vector-like pairs\(^{\text{16}}\) in (3.2). These can also be written as left-chiral supermul-

\(^{\text{16}}\)Whether these are best interpreted as exotic leptons, as additional Higgs pairs, as both, or as neither, depends on the allowed couplings and symmetry breaking. Similar statements apply to neutral singlet fields, which may play the role of Higgs particles ($S$), sterile neutrinos ($N^c$), both, or neither.
tiplets,
\[ D + D^c, \quad \begin{pmatrix} E_0^- \\ E^+ \end{pmatrix} + \begin{pmatrix} E_0^- \\ E_0^+ \end{pmatrix}. \quad (3.3) \]

Other commonly encountered pairs are isosinglet up-type quarks \( U + U^c \), isodoublet quarks \( (UD)^T + (D^c U^c)^T \), non-abelian singlets like \( E^- + E^+ \) with \( Y = Q = \mp 1 \), and shifted-charge lepton/Higgsino doublets with charges \((\pm 1, \pm 2)\). Yet other types of pairs, such as color sextets, isosinglet charge \(-4/3\) quarks, or isodoublet quarks with shifted charges \((-1/3, -4/3)\) are also in principle possible, but did not occur in the quiver survey in [44]. The possibility of fractionally-charged color singlets will be discussed separately below.

The scalar components of exotics can acquire masses by soft supersymmetry breaking. However, mass generation for the fermions is sometimes problematic. If the chiral symmetry is broken perturbatively by the expectation values of a standard model singlet \( S \) then the fermions (and the scalars if the breaking scale is larger than the supersymmetry scale) will typically acquire masses related to the symmetry breaking scale. For example, exotic pairs \( X + X^c \) would become massive through couplings like \( SXX^c \) or by higher-dimensional operators involving \( S \) if they are perturbatively allowed. Masses could also be generated by nonperturbative breaking of the chiral symmetries, e.g., by string instantons [165, 166]. In other vacua that are not consistent with our own the fermions may remain massless.

Vector pairs that are non-chiral under all symmetries are a common special case. At generic points in moduli space these would often acquire string-scale masses and are therefore irrelevant for our purposes unless the string scale is low. However, they may become light at subloci in moduli space when the combination of field expectation values that determines the mass of the vector pair goes to zero.

\( SU(2) \times U(1) \) singlets such as \( S \) and \( SU(2) \)-triplets with \( Y = Q = 0 \) can also occur. These are non-chiral under \( SU(2) \times U(1) \) and can occur either in pairs or singly. The fermions could in principle acquire either Dirac or Majorana masses.

Of course, string vacua can also involve additional matter that is standard model chiral but anomaly free. The most obvious possibility is a chiral sequential or mirror fourth family. Another is the shifted family
\[ (3, 2)_{-\frac{2}{3}}, \quad (\bar{3}, 1)_{\frac{1}{3}}, \quad (\bar{3}, 1)_{\frac{1}{2}}, \quad (1, 2)_{-\frac{1}{2}}, \quad (1, 3)_1, \quad (3.4) \]

where the numbers in parentheses are the \( SU(3) \) and \( SU(2) \) representations, and the subscript is \( Y \). However, exotics that are chiral under \( SU(2) \times U(1) \) would have to acquire masses by electroweak breaking. Experimental limits are sufficiently strong that they would require very large Higgs-Yukawa couplings, leading to Landau poles at low energy, violation of tree-level unitarity, and probable conflict with the observed Higgs couplings and electroweak precision tests.

The phenomenology of vector-like exotics has been extensively discussed. Here we mainly follow the supersymmetric treatment in [225]. Colored exotic quarks and squarks could be efficiently pair produced at a hadron collider by QCD processes if their masses are not too large, while color singlets would have much lower rates. Single production by mixing or leptoquark or diquark couplings would be kinematically favored but with smaller couplings. The relative masses of the

\[^{17}\text{A partial list includes [159, 223–248].}\]
quark and two squarks are dependent on the $D$-terms and soft breaking, with the heavier ones cascading downwards by processes like $\tilde{D}_1 \to D \tilde{G}$ or $D \to \tilde{D}_1 \tilde{\chi}^0$, where $\tilde{G}$ and $\tilde{\chi}^0$ represent gluino and neutralino, respectively.

The exotics might be able to decay into standard model particles by mixing. For example, an isosinglet down-type quark might decay into $D \to u_i W^-, D \to d_i Z$, or $D \to d_i H^0$ by $D - d_i$ mixing. However, the extra symmetries often forbid or strongly suppress this. Another possibility is decays via allowed leptoquark or diquark operators, such as $D u e^c$, or $D c u d^c$, respectively.\textsuperscript{18} Of course, proton stability requires that they should not both be present simultaneously at any significant level. The lightest exotic could also be stable at the renormalizable level, decaying only by higher-dimensional operators or nonperturbative effects. Depending on the details, exotics could decay promptly; with displaced vertices; pass out of the detector; or hadronize and stop, decaying later (a form of $R$-hadron \textsuperscript{249}). They could even be approximately stable on cosmological time scales, although that possibility is constrained by big bang nucleosynthesis.

Other possible implications of quasi-chiral exotics include flavor, precision, and Higgs physics; flavor-changing and nonuniversal effects; gauge mediation; and cosmology.

### 3.3.3 Absence of Large Representations

Perturbative string constructions typically do not allow large or unusual representations for the elementary matter multiplets in the low energy four-dimensional gauge theory. For example, simple heterotic constructions only allow singlets, fundamentals (or bifundamentals under two groups), and their conjugates. In some cases this constraint can be evaded if the group is realized at a higher Kač-Moody level \textsuperscript{250, 251}. For example, the diagonal subgroup $G \in G \times G$ corresponds to a level 2 embedding, and can yield adjoints, symmetric and antisymmetric representations, and their conjugates. However, higher embeddings are increasingly complicated and there are restrictions on their size.

Similarly, Type IIA intersecting brane constructions are restricted to singlets, fundamentals, bifundamentals, adjoints, two-index symmetric and antisymmetric representations, and their conjugates. Large representations, trifundamentals, etc., may be possible in F-theory and M-theory, but relatively little investigation has been done on the subject. Large gauge groups are to be expected in F-theory due to the prevalence of non-Higgsable clusters \textsuperscript{53, 252, 253}, however. Since they correlate strongly with being at strong string coupling \textsuperscript{254}, that is, in moving away from weakly coupled type IIB, it may be that such effects are also generic in M-theory, that is, in moving away from weakly coupled type IIA.

### 3.3.4 Fractionally Charged Color Singlets

Fractional electric charges for color singlet particles or bound states are common in string constructions, e.g., \textsuperscript{44, 57, 59, 64, 65, 255–258}, and constitute a problem rather than a signature. These may, for example, be associated with $SU(2)$ singlets with hypercharge $Y = \pm \frac{1}{2}$, nonstandard $Y$, or additional contributions to the electric charge generator $Q$ for exotics.\textsuperscript{19} For example, the

\textsuperscript{18}Unlike the possible leptoquark or diquark operators in the MSSM, these exotic operators still allow for a conserved $R$-parity and stable LSP.

\textsuperscript{19}Small, e.g., millicharge, particles can also be generated by kinetic mixing, such as between the photon and a hidden-sector dark boson \textsuperscript{191}. However, this is not necessarily a stringy effect.
Schellekens’s theorem [259] states that otherwise-realistic weakly coupled heterotic constructions require either fractional charges, an unbroken $SU(5)$, or higher-level Kač-Moody embeddings for $SU(2)$ or $SU(3)$.

The lightest fractionally-charged particle would be stable for an unbroken $Q$, and (except for very small charges) are excluded for almost all masses by direct searches [260] and cosmological constraints such as overclosing the universe [261]. However, these constraints can be evaded if the particles are too heavy to produce following inflation or are confined by additional non-abelian gauge interactions.

### 3.4 Couplings and Hierarchies

#### 3.4.1 Leptoquark, Diquark, Dilepton, and $R_P$-Violating Couplings

String constructions can allow perturbative leptoquark, diquark, or dilepton couplings between new quasi-chiral exotics and ordinary particles. Of course, the simultaneous presence of both leptoquark and diquark couplings would lead to rapid proton decay, but one or the other by itself would be possible. There are numerous implications for collider physics and for flavor changing or nonuniversal processes [262, 263], such as the hints of anomalies in $B$ decays (e.g., $\Gamma(B \to K\mu^+\mu^-)/\Gamma(B \to Ke^+e^-)$ [264], or $\Gamma(B \to D^{(*)}\tau\nu)/\Gamma(B \to D^{(*)}\ell\nu)$ [265]). Other collider implications include decay mechanisms for quasi-chiral exotics that are sometimes forbidden to significantly mix with ordinary particles [225].

The extended (beyond-MSSM) symmetries can either allow or forbid $R_P$-violating couplings amongst standard model particles. For example, $H_d$ and the standard model lepton doublets $L$ have the same standard model gauge assignments. If, however, they carry different extended quantum numbers then the couplings $H_uL$, $LLE^c$, and $LQD^c$ will be forbidden, at least at the perturbative level. The diquark coupling $U^cDD^c$ may be similarly forbidden. Conversely, in some vacua some of these couplings are allowed, implying $R_P$ and $L$ and/or $B$ violation.

#### 3.4.2 Family Nonuniversality

Family nonuniversality, typically because of different stringy origins of the fermion families, has already been mentioned in connection with $U(1)'$ couplings and with selection rules for Higgs couplings. Such effects can lead to flavor changing neutral currents or could be relevant to understanding fermion mass hierarchies. Mixing between ordinary and exotic fermions with different $SU(2) \times U(1)$ assignments will in general also lead to flavor changing effects involving the ordinary $Z$ [223].

#### 3.4.3 Mechanisms for Yukawa Hierarchies

String theories admit a number of possible origins of fermion masses and mixings (as well as other couplings, such as those leading to quasi-chiral exotic masses), and it is possible or even likely that the details of the fermion spectrum are to some extent accidental features of the location in the landscape, and not susceptible to any simple explanation.

A perhaps simpler issue is the origin of the observed large hierarchies of masses and mixings. Here, there are again various possibilities. Let us consider two very different examples.
One possibility is that the Yukawa and related couplings actually result from stringy higher-dimensional operators. For example, the various $u$-type quark Yukawas couplings

$$W_U = \Gamma^q Q_H U^c$$

(3.5)

could actually result from couplings like

$$\frac{1}{M^q_i} Q_H U^c S_1 S_2 \cdots S_n,$$

(3.6)

where $S_i, i = 1 \cdots n$, are standard model singlets (some of which could be the same) and $M_i$ is the string scale. Effective couplings of various magnitudes could arise if the $S_i$ obtain expectations values smaller than $M_i$. For example, the $t$, $c$, and $u$ Yukawas could be associated with operators with $n = 0$, 2, and 5, respectively, and $\langle S_i \rangle / M_i \sim 0.1$. This sort of thing often occurs in heterotic constructions (e.g., [59]), with the selection rules on the allowed operators controlled by additional gauge symmetries. The magnitudes of the $\langle S_i \rangle$ are obtained by minimizing the scalar potential in the presence of a non-zero Fayet-Iliopoulos term for an anomalous $U(1)'$, a process known as vacuum restabilization [162].

This mechanism is basically a stringy version of the Froggatt-Nielsen mechanism [266].

In Type IIA intersecting brane constructions, on the other hand, chiral fermions are trapped at the intersections of D6 branes. Consider, for simplicity, the case in which the extra dimensions form three 2-tori. Three-point couplings result when three stacks of branes form triangles of area $A_I$ in the $I$th torus, with the Yukawa couplings proportional to $\prod_I \exp(-A_I / 2\pi\alpha')$, where $\alpha'$ is the string tension [267, 268]. Thus, modest hierarchies of these areas can lead to exponential hierarchies in the Yukawa couplings. There may also be selection rules such that some of the couplings are forbidden at the perturbative level by the extra anomalous (or non-anomalous) $U(1)'$s, but in some cases very small couplings can emerge from nonperturbative string instanton effects.

Of course, there are also other possibilities, such as multiple Higgs bosons, the ordinary (field-theoretic) Froggatt-Nielsen mechanism, etc, that are not unique to string theory, and exponentially-small couplings could have other, non-stringy, origins [271].

We also remark that there are sometimes relations between Yukawa couplings that are often different from familiar GUT-type relations.

### 3.4.4 Nonstandard Neutrino Mass Mechanisms

It is nontrivial to incorporate conventional bottom-up type models for neutrino mass in simple perturbative string constructions. For example, even when there is an underlying grand unification symmetry (as in heterotic models), the GUT symmetry is often broken in the extra dimensions. Moreover, it is difficult or impossible to obtain the large representations, such as the 126 of

---

20By stringy operator we mean that the inverse masses are due to the underlying string dynamics rather than the exchange of a massive particle in the four-dimensional field theory.

21The restabilization also reduces the gauge symmetry and generates string-scale masses for some of the exotics.

22A roughly uniform distribution of such areas could lead to the approximately scale-invariant spectrum of fermion masses, consistent with observations [4, 269].

23For detailed studies see [270].

24The absence of the 126 of SO(10) is a theorem for free-field heterotic constructions [251].
SO(10) that are often employed in model building to generate large Majorana mass terms such as $W \sim S_\alpha N^c N^c$ [suppressing $SO(10)$ indices] for the right-handed (sterile) neutrinos $N^c$, or the non-supersymmetric equivalent. More likely are that both Majorana and Dirac mass terms that can lead to a seesaw model for the active neutrinos $N$ are generated by stringy higher-dimensional operators analogous to (3.6),

$$W = \lambda_D \frac{1}{M_s^6} L H_u N^c S_1 S_2 \cdots S_p + \lambda_S \frac{1}{M_s^q} N^c N^c S_1 S_2 \cdots S_q. \quad (3.7)$$

However, the additional gauge symmetries sometimes restrict such operators so that the resultant seesaw masses for the active neutrinos are too small to be relevant [272]. There are a few successful examples [75] involving rather large $p$ and $q$, but these involve of $O(100)$ right-handed neutrinos $N^c$ and are therefore quite different from the conventional seesaw.

Type IIA constructions are even more difficult, because they conserve lepton number at the perturbative level [62, 66]. However, $L$ can be violated by nonperturbative string instantons [165, 166], possibly allowing a Majorana mass for the $N^c$. A direct stringy Weinberg operator $\lambda_f N H_u N H_u / M_s$ [273] is another possibility [274], although there would have to be a low string scale for it to be sufficiently large, i.e., $M_s / \lambda_f \sim 10^{14}$ GeV. Alternatively, exponentially-small Dirac masses can be generated at the perturbative level if there are large intersection areas $A_I$ [40, 72], or nonperturbatively by string instantons [275].

There are other (not necessarily stringy) possibilities for small Dirac or Majorana masses (or even mixing between active and sterile neutrinos) from higher-dimension field-theoretic or stringy operators in the superpotential or Kähler potential [276, 277], non-holomorphic soft supersymmetry breaking terms [278], or large dimensions. The entire subject is reviewed in [279].

3.4.5 Perturbative Global Symmetries from Anomalous $U(1)'$

Exact global symmetries are believed to be incompatible with string theory or quantum gravity [15]. However, as has already been described in Section 3.2.1, constructions often involve anomalous $U(1)'$ factors. The associated $Z'$ gauge bosons typically acquire string-scale masses, but the symmetry survives as a global symmetry on the perturbative couplings. Such global $U(1)'$'s are only approximate, and can be broken by non-perturbative string instanton effects [165–167], which may, however, be very small. Possible implications include Yukawa hierarchies [270], suppressed Majorana neutrino masses [279], and the $\mu$ problem.

3.5 Additional Issues

3.5.1 Grand Unification and Gauge Unification

Most string compactifications do not yield canonical grand unified theories in the low energy four-dimensional theory. Although heterotic constructions involve an underlying grand unification that can lead to canonical grand unification via a special choice of gauge field background (as encoded in a holomorphic vector bundle), the symmetry may be broken in the compactification, leading to a direct compactification to the standard model, MSSM, or extended gauge group. Direct compactifications may have some advantages or disadvantages compared to traditional GUTs, such as Orbifold GUTs [280], involving extra dimensions, model some of the relevant physics.
as non-standard Yukawa relations (with obvious implications for the fermion spectrum) or the absence of the doublet-triplet problem.

Even in the case of a full grand unified gauge group in four dimensions there may not exist the adjoint or other Higgs representations needed to break the GUT symmetry. They are also unlikely to yield very large groups or large representations, such as the 126-dimensional Higgs representation often invoked in $SO(10)$ models of neutrino mass [251].

Weakly coupled type I and II brane constructions generally do not have an underlying grand unification\textsuperscript{26} except for special cases, such as a full $U(5)$ stack of branes in Type IIA (with, however, a perturbative $t$-quark Yukawa forbidden by the $U(1)$). $SO(10)$ GUTs cannot be achieved at weak coupling due to the absence of the 16-dimensional representation.

Both $SU(5)$ and $SO(10)$ GUTs can be achieved in F-theory, however, due to strongly coupled effects [79, 80]. The phenomenology of these models has been studied extensively; see [83] and references therein. Interestingly, they may only arise on subloci in complex structure moduli space [281], which in simple geometries occur at high codimension [282] but at lower codimension in more typical geometries [283]. On the other hand, $SU(3)$ and $SU(2)$ are the only $SU(N)$ groups that arise geometrically in F-theory at generic points in complex structure moduli space [281], suggesting direct compactification to the standard model rather than a GUT.

One consequence of the absence of canonical grand unification is that string constructions do not necessarily preserve the simple and successful gauge unification of the MSSM [284]. In heterotic constructions the usual $SU(5)$-type boundary condition\textsuperscript{27} at the unification scale $M_X$,

\[
\frac{1}{k_y} \frac{1}{\alpha_y} = \frac{1}{k_2} \frac{1}{\alpha_2} = \frac{1}{k_3} \frac{1}{\alpha_3}, \quad \text{with} \quad (k_y, k_2, k_3) = (\frac{5}{3}, 1, 1),
\]

may be modified by higher Kač-Moody levels [285]. For the most common case of canonical $SU(2)$ and $SU(3)$ embeddings, (3.8) is modified to allow an arbitrary positive $k_y$, depending on the hypercharge embedding. Of course, $k_y \neq 5/3$ can in principle be compensated by other effects, e.g., the absence of low-scale supersymmetry [286]. Another stringy effect is that $M_X$ is expected to be close to the string scale, modulo threshold effects (which can be large due to string excitations). Effects such as exotic particles, lack of supersymmetry, multiple thresholds, and low and (conventional) high-scale threshold corrections, which can also be present in non-string theories, are also possible.

Similarly, most brane construction do not predict the normalization in (3.8). For example, in the intersecting brane constructions the gauge normalization depends on the volumes of the three-cycles wrapped by the stack, and so there need be no direct relation between the different gauge factors (unless there is a full $U(5)$ stack). In some special cases [4, 287, 288], however, there is a relation

\[
\frac{1}{\alpha_y} = \frac{1}{\alpha_2} + \frac{2}{3} \frac{1}{\alpha_3},
\]

which is weaker than but compatible with (3.8).

\textsuperscript{26}See [4] for an interesting discussion of the emergence of certain GUT-type features in brane models.

\textsuperscript{27}$\alpha_Y \equiv g^2 / 4\pi$. 
3.5.2 Low String Scale

There have been many suggestions that the fundamental scale $M_F$ of nature could be much smaller than the Planck scale $M_P$, perhaps as small as $1 - 100$ TeV, therefore solving or alleviating the electroweak hierarchy problem. This could occur if the relation between $M_F$ and gravity is modified by the existence of additional space dimensions that are large (e.g., [289, 290]) compared to the simple dimensional estimate $L \sim o(M_F^{-1})$, or warped [291], although the former case introduces a new hierarchy problem of why $L$ is so large. In the string context, this corresponds to a string scale $M_s \ll M_P$ with some of the extra dimensions large [292]. Although such vacua may be rare in the string landscape, they would not only be relevant to the hierarchy problem, but could even yield observable signatures at the LHC or other colliders if $M_s$ were of $O$(TeV). These could include Z′s from anomalous $U(1)′$s [293], TeV-scale black holes [294], and various stringy excitations [295–298], including Kaluza-Klein states, winding modes, and Regge trajectories of higher-spin states. For a general review, see [300].

3.5.3 Environmental Selection

An enormous landscape of vacua would be a natural home for ideas of the multiverse and environmental selection [4, 46, 47, 88, 90, 91, 146, 301–315]. In particular, some aspects of our observed world could be determined by anthropic considerations. Anthropic reasoning is probably the most plausible explanation that has been advanced for a tiny but non-zero cosmological constant [110, 316], and could also resolve other fine-tunings such as the smallness of the electroweak scale and values of the proton, neutron, electron, and neutrino masses [317–319]. It could also explain discrete choices, such as the existence of precisely three large space dimensions (necessary for stable planetary orbits [320]). It should also be emphasized that our observed physics need not be anthropically unique: there could well be many other anthropically-suitable vacua [321, 322]. Furthermore, anthropic considerations would not necessarily have to explain everything: some aspects (e.g., the precise values for the heavy quark masses) could well be essentially random, while others (e.g., the smallness or vanishing of the strong CP phase) could have conventional dynamical or symmetry explanations.

The multiverse concept requires not only the existence of an enormous landscape, but also that many of the vacua are de Sitter and metastable. Some aspects related to metastability of the anti D3-brane de Sitter uplift are still debated (see, e.g., [108, 323, 324] and [325, 326] for arguments pro and con, respectively, and [327] for a discussion on flux compactifications giving rise to rolling solutions), but it is a technical issue that should eventually be resolved. Furthermore, there are other uplift mechanisms other than from anti D3-branes.

Another requirement is that there should be a method for sampling the different vacua. Some form of eternal inflation seems ideally suited. At present, inflationary ideas are supported by most cosmologists and observations [89, 328, 329], but see [330] for a contrary point of view. One important prediction of eternal inflation is that our Universe should almost certainly be either flat or have a small negative curvature [331–333].

Like string theory itself, the ideas of the multiverse and environmental selection are controversial and difficult to test experimentally. Our view, however, is that the combination of the existence...
of the string landscape and the possibility of eternal inflation, together with its solution to the cosmological constant problem and the fact that it is a quantum theory of gravity, are sufficiently attractive that one must consider the multiverse and environmental selection very seriously. It is important to pursue every direct or indirect experimental probe and theoretical exploration of these ideas.

On the other hand, even if environmental selection does play a role in accounting for certain aspects of observed physics, other aspects may not admit such an explanation. In these circumstances, mechanistic or remnant interpretations may play an important role.

### 3.5.4 Other Possible Remnants/Effects

We briefly mention a few other possibilities.

- The intrinsic nonlocality of string theory opens up the possibility of spontaneous Lorentz and \( CPT \) violation (e.g., [334]), with implications for the propagation, decays, and (ν) oscillations of high energy γ’s, e±’s, ν’s, and gravity waves [335], and for alternative versions of baryogenesis (e.g., [336]).

- The possible time or space variation of moduli could in principle lead to observable variations of the fundamental “constants” of nature (e.g., [337, 338]).

- Perturbative string theory gives a simple explanation [339] of the apparent absence of massless continuous spin representations of the Poincaré group [340].

### 4. Discussion

String theory is an extremely promising and attractive candidate for a consistent quantum-mechanical unification of gravity and the other interactions. Due to the enormous landscape of vacua and the likely high string scale, it will probably not be possible to directly confirm or falsify it in detail. However, in our view the advantages are so large as to motivate every type of exploration of possible evidence for or against the basic ideas. One possibility involves theoretical progress, perhaps finding a “smoking gun”, or finding a particular vacuum that corresponds in great detail to the observed standard model or MSSM. However, so far there is no hint of a smoking gun. Furthermore, given the enormity of the landscape it is unlikely, in the absence of a theoretical breakthrough, that the exact vacuum corresponding to our world will be found.

Another direction, explored in this article, involves the possibility that the standard models of particle physics or cosmology are not the final story, and that some types of additional physics are much more likely than others in the string landscape, especially if one restricts to vacua in which the low energy SM and cosmological degrees of freedom are essentially the ones that emerge directly at the string scale and not as composites. In particular, many semi-realistic constructions contain remnant physics, by which we mean new particles, interactions, or features that are present due to the details of the compactification. Certain types of remnants (e.g., axions, moduli, large non-abelian gauge sectors, \( Z' \)’s, quasi-chiral exotics, extended Higgs/Higgsino sectors), occur very frequently. While not unique to string theory, their observation would be at least suggestive. This would especially be the case if they are unrelated to any standard model problem and are not part
of a more complete low-energy theory. In such a scenario they would appear ad hoc to a bottom-up model builder, who would likely not add them to their theory. Conversely, features such as large representations are extremely rare in the string landscape, and their observation would constitute evidence against string theory. For any feature that is firmly determined to be in the swampland, its observation would constitute falsification of string theory.

We conclude with a parable. Many string vacua look nothing like our world, they are darts that missed the dartboard entirely. Other darts get somewhat closer and hit the board, while still others may be scattered around a bullseye created by theorists that represent the “simplest” or “most natural” models. But Nature is not beholden to the theorists’ bullseye. The true bullseye is determined experimentally, and until we do so we must take near misses seriously, particularly if they typically have one (or more) of a few simple features. Hence, remnants.

Acknowledgments. J.H. thanks C. Long for useful discussions. He also thanks the students at TASI 2017 for many interesting discussions, and Mirjam Cvetič, Tom DeGrand, and Igor Klebanov for organizing such an excellent school. J.H. is supported by NSF grant PHY-1620526.

A. String Theory Calculations and Formalism

The goal of the rest of this review has been to discuss broadly what is currently known about string remnants, citing the original literature and presenting the material in a way accessible to any theoretical high energy physicist.

In this appendix we present some additional details for the interested reader.

A.1 Dark Gauge Sectors in F-theory

In this section we wish to show that large gauge sectors in F-theory are completely generic, due to the prevalence of so-called Non-Higgsable clusters, and furthermore that this almost always arises at strong coupling. We will focus entirely on F-theory geometry, which sets the foundation on which all F-theory compactifications are built. See, e.g., [281] for a discussion of how fluxes may modify aspects of the physics that are determined geometrically.

A.1.1 F-theory Geometry

F-theory [92, 341, 342] is a generalization of Type IIB string theory that allows for strong coupling regions in the extra dimensions of space \( B \), as encoded in a holomorphically varying axiodilaton field

\[
\tau = C_0 + i e^{-\phi}.
\]

(A.1)

One central observation of [92] is that it is convenient to encode this field profile in the complex structure of an elliptic curve that is fibered over \( B \). The Calabi-Yau manifold \( X \) of the F-theory compactification is therefore an elliptic fibration

\[
E \hookrightarrow X \xrightarrow{\pi} B.
\]

(A.2)

So-called birational transformations do not change the complex structure of the smooth elliptic fibers, and therefore the axiodilaton is invariant under such transformations. By a theorem of
Consider a point \( p \) in the space and \( \pi \) the projection to \( \mathbb{C}^2 \). Above a generic point \( p \in B, \pi^{-1}(p) =: E_p \) is a smooth elliptic curve. On the discriminant locus \( \Delta = 4f^3 + 27g^2 = 0 \), points \( p \in \{ \Delta = 0 \} \) have \( \pi^{-1}(p) \) singular. This can be seen directly from the cubic \( v(x) = x^3 + fx + g \).

Consider a point \( p \) away from \( \Delta = 0 \), the cubic has three distinct roots and the fiber is a double cover of \( \mathbb{C} \) branched at the roots of \( v(x) \) and also a point at infinity. Such a double cover is a torus, where a basis of one-cycles may be defined by connecting the roots of the cubic and considering that interval to be the projection of a one-cycle in the double cover. Visualizing this, suppose that the roots are \( p_1 \) and \( p_2 \), and one travels along the bottom sheet from \( p_1 \) to \( p_2 \), and then travels back from \( p_2 \) to \( p_1 \) along the top sheet; this defines an \( S^1 \). Now consider any path from \( p \) to a point inside \( \Delta = 0 \). By definition \( v(x) \) has a degenerate root, signaling that a circle of the type just constructed has collapsed and the corresponding fiber is singular.

From the point of view of string theory, the collapsed fiber gives rise to a monodromy \( M \in \text{SL}(2, \mathbb{Z}) \) associated to the fiber, which means that taking a circle around \( \Delta = 0 \) induces a map on \( \tau \mapsto \frac{a\tau + b}{c\tau + d} \) for \( \tau \mapsto \frac{a\tau + b}{c\tau + d} \).
on the axiodilaton. Such a monodromy signals the presence of an object magnetically charged by the axiodilaton, that is, a seven-brane. In weakly coupled Type IIB compactifications these may only be D7-branes and O7-planes.

In F-theory there is a richer spectrum of seven-branes than in weakly coupled Type IIB. They are classified by Kodaira’s classification of codimension one singular fibers, which are determined by the order of vanishing (OOV) of \( f \), \( g \), and \( \Delta \), labeled respectively by \( l \), \( m \), and \( n \) in Table 1. In each case the OOV determines an ADE singularity in the elliptic fibration and, depending on the details of codimension two structure, the gauge algebra on the seven-brane. Unlike weakly coupled Type IIB, seven-branes in F-theory support the exceptional simple Lie gauge groups \( G_2 \), \( F_4 \), \( E_6 \), \( E_7 \), \( E_8 \). One can see that these are not weakly coupled by the study of the \( j \)-invariant of the elliptic curve. It is given by

\[
j = \frac{4f^3}{4f^3 + 27g^2},
\]

and in the cases of the exceptional algebras \( f \) and \( g \) both have non-zero orders of vanishing, forcing \( j \) to be either 0 or 1. The \( j \)-invariant is a function of \( \tau \), and in the case \( j = 0 \) \( (j = 1) \) may be inverted to give \( \tau = \exp(2\pi i/3) \) \( (\tau = i) \), up to an \( SL(2,\mathbb{Z}) \) transformation necessary to map \( \tau \) into the fundamental domain. These values of \( \tau \) have \( g_s = 2/\sqrt{3} \) and \( g_s = 1 \), respectively; the exceptional seven-branes are strongly coupled. For a recent in-depth discussion of strongly coupled effects, particularly in the presence of the elliptic fibers that may arise for non-Higgsable clusters, see \[254\].

A.1.2 Non-Higgsable Clusters

In recent years it has been discovered that for nearly all toric extra dimensional spaces \( B \), the most general form of \( f \) and \( g \) appears as

\[
f = F \prod_i x_i^{l_i}, \quad g = G \prod_i x_i^{m_i},
\]

where \( F \) and \( G \) are sections of appropriate bundles, \( x_i \) are homogeneous coordinates on the toric variety\[29\], and \( l_i \) and \( m_i \) are positive. Similar statements can be made for bases that aren’t toric (see, e.g., \[344\]), but it is simplest here to discuss the toric case. When \( f \) and \( g \) have this form, the discriminant may be written

\[
\Delta = \tilde{\Delta} \prod_i x_i^{n_i},
\]

where \( n_i = \min(3l_i, 2m_i) \). From Table 1 it is easy to see that there are non-trivial Kodaira fibers above each locus \( x_i = 0 \) appearing in the discriminant, and therefore these loci are also the locations of seven-branes.

Recall from your string theory course that the relative positions of D-branes determines the mass of the W-boson arising from the string stretched between them. Separating two stacks of D-branes therefore corresponds to a non-abelian Higgs mechanism. In the current picture, however,

\[29\] For an introduction to toric geometry, see, e.g., \[343\]. Here, it is critical to know that toric varieties are generalizations of weighted projective spaces that may have many scaling relations for the homogeneous coordinates, rather than one. Like weighted projective spaces, for any given scaling relation the different homogeneous coordinates may have different scaling weights.
there are non-trivial seven-branes on $x_i = 0$ for the most general values of the coefficients of $f$ and $g$, i.e., at generic points in the complex structure moduli space $\mathcal{M}_{cs}(X)$ of the elliptic fibration. Therefore the branes cannot be split at all, and there is no Higgs mechanism!

Such seven-branes are called geometrically non-Higgsable seven-branes (NH7) and a collection of intersecting NH7's is called a geometrically non-Higgsable cluster. The modifier "geometrically" arises is unnecessary in six-dimensional F-theory compactifications, where all known Higgs mechanisms are via complex structure deformation, but it is important in four-dimensional compactifications and two-dimensional compactifications since there the Higgs mechanism may arise in other ways, such as T-branes [345]. Henceforth we will drop the modifier and refer to these configurations as either non-Higgsable seven-branes or non-Higgsable clusters. Similarly, since T-branes and fluxes may break down the gauge group realized geometrically on the seven-brane, such gauge groups are often referred to as the geometric gauge group. Henceforth we will also drop this modifier, and it is to be understood that all mentioned gauge groups are the geometric gauge groups prior to any flux or T-brane breaking.

Non-Higgsable seven-branes are almost always strongly coupled. Since an NH7 arises along $x_i = 0$ when $l_i$ and $m_i$ are both positive, we see from the form of the $j$-invariant that neither all non-Higgsable seven-branes have

$$\tau = i \quad \text{or} \quad \tau = e^{2\pi i/3}, \quad (A.11)$$

in which case the string coupling is $O(1)$ on the seven-brane. The lone exception is $l_0^3$ in the case that $(l_i,m_i) = (2,3)$. Then the factors of $x_i$ drop out of the $j$-invariant and a weak coupling limit may be taken. Such configurations are seven-branes that become 4 D7-branes on top of an O7-plane in the weak coupling limit.

The gauge groups that may arise on non-Higgsable seven-branes are limited. From Table 1, we see that the possibilities are

$$G \in \{E_8, E_7, E_6, F_4, D_4, B_3, G_3, A_2, A_1\}, \quad (A.12)$$

This unusual sequence of Lie groups is known as Deligne’s exceptional series, and it also arises in 4d $\mathcal{N} = 2$ SCFTs [346] as the flavor symmetries of SCFTs whose Higgs branch is the one-instanton moduli space of the instantons of the flavor symmetry. It is interesting to note that $SU(3)$ and $SU(2)$ are the only $SU(n)$ groups that may exist for generic values of the moduli, which is potentially of phenomenological interest. For example the entire $SU(3) \times SU(2)$ spectrum of the standard model may be realized at a single intersection of such seven-branes [281].

There is a rich literature on non-Higgsable clusters. They are very well understood in six-dimensional compactifications, beginning with the seminal works [45, 341, 347] and also later studies [252, 347–352] across a number of topics. Non-Higgsable clusters conceptually change the task of moduli stabilization since gauge symmetry can exist at generic points in moduli space [281], i.e., without tuning to a non-trivial codimension in complex structure moduli space [52, 282, 283]. They also cannot geometrically give rise to GUTs, but may allow for a simple realization of the Standard Model in F-theory [281]. In four dimensions, non-Higgsable clusters are particularly interesting [344] due to realizing new structures that do not arise in higher dimensions, such as
loops [45]. NHCs appear in the geometry that is believed to give rise to the largest number of flux vacua [45], as well as in recently developed large ensembles of geometries [53, 252, 253]; see also [353]. In [124] it was shown that supervised machine learning can be used to make progress in understanding large ensembles of geometries with non-Higgsable clusters.

Non-Higgsable clusters can be complex. For example, [252] presented a Monte Carlo analysis of tens of millions of bases and associated non-Higgsable clusters, and found that a typical NHC might look like the figure displayed in section 3.2.2, which is the non-Higgsable cluster associated to a particular toric threefold base. Each node represents an NH7 with the noted gauge group, and an edge is drawn between any two NH7 that intersect. Note the presence of multiple disconnected gauge sectors, complexity within the graph itself including both loops and long chains, and five different simple Lie group factors that appear. It is the prevalence and complexity of such structures that motivate dark gauge sectors.

It will help with physical intuition to turn to some examples.

All of our examples will be in the case that \(B\) is a toric variety. Suppose that \(B\) is a toric \(n\)-fold, which has real dimension \(2n\). Then \(f\) and \(g\) may be computed from the polytopes

\[
\Delta_f = \{ m \in \mathbb{Z}^n \mid m \cdot v_i + 4 \geq 0 \ \forall \ i \} \quad \text{and} \quad \Delta_g = \{ m \in \mathbb{Z}^n \mid m \cdot v_i + 6 \geq 0 \ \forall \ i \},
\]

which give rise to collections of monomials via

\[
m_f \in \Delta_f \mapsto \prod_i x_i^{m_f \cdot v_i + 4} \quad \text{and} \quad m_g \in \Delta_g \mapsto \prod_i x_i^{m_g \cdot v_i + 6}.
\]

From this, the most general form of \(f\) and \(g\) may be computed, and the structure of associated non-Higgsable clusters may be studied.

Consider one of the first examples of a non-Higgsable cluster \([341]\). The base of the elliptic fibration is an algebraic surface, the Calabi-Yau is a threefold, and the F-theory compactification has six non-compact directions. Specifically, the base is \(B_2 = F_3\), the third Hirzebruch surface. This surface may be modeled with homogeneous coordinates \(x_1, x_2, x_3, x_4\) with two \(U(1)\) charges,

\[
\begin{array}{c|cc}
Q_1 & Q_2 \\
\hline
x_1 & 1 & 0 \\
x_2 & 1 & 0 \\
x_3 & 3 & 1 \\
x_4 & 0 & 1 \\
\end{array}
\]

subject to the additional condition that \(x_1 = x_2 = 0\) and \(x_3 = x_4 = 0\) are removed from the geometry; in technical terms, the Stanley-Reisner ideal of the associated toric variety is \(SRI = \langle x_1 x_2, x_3 x_4 \rangle\). The anticanonical class \(-K_{F_3}\) has degree \((5, 2)\) and therefore \(f\) and \(g\) are polynomials of degree \((20, 8)\) and \((30, 12)\), respectively.

The existence or non-existence of non-Higgsable clusters is determined by studying the most general form of \(f\) and \(g\). This can be done systematically by determining all monomials associated with points in \(\Delta_f\) and \(\Delta_g\) as described above, but we will instead be descriptive since it better illustrates the central point. Suppose there was a monomial \(m\) in \(f\) with a factor \(x_3^4\). Since every monomial in \(f\) must be of degree \((20, 8)\), we deduce that \(m\) must also have a factor \(x_3^8\). But \(Q_1(x_3^8) = \)
24, which overshoots the 20 required for \( f \). Since all of the charges are positive, this means that some of the remaining coordinates must be in the denominator, and therefore they are not global sections; i.e., this monomial \( m \) is not defined everywhere on \( \mathbb{P}^3 \), and therefore \( m \notin \Gamma(-4K_{\mathbb{P}^3}) \). Similar arguments show that there is no monomial in \( f \) of the form \( x_4^3 \). However, monomials with \( x_4^3 \) may appear since this would require \( x_3^4 \) and \( Q(x_3^4) = 18 < 20 \). This analysis can also be done for \( g \), with the result that

\[
f = x_4^2 F \quad g = x_4^2 G,
\]

where \( F \) and \( G \) are sections of \( \Gamma(-4K_{\mathbb{P}^3} - 2D_4) \) and \( \Gamma(-6K_{\mathbb{P}^3} - 2D_4) \). The discriminant is

\[
\Delta = x_4^4 (4F^3 x_3^2 + 27G^2) =: x_4^4 \tilde{\Delta}.
\]

from which we see that we have a seven-brane with a type IV Kodaira fiber on \( x_4 = 0 \). In some cases monodromy may reduce the gauge group of a type IV fiber from \( SU(3) \) to \( SU(2) \), but in this example it does not occur and \( G = SU(3) \).

This \( SU(3) \) seven-brane exists for generic values of the complex structure moduli, as we have seen directly in this computation, which is easily understood in field theory. First, the number of adjoint hypermultiplets is \( g(C) \), the genus of the curve \( C \) wrapped by the seven-brane. Here the \( C = \{ x_4 = 0 \} \) and on \( C \) we have \( x_3 \neq 0 \), which leaves \( (x_1, x_2) \) as the remaining coordinates on \( C \), subject to the constraint that \( x_1 = x_2 = 0 \) does not exist. This, together with the charges of \( (x_1, x_2) \), show that \( C \) is a \( \mathbb{P}^1 \) and therefore \( g(C) = 0 \); there are no adjoint hypermultiplets. Charged hypermultiplets may also arise at intersections of the seven-brane with other non-abelian seven-branes or with the overall residual seven-brane \( \tilde{\Delta} \). In this example there are no other non-abelian seven-branes and a short computation shows that \( \tilde{\Delta} = x_4 = 0 \) does not exist in the geometry. In summary, there are no charged hypermultiplets, the gauge theory on the seven-brane at \( x_4 = 0 \) is pure \( G = SU(3) \) super Yang-Mills, and the absence of matter implied that the gauge theory is non-Higgsable.

[341] also studied F-theory compactifications on \( \mathbb{F}_n \) for all \( n = 0, \ldots, 12 \). The charges are

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( x_3 )</th>
<th>( x_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Q_1 )</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>( Q_2 )</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

and \( f \) and \( g \) are of degree \( (4(n + 2), 8) \) and \( (6(n + 2), 12) \), respectively, with the same Stanley-Reisner ideal as before. There is a non-Higgsable seven-brane on \( x_4 = 0 \) \( \forall n \geq 3 \), satisfying

\[
\begin{array}{c|cccccccccccc}
\hline
n & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
G & SU(3) & SO(8) & F_4 & E_6 & E_7 & E_7 & E_8 & E_8 & E_8 & E_8 \\
\hline
\end{array}
\]

The case \( n = 7 \) is the most interesting. Here \( \tilde{\Delta} = x_4 = 0 \) does exist, signifying the presence of matter, but a detailed analysis shows the matter is a single half-hypermultiplet in the \( 56 \) of \( E_7 \). This representation is pseudoreal, and with only a single half-hypermultiplet there is no gauge invariant holomorphic combination of the matter fields, signaling the absence of D-flat directions.
and therefore also the presence of non-Higgsability. The rest of the cases are non-Higgsable due to being pure super Yang-Mills.

The origin of non-Higgsability in these examples turns out to cover all of the six-dimensional examples. In [45] the possible building blocks of 6d non-Higgsable clusters were classified, with the result that gauge sectors are either pure super Yang-Mills, or have so little matter that there is no D-flat direction. Thus, non-Higgsable clusters are completely understood from the point of view of gauge theory in 6d, as determined by an exact dictionary between the geometry and the gauge theory.

In 4d and 2d compactifications the dictionary and physical origin of non-Higgsability is not yet completely understood. One natural possibility is that the matter spectrum and low energy superpotential together lift any would-be flat directions, so that the theory is non-Higgsable. However, the authors know of no reason that such a perturbative Lagrangian explanation would give rise to Deligne’s exceptional series, and the story may be further complicated by features that arise in 4d and 2d compactifications that do not arise in 6d compactifications, such as D3-brane tadpole cancellation. Obtaining a full understanding of geometric non-Higgsability is therefore an important subject in ongoing research.

The geometries giving rise to 4d and 2d compactifications are relatively well understood, however, and for the remainder of our discussion we will study aspects of large ensembles and how they motivate dark gauge sectors.

A.1.3 The Genericity of Non-Higgsable Clusters and the Scarcity of Weak Coupling

We will now discuss an ensemble of F-theory bases $B$, where in each case the most general Weierstrass model over $B$ was studied, including the associated structure of non-Higgsable clusters. This ensemble is closely related to two other ensembles that we will mention briefly.

In all three ensembles, the bases $B$ are toric varieties and the ensemble is generated by topological transitions to other bases known as blowups from some initial bases. We must therefore review some properties of blow-ups. For the general case of blow-ups, see, e.g., [354].

Toric blow-ups in toric threefolds, on the other hand, may be visualized and encoded combinatorially in the structure of appropriate subdivisions of intervals or triangles, which correspond to curve and point blow-ups, respectively. In both types of blow-ups, the curve or point is related by a new divisor that is known as the exceptional divisor, and since the divisor is a toric divisor it may be encoded in the structure of a new ray. The associated topological transition changed $h^{11}$ by 1, and the change in $h^{31}$ may be approximated by compute $\Delta_f$ and $\Delta_g$ on both sides of the transition.

Let us concretely consider the case of a curve blow-up, given two three-dimensional cones of the associated toric variety with vertices $(v_0, v_1, v_2)$ and $(v_1, v_2, v_3)$. To perform the blowup, we add a new ray $v_e = v_1 + v_2$ and replace the original 2 cones with 4 cones generated by

$$
(v_0, v_1, v_e), (v_0, v_2, v_e), (v_1, v_e, v_3), (v_2, v_e, v_3).
$$

This procedure may be iterated by subdividing the new cones in the same way, yielding the diagram shown in Figure 1 after two additional blow-ups. Some of the structure in the figure, such as the line between $v_1$ and $v_2$, will be discussed when polytopes are introduced.
Point blow-ups proceed in a similar way. Consider a cone generated by \( (v_1, v_2, v_3) \). It may be blown up by adding the new ray \( v_e := v_1 + v_2 + v_3 \), where the old cone structure is replaced by

\[
(v_1, v_2, v_3) \rightarrow (v_2, v_3, v_e) \rightarrow (v_3, v_1, v_e). \tag{A.18}
\]

One can visualize this procedure via the diagram

![Diagram](image)

where the vertices have been projected into a common plane in order to make the subdivision structure clear. The vertices in Figure 1 could similarly be projected onto the edge between \( v_1 \) and \( v_2 \).

Having introduced toric blowups, we now turn to a discussion of a concrete ensemble. However, we strongly urge the reader to also study the ensembles presented in excellent work by Taylor and Wang [252, 253]. The ensemble presented here and the ensembles of Taylor and Wang have different strengths, but the same qualitative phenomena: non-Higgsable clusters are completely generic and typically arise in multiple disconnected components, giving rise to large gauge groups and complex dark sectors.

**An Algorithmic Ensemble of \( \frac{4}{3} \times 2.96 \times 10^{755} \) Geometries**

We now present the ensemble of [53]. Compared to [253], it has the disadvantage of being restrictive in the sense that a certain “height” bound is imposed that is sufficient to have a good F-theory base, but not necessary, and thus is likely significantly undercounts F-theory geometries.

On the other hand, though the height bound is restrictive, it is concrete enough that local patches consistent with the height bound can be classified, giving a concrete construction algorithm for an exponentially large number of geometries using tree-like structures. Under the assumption that an \( O(1) \) number of the bases in this “tree” ensemble support at least one consistent flux solution, which is rather conservative since a Bousso-Polchinski argument would imply there are many, this ensemble gives an exact lower bound on the number of F-theory geometries

\[
\# \text{F-theory Geometries} \geq \frac{4}{3} \times 2.96 \times 10^{755}. \tag{A.19}
\]

Furthermore, by utilizing the construction algorithm, [53] was able to demonstrate universality in the ensemble. The authors find this avenue promising because — as is the spirit of remnants — demonstrating universality is a natural avenue for making predictions from string theory, and being able to do so in an exponentially large ensemble is often difficult. In such cases it may be possible to apply various techniques from computer science, such as supervised machine learning or genetic algorithms [120, 123, 124, 355]. See section 2.2 for further discussion.
Let us now describe how the ensemble is constructed.

Consider a smooth weak Fano toric threefold, where weak Fano means that any holomorphic curve $C$ satisfies $-K \cdot C \geq 0$, where $-K$ is the anticanonical class. Such a toric threefold may be encoded in a fine regular star triangulation of a 3d reflexive polytope $\Delta^\circ$, via the associated fan formed from 3d and 2d cones corresponding to faces and edges on the facets of $\Delta^\circ$. This is an important distinction: the facets are codimension one faces of $\Delta^\circ$ that may have interior points that give rise to non-trivial triangulations. The triangles and lines in that triangulation are also referred to as faces and edges. Henceforth, “face” and “edge” means the face and edge of the triangulation, and we will refer to the codimension one faces of $\Delta^\circ$ as facets. For any facet, the number of faces and edges in any of its triangulations is independent of the triangulation (as long as all points are used) and is a simple function of the number of facet vertices and number of lattice points in the facet boundary.

**Figure 1**: A height-2 and two height-3 blowups of an edge generated by the vertices $\{v_1, v_2\}$.

Topological transitions to distinct threefolds arise via smooth toric blowups of curves or points, as discussed above. Consider a curve associated with a 2d cone generated by $v_1$ and $v_2$. Viewing the facet head on, there would be a corresponding interval

and the corresponding curve may be blown up. This subdivides the interval by adding a point on the edge and its corresponding “height”, giving a sequence of blowups

where the edge between $v_1$ and $v_2$ is the edge in a facet, and, for example, the point labeled by 2 is the ray $v_1 + v_2$ projected into the facet. The sequence of blowups pictured here will be referred
to colloquially as “trees,” and accordingly the dashed green lines correspond to new edges that arise from the subdivisions associated to the blowup. Then any additional ray takes the form \( v_e = av_1 + bv_2 \) and the combination

\[ h_i = a_i + b_i \]  

(A.20)
is the “height” of the associated ray. Since the sequence forms a “tree,” we refer to each ray as leaf, and in particular leaves that sit in the original facet are on the ground, and sometimes referred to as “roots.”

Similarly, point blowups can be visualized via the subdivision of one face in the triangulation of a facet. The first smooth toric blowup would have \( v_e = v_1 + v_2 + v_3 \) and the associated diagram is

![Diagram](image)

where again the heights are labeled. The collection of trees built on edges (faces) will be referred to as edge trees (face trees).

Figure 2: The largest facets for \( \Delta_1 \) and \( \Delta_2 \), respectively, each with an arbitrary triangulation.

A valid F-theory base must have no divisors with OOV \( (4, 6, 12) \) for \( (f, g, \Delta) \); see the appendix of [53] for a complete discussion of this and related issues. In this ensemble, restricting any leaf to have height \( h \leq 6 \) is sufficient to avoid the existence of \( (4, 6) \) divisors, and thus any base built by putting \( h \leq 6 \) edge trees on edges and \( h \leq 6 \) face trees on face is a valid F-theory base from the point of view of the \( (4, 6) \) condition. Building an ensemble this way then necessitates classifying all such face trees and edge trees. The number with \( h \leq N \) as a function of \( N \) is given in Table 2.

An ensemble \( S_{\Delta^e} \) may then be built from any 3d reflexive polytope \( \Delta^e \) as follows. Take a fine regular star triangulation \( \mathcal{T}(\Delta^e) \) of \( \Delta^e \), which defines an initial toric variety. For each edge or face, place an edge tree or face tree with \( h \leq 6 \) for each such tree. This defines a large sequence of
Table 2: The number of possible edge trees and face trees as a function of the maximal height \( h \).

<table>
<thead>
<tr>
<th>( N )</th>
<th># Edge Trees</th>
<th># Face Trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>4231</td>
</tr>
<tr>
<td>6</td>
<td>82</td>
<td>41,873,645</td>
</tr>
</tbody>
</table>

 blowups from the initial toric variety, and the set of all ways of doing this defines \( S_{\Delta^e} \). Using Table 2, the cardinality of the ensemble is

\[
|S_{\Delta^e}| = 82^{#E \text{ on } \mathcal{T}(\Delta^e)} \times (41,873,645)^{#F \text{ on } \mathcal{T}(\Delta^e)},
\]

where \( #E \) and \( #F \) are the number of total edges and faces on \( \mathcal{T}(\Delta^e) \).

By studying all 4319 reflexive polytopes, the cardinality of each ensemble \( S_{\Delta^e} \) can be computed. Two polytopes \( \Delta_1^e \) and \( \Delta_2^e \) dominate the ensemble; they are the convex hulls of the vertex sets

\[
\begin{align*}
S_1 &= \{(-1,-1,-1), (-1,-1,5), (-1,5,-1), (1,-1,-1) \} \\
S_2 &= \{(-1,-1,-1), (-1,-1,11), (-1,2,-1), (1,-1,-1) \}
\end{align*}
\]

and the cardinalities of the associated ensemble of geometries are

\[
|S_{\Delta_1^e}| = \frac{2.96}{3} \times 10^{755} \quad |S_{\Delta_2^e}| = 2.96 \times 10^{755}.
\]

The factor of 3 in the denominator arises from a \( \mathbb{Z}_3 \) symmetry of the polytope that extends to a \( \mathbb{Z}_3 \) action on the associated toric varieties in the ensemble. These give the exact lower bound

\[
\# \text{ F-theory Geometries} \geq \frac{4}{3} \times 2.96 \times 10^{755}
\]

that was mentioned above. The number may be enlarged in a number of ways, discussed in [53], and in particular this is a subset of the ensemble considered in [253].

A number of universal physical features are known about this ensemble. We refer the interested student to the original literature for details, and simply cite some of the results.

In [53] it was shown that the probability that a base \( B \) in the ensemble has a non-Higgsable cluster is

\[
P(\text{NHC}) \geq 1 - 1.07 \times 10^{-755},
\]

so non-Higgsable clusters are completely generic in the ensemble, and understanding why they are implies that they should be generic in the full set of F-theory geometries. With probability above 0.9999995, the geometric gauge group of the non-Higgsable clusters is

\[
G \geq E_8^{10} \times F_4^{18} \times U^8 \times F_4^{H_2} \times G_2^{H_6} \times A_1^{H_4},
\]
Figure 3: The largest facet $F_1$ of the 3d reflexive polytope $\Delta^0$, with an arbitrary triangulation. A type II fiber above the divisor corresponding to $v_0$ places strong constraints on the existence of the Sen limit.

in which case $\text{rk}(G) \geq 160$. $H_i$ is the number of height $i$ leaves above leaves on the ground that carry $E_8$, and the $\geq$ means that $G$ can enhance beyond this, but this is the minimal structure. Notably absent are $SU(3)$, $E_6$, $SU(5)$, and $SO(10)$, which are interesting from the perspective of particle physics. One solution is that flux may break down the geometric gauge group to these factors, or they may arise geometrically with some smaller probability. For example, by random sampling the ensemble it was observed that $E_6$ arises on a distinguished divisor with probability one in a few thousand.

In [124] a supervised machine learning algorithm known as logistic regression was used to study those random samples. With 98% accuracy, the trained algorithm was able to distinguish whether or not the vertex $v_{E_6}$ associated to a distinguished divisor had an $E_6$ gauge group. By studying how the machine made its decision, a conjecture was generated that was rigorously proven, so that it can be computed explicitly that

$$P(E_6 \text{ on } v_{E_6} \text{ in } T) \simeq 0.00059128,$$

in which case the number of bases with $E_6$ on that divisor in the ensemble is $5.83 \times 10^{751}$. Through conjecture generation, a numerical machine learning result was turned into a rigorous result about the physics of an exponentially large ensemble.

Finally, it is interesting to understand whether or not weakly coupled limits are possible in the ensemble. Since nearly all geometries have non-Higgsable clusters, and non-Higgsable clusters are intimately tied to strong string coupling [254], at generic points in their moduli space the
theory is strongly coupled in $g_s$. Technically, this means that $g_s$ varies over the spacetime, but there are regions in the extra dimensions where $g_s$ is $O(1)$. However, this does not preclude the possible existence of a weakly coupled Sen limit. A necessary condition for the existence of a Sen limit is that the non-Higgsable clusters must have singularities no worse than Kodaira $I_0^*$, and any singularities less severe than $I_0^*$ must enhance to $I_0^*$. Recall that $I_0^*$ may become 4 D7 on top of an O7 in the weak coupling limit.

In [114] it was show that this necessary condition is effectively never satisfied. Specifically,

$$P(\exists \text{ a Sen limit}) \leq 3.0 \times 10^{-391}. \quad (A.27)$$

The authors believe that this result is the first quantitative statement about the size of the weakly coupled Type IIB string landscape within its larger F-theory context.

This fact poses a sharp question for future research: do the qualitative remnants and lessons from weakly coupled Type IIB compactifications survive the F-theory limit? For some remnants, such as the axions, dark gauge sectors, and standard model exotics, the answer seems to be yes. For example, the existence of large dark gauge sectors is strongly motivated by the genericity of non-Higgsable clusters discussed in this section.

### A.1.4 Dark Gauge Sectors and Glueballs

The results of the previous sections strongly motivate hidden gauge sectors in string theory. The results were specific to the broad (relative to Type IIB) context of F-theory, but qualitatively match results in other corners of the landscape, which also regularly have large hidden gauge sectors. Given what is currently known, it is reasonable to expect that large hidden gauge sectors are extremely common in the landscape as a whole, with a wide variety of ultraviolet gauge couplings as determined by moduli stabilization, as well as a variety of matter spectra, which in turn sets whether the gauge sector is asymptotically free. Some of the phenomenological considerations related to hidden gauge sectors were discussed in the main text.

Here we focus on the possibility that dark matter comes in a dark confining gauge sector. Whether or not this arises in a particular construction depends on the beta function and matter spectrum, but with the plethora of gauge sectors arising in string theory it seems reasonable to expect that some of them confine. For simplicity, we will also assume that pure Yang-Mills theory arises in the infrared, so that it is Yang-Mills that confines. In the presence of matter fields, or alternatively if the SUSY breaking scale is below the confinement scale, the spectrum of composite dark particles can become significantly more complex; see, e.g., [356]. However, dark pure Yang-Mills will suffice to make some of qualitatively important points clear, and it is also well motivated.

Consider a scenario [204] in which a dark Yang-Mills sector with gauge group $G$ and confinement scale $\Lambda$ is reheated in the early universe to a temperature $T'_{rh} > \Lambda$. This dark sector interacts gravitationally with the standard model and may also have other interactions. It is a thermal bath of dark gluons and cools as the universe expands. At a temperature $T'_\Lambda \simeq \Lambda$ the gluons confine into glueballs. The effective theory of the glueballs may have $3 \rightarrow 2$ interactions, which change the dependence of the dark sector temperature $T'$ on the scale factor $a(t)$, relative to the scale-factor dependence of the temperature $T$ of a bath of non-interacting non-relativistic particles. The
The temperature ratio is
\[ \frac{T'}{T} \propto \frac{a}{\ln(a)}. \quad (A.28) \]
The dark sector stays warm because the glueballs “cannibalize” themselves, i.e., the $3 \to 2$ interactions increase the average kinetic energy of glueballs while depleting their number. When those interactions cease to be effective, the glueballs freeze out and leave a relic. The relic abundance of glueballs is approximately
\[ \Omega_{\text{gb}} h^2 \approx \frac{\Lambda}{3.6 \, \text{eV} \, \xi}, \quad (A.29) \]
where $\xi := s/s'$ is the ratio of comoving entropy densities between the visible and dark sector, which remains constant throughout the process. Its value is set by reheating, and if reheating is not preferentially into one sector then $\xi = 1$. The approximation here is reasonable because the dependence of the relic abundance on other parameters is rather weak; see [208] for an exact expression and a discussion. We see that in the democratic scenario $\xi = 1$ a confinement scale $\Lambda \approx O(10) \, \text{eV}$, is required to saturate the relic abundance.

Now consider the fact that string compactifications often have many such gauge sectors. Each has an ultraviolet gauge coupling $\alpha_{UV}$ that is determined by moduli stabilization, and accordingly a compactification with many non-abelian sectors that confine will have a wide variety of confinement scales. From (A.29) it is clear that as the number of confining dark sectors goes up, the likelihood that at least one will oversaturate the relic abundance increases, since the confinement scales will have some non-trivial distribution determined by moduli stabilization. However, it only takes one oversaturating confining sector to rule out the model, and therefore it seems that there is a dark glueball problem in the landscape.

One way out is for the dark matter to decay, but for sectors that only couple to the standard model via moduli, axions, and gravity the associated decay is usually too slow [208], though in non-thermal cosmologies the problem is partially alleviated [212]. In cases where all of the confining sectors have sufficiently strong interactions with the visible sector, for example via light mediators, the problem can also be alleviated; see [204–214] for a broad dark glueball literature that considers a number of different possible interactions with the standard model. Another way out of the problem is to make $\xi$ extremely large, which would constitute an extremely asymmetric reheating process that reheats the visible sector, but barely reheats the dark sectors. See [208] for necessary values of $\xi$ under certain circumstances, and [357, 358] for recent studies of asymmetric reheating that demonstrate both interesting possibilities and difficulties.

None of the proposed mechanisms for avoiding the problem are completely compelling in light of a large landscape with many vacua, and the typicality of many non-abelian gauge sectors. Further alleviating the problem is an interesting direction for future research, and it would be interesting to understand how solutions correlate with other remnants.

### A.2 Axions

Axions are ubiquitous in string compactifications, due to the many $p$-form gauge fields that may give rise to axions upon compactification. In this section we review the general formalism that gives rise to axions in 4$d$ compactifications, and will then comment on how they arise in specific
corners of the landscape. We will also discuss some aspects of axion cosmology, including mass scales and ultralight axion dark matter.

Consider a $D$-dimensional quantum field theory with $D > 4$ and a $p$-form gauge field with field strength of the form

$$\int d^D x \, dC_p \wedge \ast dC_p.$$  

(A.30)

$C_p$ transforms under the gauge symmetry as

$$C_p \mapsto C_p + d\Lambda_{p-1}$$  

(A.31)

$(p-1)$-form $\Lambda$. $C_p$ may be present in other interactions such as Chern-Simons terms, but what matters here is that it is a dynamical higher dimensional field with a gauge symmetry.

Now consider compactification of this theory to four dimensions by the ansatz

$$X_D = X_4 \times M,$$  

(A.32)

where $X_4$ is a non-compact four-manifold and is a smooth $(D-4)$-manifold. $C_p$ may have terms with $p$ legs along $M$, in which case we expand in a basis of non-trivial $p$-forms as

$$C_p = \sum_i a_i(x) \omega_i \quad \omega_i \in H^p(X,\mathbb{Z}),$$  

(A.33)

where $x$ is the coordinate on $X_4$. The four-dimensional fields $a_i(x)$ are axions, and

$$\# \text{ axions} = b_p(M),$$  

(A.34)

where $b_p(M)$ is the $p^{th}$ Betti number of $M$. The axions will appear in the four-dimensional effective according to the details of the $D$-dimensional theory, which often includes the interactions of a higher dimensional supergravity theory in string contexts.

The basic setup occurs essentially everywhere in the landscape. For example, in different corners the field $C_p$ may be considered to be

- **M-theory.** $C_3$, the three-form of 11$d$ SUGRA.
- **F-theory and IIB.** $B_2$, the NSNS 2-form. $C_0, C_2, C_4, B_2$. There are the RR 01-form, 2-form, and 4-form, and the NSNS 2-form.
- **IIA** $C_1, C_3, C_5, B_2$. There are the RR 11-form, 3-form, and 5-form, and the NSNS 2-form.
  
  Note that in case where $M$ is a Calabi-Yau threefold with holonomy $SU(3)$ (rather than a subgroup) there are no axions from $C_1$ or $C_5$ since $b_1(M) = b_5(M) = 0$.

- **Heterotic.** $B_2$, the NSNS 2-form.

The authors are not aware of a well-motivated corner of the landscape that does not give rise to axions.

All concrete knowledge of string geometries to date imply that many axions should be expected. Let us discuss three ensembles to demonstrate the point. In [359], Kreuzer and Skarke classified four-dimensional reflexive polytopes. A toric variety may be associated in a canonical
way to each fine regular star triangulation of this polytope, and there is a family of smooth Calabi-Yau threefolds that arise as hypersurfaces in this toric variety. There are 30,108 distinct Hodge pairs \((h^{11}, h^{21})\) for these threefolds, and typically both Hodge numbers are in the tens or hundreds. In recent years there has also been significant progress in constructing \(G_2\) manifolds. These are seven-manifolds with holonomy \(G_2\), which are the relevant manifolds for 4\(d\) \(\mathcal{N} = 1\) compactifications of 11\(d\) SUGRA or M-theory. Both \(G_2\) manifolds and Calabi-Yau threefolds have metrics with specific holonomy groups, but unlike Calabi-Yau threefolds there is no simple topological criterion that can be checked to guarantee the existence of an appropriate metric; there is no analog of Yau’s theorem. However, in a construction known as the twisted connected sum construction of \(G_2\) manifolds (due to Kovalev \[360\], Corti Haskins Nordstrom Pacini \[361\] and others), there are now at least 50 million compact \(G_2\) manifolds. In \[361\] it was shown that \(b_3\) is often of \(O(30)\), but it was also argued convincingly that this ensemble is only scratching the surface. In the simplest generalizations which still used so-called “orthogonal gluing”, all such \(G_2\) manifolds will have \(b_3 \geq 23\), and the number could be much higher. See \[50, 362–366\] for physics works on twisted connected sum \(G_2\) manifolds. Finally, typical geometries in the previously mentioned F-theory ensembles \[53, 252, 253\] have \(h^{11} \simeq O(1000)\), which gives rise to that number of axions via reduction of the RR 4-form \(C_4\). In summary, this evidence points to the existence of at least \(O(10)\) axions, and likely \(O(10)\) or \(O(1000)\). Henceforth we will refer to “hundreds” of axions for brevity. These many-axion scenarios in string theory are sometimes referred to as an axiverse \[137\].

Each axion enjoys a perturbative shift symmetry, the origin of which is the gauge symmetry of \(C_p\). However, the shift symmetry may be broken by Euclidean instantons that couple to \(C_p\) in the extra dimensions, an therefore generate non-perturbative corrections to the 4\(d\) effective action that involve the instantons. For example, non-perturbative superpotentials of the form

\[
W_{np} = A(\phi)e^{-T} + \ldots
\]

may arise, where \(a = Im(t)\) and \(A(\phi)\) is a holomorphic function of other superfields in the theory.

Since the Euclidean instanton couples to \(C_p\), it is a \((p-1)\)-brane instanton, which in the respective corners may be

- **M-theory.** \(M2\)-brane instantons.

- **F-theory and IIB.** Worldsheet instantons, \(ED(-1)\)-instantons, \(ED1\)-instantons, and \(ED3\)-instantons.

- **IIA** \(ED0\)-instantons, \(ED2\)-instantons, \(ED4\)-instantons, and worldsheet instantons.

- **Heterotic.** String worldsheet instantons.

Each theory with a \(B_2\) field also has a dual \(B_6\) field that charges NS5-branes, which may give rise to NS5-brane instanton corrections.

**Masses, Axion Cosmology, and Dark Matter**

Axions may have a very wide range of masses in string theory. In practice it is very hard to actually compute the masses, since it requires all of the details of moduli stabilization. However,
in the two most-studied moduli stabilization scenarios of KKLT [108] and LVS it has been done. A variety of masses are possible; see, e.g., [139] and references therein.

We emphasize the importance and difficulty of the problem. Computing axion masses in KKLT and LVS, as well as in more general \( \mathcal{N} = 1 \) compactifications, requires a number of essential steps.

- Find all (or all leading, if an approximation can be justified) instanton corrections. For simplicity, consider instanton corrections to the superpotential,

\[
W_{NP} = \sum_i \mathcal{O}_i e^{-Q_i/T_i},
\]

where the details of the Q-matrix are determined by the cycle class, and systematically determining the cycle class of contributing instantons is computationally complex in analogy to the negative solution to Hilbert’s tenth problem [118].

- Compute the full scalar potential, including instantons. In supersymmetric theories this includes the F-term scalar potential

\[
V_F = e^K \left( K^{IJ} D_I W D_J \bar{W} - 3 |W|^2 \right)
\]

which itself becomes a lengthy polynomial in the hundreds of axion fields. Thus, moving from the superpotential to the scalar potential is labor intensive for typical numbers of axions and instantons.

- Minimize the potential and search for metastable de Sitter vacua. This is a very difficult problem for polynomials of hundreds of variables.

For large values of Hodge numbers, which is where most of the landscape is expected to be located via, e.g., arguments of Bousso-Polchinski [110], it is difficult to systematically compute instanton corrections, and therefore also the scalar potential and its minima.

Systematic computation of axion masses, decay constants, and couplings are critical to furthering our understanding of the phenomenology and cosmological implications of axions in the string landscape. However, recent state-of-the-art work [367] has made progress on efficiently studying certain instanton zero modes for Kreuzer-Skarke Calabi-Yau threefolds at large \( h^{11} \), including the example with largest \( h^{11} \), which has \( h^{11} = 491 \). These instanton zero modes are important for determining whether or not there is an instanton correction to the superpotential, and with additional progress on other zero modes it may be possible in the next few years to systematically compute instanton corrections at large \( h^{11} \) in Type IIB compactifications. This would not solve the problem of actually finding the minima, but it is a necessary first step and a promising direction.

Since it is difficult to do the concrete string calculations given current mathematical technology, it is worthwhile to consider effective field theory arguments for axion masses, where parameters in the effective theory are motivated by string considerations. We proceed in three steps. First, determine all symmetries of the EFT. A minimal assumption for a realistic EFT with an axion is that the symmetries are \( G_{SM} = SU(3) \times SU(2) \times U(1) \) and the shift symmetry of the axion. Next, determine operators that may give rise to axion masses consistent with the symmetries of the EFT. For the above minimal possibility, we have

\[
V_a = A \left( \frac{h^T h}{2 \pi^2 a} \right)^n \cos \left( \frac{a}{F} \right)
\]
in which case the axion mass

$$m_a = \sqrt{A} \left( \frac{\Lambda_{EW}}{\Lambda} \right)^{n-1} \left( \frac{\Lambda}{F} \right) \Lambda_{EW},$$  

(A.39)

where $A$ is a Wilson coefficient, $h$ is the Standard Model Higgs field, $F$ is the axion decay constant, $\Lambda$ is the ultraviolet cutoff, and $n$ is a non-negative integer whose value is determined by the structure of instanton zero modes. The axion mass is a function of the Wilson coefficient, the hierarchy between the electroweak scale and the UV cutoff, the ratio of the cutoff to the axion decay constant, and the electroweak scale itself. In the $n = 0$ case that does not involve Higgs fields, the electroweak scale drops out, as it should.

In string compactifications there is a distribution of axion decay constants, see, e.g., [138] and often the UV cutoff and mean of the decay constant distribution are an order of magnitude or two below the Planck scale. (More specifically, in certain ensembles of well-understood geometries, while the largest decay constant may be near the Planck scale, the mean may be much lower). We therefore take

$$\Lambda \simeq F \simeq B M_p,$$  

(A.40)

where $B \sim 1/100 - 1/10$. Then the axion mass is

$$m_a \simeq \sqrt{A} \Lambda_{EW} \left( \frac{\Lambda_{EW}}{M_{pl}} \right)^{n-1} B^{1-n},$$  

(A.41)

and depends on the Wilson coefficient, the parameter $B$, and electroweak scale, and the electroweak hierarchy. Under standard effective field theory arguments the Wilson coefficient $A$ is expected to be $O(1)$, and therefore the axion mass scale is set by the electroweak scale and hierarchy, up to the powers of $B$ that play a minor role relative to the hierarchy. Estimating the scales by taking $A = 1$, $\Lambda_{EW} = 10^2$ GeV, $\Lambda_{EW}/(BM_{pl}) = 10^{-16}$, we obtain axion mass scales

$$n = 0 : \quad m_a \simeq 10^{27} \text{ eV} \simeq M_{pl}$$

$$n = 1 : \quad m_a \simeq 10^{11} \text{ eV} \simeq \Lambda_{EW}$$

$$n = 2 : \quad m_a \simeq 10^{-5} \text{ eV} \simeq m_\nu$$

$$n = 3 : \quad m_a \simeq 10^{-21} \text{ eV} \simeq m_{pl}$$

$$n = 4 : \quad m_a \simeq 10^{-37} \text{ eV}.$$  

(A.42)

We see the axion mass generated by this operator is Planck scale, electroweak scale, neutrino scale, so-called “ultralight” scale, and $10^{-37}$ eV in the case $n = 0, 1, 2, 3, 4$, respectively.

This analysis should concern the astute reader: in the typical case $n = 0$ that does not utilize the Higgs field in the axion mass operator, why is the axion at Planck scale, and how does this square with the notion that axions should be light? The caveat lies in the assumption $A = 1$. In gauge theoretic or string theoretic setups, $A$ is exponentially suppressed since the operator is generated by an instanton. $A$ is an instanton suppression factor, and therefore the mass scale conclusions (A.42) are correct up to this critical instanton suppression factor. In string compactifications, the instanton suppression factor is usually determined by the volume of the cycle wrapped by the instanton in the extra dimensions, which itself is set by the vacuum expectation value of four-dimensional scalar
fields such as Kähler moduli in the case of Calabi-Yau threefolds. In a metastable vacuum, however, these moduli are stabilized and the volumes and instanton suppression factors are determined; see, e.g., [368].

To determine axion mass spectra concretely it is therefore critical to know distributions of instanton suppression factors in metastable vacua. Most metastable vacua are expected to occur in cases where the number of moduli controlling extra dimensional volumes is large, such as Kähler moduli in the case of Calabi-Yau manifolds and $G_2$ moduli in the case $G_2$ manifolds. This is an important direction for future research, and until such distributions are determined one can only do analyses at low numbers of moduli or using EFT arguments with the critical caveat that we have discussed.

Axions may be cosmologically relevant. For example, depending on details of the model they may be either the inflaton or the dark matter. We will focus on the possibility of axion dark matter, and refer the reader to [154] and references therein for a detailed treatment of axion inflation and axion monodromy inflation in string compactifications.

Consider evolution of an axion $\phi$ in a Friedmann-Robertson-Walker (FRW) universe with scale factor $a(t)$. The equation of motion for the axion is

$$\ddot{\phi} + 3H \dot{\phi} + m_\phi^2 \phi = 0,$$

where $H = \dot{a}/a$ is the Hubble parameter and $m_\phi$ is the axion mass. If $H \gg m_\phi$ then Hubble friction freezes $\phi$ and $w_\phi = -1$, where $w_\phi = p_\phi/\rho_\phi$ with $p_\phi$ and $\rho_\phi$ the axion pressure and energy density. On the other hand, if $H \ll m_\phi$ then $\phi$ oscillates about its minimum and behaves like matter (or “dust”), with an equation of state $w_\phi = 0$.

For the axion to be the dark matter in our universe, as opposed to some other matter-like relic, it must have a number of non-trivial properties. We will focus on two: behaving like matter prior to matter-radiation equality, and saturating the dark matter relic abundance.

Recall that big bang nucleosynthesis (BBN) makes successful predictions for light element relic abundances. For this to occur, the universe must be radiation dominated at BBN time. On the other hand, at later times dark matter comes to dominate the energy density of the universe, as is critical for the formation of large scale structure. It occurs for a simple reason. Recall that the energy density of a perfect fluid in an FRW universe scales with $a(t)$ as

$$\rho \propto \frac{1}{a^{3(1+w)}}.$$

Since $w_r = 1/3$ and $w_m = 0$ for radiation and non-relativistic matter, respectively, $\rho_r \propto 1/a^4$ while $\rho_m \propto 1/a^3$. Therefore as the universe expands via growth of $a(t)$, at some point

$$\frac{\rho_m(t)}{\rho_m(t) + \rho_r(t)}$$

trends to 1 and energy density is almost entirely in matter. This occurs at the time of MRE, and a necessary condition for this to occur is that the dark matter is actually behaving like matter at or
prior to MRE time. Since at MRE time $H(a_{\text{MRE}}) \approx 10^{-28}\text{eV}$, we have

$$m_\phi \geq 10^{-28}\text{eV}.$$  \hfill (A.46)

An axion lighter than this scale could exist and potentially cause cosmological problems, but it cannot be the dark matter that drives structure formation in our universe.

If the axion is all of the dark matter, or a non-trivial component of it, it must saturate or slightly undersaturate the observed dark matter relic abundance \cite{369}

$$\Omega_{\text{cdm}}h^2 = 0.12,$$  \hfill (A.47)

where the Hubble constant today is $H_0 = h \times 100\text{km/s/Mpc}$ where $h \simeq 0.7$. For an axion relic produced by the misalignment mechanism with $O(1)$ misalignment angle, its abundance is

$$\Omega_{\text{ax}}h^2 = 0.7^2 \times 1 \times \left(\frac{F}{10^{17}\text{GeV}}\right)^2 \left(\frac{m_\phi}{10^{-22}\text{eV}}\right)^{1/2}.$$  \hfill (A.48)

The experimental constraint is

$$\Omega_{\text{ax}}h^2 \leq \Omega_{\text{cdm}}h^2 = 0.12$$  \hfill (A.49)

For the decay constants a few orders of magnitude below the Planck scale, as is the case for many string theoretic axions, we see that the relevant axion mass scale is $m_\phi \simeq 10^{-22}\text{eV}$. This mass scale is the relevant mass scale for so-called fuzzy dark matter, which is a form of ultralight dark matter that may account for a number of potential astrophysical anomalies on galactic scales. In \cite{370} it was shown that string axions may give rise this scenario due to large instanton suppression factors; we refer the reader to \cite{370} for a detailed account of the scenario and related astrophysical issues.

One of the main results is that for decay constants typical in string theory, axions of the mass scale relevant for fuzzy dark matter saturate the relic abundance. While perhaps not as strong of a coincidence as the so-called WIMP miracle, it is certainly interesting. In \cite{370} the mass scale was accommodated by a very large instanton suppression factor. In \cite{371} concrete SUSY models were presented and it was shown that stringy instantons that couple the axion to the Higgs can trade very large instanton suppressions for smaller instanton suppressions and powers of the electroweak hierarchy; this is the EFT analysis presented briefly above.

A.3 Moduli

In four-dimensional compactifications of string theory that preserve $\mathcal{N} = 1$ supersymmetry all scalars must be complex, and therefore pseudoscalar axions must be accompanied by scalar moduli. Varying the vacuum expectation value of these moduli fields corresponds to varying the shape or size of certain cycles in the extra dimensions. Moduli and axions together correspond to flat directions in the leading order scalar potential, and therefore they must be stabilized by subleading effects such as fluxes or instantons in order to avoid fifth force constraints. Like axions, the number of moduli fields is related to topological numbers of the compactification manifolds $M$. Their existence is a generic prediction of string compactifications.

If excited in the early universe, moduli can have important cosmological consequences. Here we briefly review moduli in the case of Calabi-Yau threefolds and the cosmological consequences
of these and other moduli. We refer the reader to [134, 135] for further details related to the discussion in this section.

Let $M$ be a compact Calabi-Yau threefold. These are six-manifolds with a Ricci-flat Kähler metric. In the generic case they have holonomy $SU(3)$, though in some cases (for example, $M = K3 \times T^2$) the holonomy is $SU(2)$, and also a covariantly constant spinor. Actually finding such a metric on the compact space is notoriously difficult, and currently there are no known non-trivial examples. Fortunately, Yau’s theorem guarantees the existence of such a metric, even if it cannot be computed. The theorem is that such a metric exists if the first Chern class of the tangent bundle $TM$ is trivial, i.e.,

$$c_1(TM) = 0.$$  \hfill (A.50)

This topological condition is straightforward to check in many examples, for example if $M$ is a hypersurface or complete intersection in a higher-dimensional toric variety. Compactifications of heterotic and Type II string theory on $M$ have $\mathcal{N} = 1$ and $\mathcal{N} = 2$ supersymmetry, respectively, and in Type II compactifications the supersymmetry may be broken to $\mathcal{N} = 1$ by the introduction of $1/2$-BPS branes and orientifolds. The choice of $M$ gives a vacuum in four dimensions [99].

From the point of view of the $10d$ theory, Calabi-Yau moduli correspond to infinitesimal deformations of the Ricci-flat Kähler metric that lead to another Ricci-flat Kähler metric, and therefore to another Calabi-Yau manifold and associated vacuum state. Under Kaluza-Klein reduction, metric deformations become $4d$ scalar fields that are the moduli fields. There are two essential types of deformations: complex structure deformations and Kähler deformations. These give rise to complex structure and Kähler moduli whose number are

$$\# \text{ Complex Structure Moduli} = h^{21}(M)$$

$$\# \text{ Kähler Moduli} = h^{11}(M),$$  \hfill (A.51)

where $h^{21}(M)$ and $h^{11}(M)$ are the Hodge numbers of the Calabi-Yau threefold; see, e.g., [354]. These typically number in the dozens or hundreds, as discussed. The largest known database of Calabi-Yau threefolds are those corresponding to hypersurfaces in Gorenstein Fano toric varieties. The number is not known, but the Gorenstein Fano toric varieties can be associate to fine regular star triangulations of $4d$ reflexive polytopes. The latter have been classified by Kreuzer-Skarke [359], and there are 473,800,776 of them.

From the point of view of $4d$ cosmology, the critical difference between the modulus and axion field that belong to the same complex scalar is that the axion enjoys a shift symmetry while the modulus does not. The axion therefore receives its mass from non-perturbative effects (instantons) that break the perturbative shift symmetry, whereas many effects may contribute to the modulus mass. For this reason, moduli are usually heavier than axions, and in many models the lightest modulus $\phi$ satisfies

$$m_\phi \simeq m_{3/2},$$  \hfill (A.52)
where $m_{3/2}$ is the gravitino mass. This relationship is typical of models with gravity mediated supersymmetry breaking. In some models, which are often referred to as sequestered models \[372\], the relationship may be broken, but sequestering is difficult to achieve in string compactifications \[373\].

A central aspect of cosmologies with moduli is the so-called cosmological moduli problem \[126–128\]. Similar to the case of an oscillating axion, when the modulus oscillates about its minimum it behaves like cold matter and will therefore come to dominate the energy density of the universe, due again to the $\rho_0 \propto 1/a^3$ scaling. If the eventual decay of the modulus does not reheat the universe to a temperature above the BBN temperature $T_{BBN} \simeq 5\,\text{MeV}$ then the successful predictions of nucleosynthesis are spoiled. This problem is known as the cosmological moduli problem. One solution is to eliminate moduli from the theory, but this is not well-motivated from the perspective of string theory.

The simplest solution to the cosmological moduli problem in a string-motivated scenario is for the modulus to be massive enough that the reheat temperature associated with its decay is above the BBN temperature. Since moduli couple to the visible sector via non-renormalizable operators we will study decays that use a dimension-5 operator. The rate is

$$\Gamma_\phi = c_\phi^2 \frac{m_\phi^3}{M_{pl}^2}, \quad (A.53)$$

where $c_\phi$ is the Wilson coefficient of the effective operator. If the $\phi$ decay and subsequent thermalization are instantaneous then the reheat temperature is

$$T_{rh} = \left[ \frac{8}{90} \frac{\pi^3}{g_*} \frac{1}{M_{pl} \Gamma_\phi} \right]^{1/2} \simeq 10 c_\phi \left( \frac{m_\phi}{100\,\text{TeV}} \right)^{3/2} \text{MeV}. \quad (A.54)$$

The nucleosynthesis bound $T_{rh} \geq T_{BBN} \simeq 5\,\text{MeV}$ gives

$$m_\phi \gtrsim 100 \left( \frac{1}{2c_\phi} \right)^{2/3} \text{TeV} = 63\,\text{TeV}, \quad (A.55)$$

where the last equality is under the assumption $c_\phi = 1$. The bound can be moved slightly due to a factor of 2 uncertainties in $T_{BBN}$ or changes in the Wilson coefficient or assumption of instantaneous thermalization. However, the qualitative conclusion that

$$m_\phi \gtrsim O(50)\,\text{TeV} \quad (A.56)$$

does not depend strongly on such modifications.

Even if the cosmological moduli problem is avoided by sufficiently heavy modulus, there are at least two significant ways in which particle physics and cosmology are affected by the modulus.

Consider a theory where SUSY solves the electroweak hierarchy problem that also has a modulus satisfying the bound \[(A.56)\]. If the model gives rise to the relation

$$m_\phi \simeq m_{3/2} \quad (A.57)$$

then the bound implies that the supersymmetry breaking scale is an order of magnitude above the TeV scale. Such scenarios have been extensively studied and typically exhibit heavy scalar partners of the standard model fermions \[5, 192, 374, 375\].
The cosmological implications of the heavy modulus do not require weak-scale SUSY, but are compatible with it. In a standard cosmological history the universe is radiation dominated during nucleosynthesis and the first phase of matter domination occurs at matter-radiation equality, as discussed. However, it is also possible that there is a phase of matter domination prior to nucleosynthesis, in which case other aspects of cosmology may also be affected. For example, WIMP dark matter produced by thermal freezeout has a freezeout temperature
\[ T_{fo} \approx \frac{m_\chi}{20}, \tag{A.58} \]
where \( \chi \) is the WIMP. If the relation
\[ T_{BBN} \leq T_{rh} \leq T_{fo} \tag{A.59} \]
is satisfied then the decaying modulus avoids the cosmological moduli problem, but reheats the universe to a temperature below the scale necessary to produce \( \chi \) via by thermal freeze-out. Instead, the dark matter \( \chi \) may be produced directly by modulus decay, in which case the dark matter phenomenology changes, in particular predictions for indirect detection. If \( m_\chi \) is electroweak scale, as is typical for WIMPs, the window (A.59) is three orders of magnitude.

A.4 Exotics and Consistency Conditions

In this section we discuss certain string consistency conditions, their relation to consistency conditions in field theory, and their implications for particle spectra. We will begin with a concrete study of consistency conditions and will then proceed to understand them multiple points of view. We will then discuss implications of the conditions for embeddings of the minimal supersymmetric standard model (MSSM) into string theory. We will discuss multiple duality frames, but since the derivation relevant to us is straightforward in Type IIA compactifications, we begin there.

String Consistency Conditions

Consider 4d compactifications of the Type IIA superstring on a Calabi-Yau threefold \( M \) with O6-planes and D6-branes. The orientifold requires an antiholomorphic involution
\[ \sigma : M \rightarrow M \tag{A.60} \]
such that the fixed locus of \( \sigma \) is a three-cycle
\[ \sigma_{\text{fixed}} = \pi_{O6} = \pi_{O6} \in H_3(M, \mathbb{Z}). \tag{A.61} \]
The spacetime-filling O6-plane on \( \pi_{O6} \) is magnetically (electrically) charged under the the Ramond-Ramond one-form (seven-form) \( C_1 \) (\( C_7 \)). Since \( M \) is compact, Gauss’ law requires that there are other charged objects on \( M \) so that the net \( C_7 \)-charge is zero. D6-branes may be added to the compactification for this purpose, and the Gauss law condition — also known as D6-brane tadpole cancellation — is
\[ \sum_a N_a \left( \pi_a + \pi'_a \right) = 4\pi_{O6}, \tag{A.62} \]
where there are \( N_a \) D6-branes on the three-cycle \( \pi_a \) and also the orientifold image cycle \( \pi'_a \). Finally, if each \( \pi_a \) is a special Lagrangian submanifold and an additional condition is satisfied, presence of
the D6-branes and O6-plane breaks the $\mathcal{N} = 2$ supersymmetry of the Calabi-Yau background to $\mathcal{N} = 1$.

For general $\pi_a$ the stack of $N_a$ D6-branes carries a $U(N_a)$ gauge group, though it may carry a symplectic or special orthogonal group if $\pi_a$ is orientifold invariant. We consider the case of unitary groups, and therefore

$$G = \prod_a U(N_a).$$

The trace $U(1)_a \subset U(N_a)$ is often anomalous, in which case it receives a mass via the Stückelberg mechanism. However, under certain conditions a linear combination

$$U(1)_q = \sum_a q_a U(1)_a$$

is left massless; see, e.g., [167] for additional details. Charged chiral superfields may arise at the intersection of D6-brane stacks or at the intersection of D6-branes with the O6-plane. The possible representations are bifundamentals of $U(N_a) \times U(N_b)$ or two-index symmetric or antisymmetric tensor representations of $U(N_a)$. The chiral index of each such representation are counted by intersection numbers in $M$, which are presented in Table 3.

We may now study the relationship between string consistency conditions and constraints on chiral spectra. The Gauss law (D6-brane tadpole cancellation) \((A.62)\) is necessary for consistency of the string compactification, and it puts a single constraint on the homology cycles wrapped by the D6-branes and O6-plane. This equation that constraints homology may be turned into a number of equations that constrain integers representing the chiral spectra by intersecting \((A.62)\) with $\pi_b$, in which case a short computation using Table A.4 yields

$$T_b := \#b - \#\bar{b} + (N_b + 4) (\#\Box_b - \#\Box_b) + (N_b - 4) (\#\Box_b - \#\Box_b) = 0.$$ \((A.65)\)

Since this equation follows from \((A.62)\) is is necessary for D6-brane tadpole cancellation. It constrains the chiral spectrum of the 4d $U(N_b)$ gauge theory arising from $N_b$ D6-branes on $\pi_b$.

Consistency constraints on chiral spectra are commonplace in quantum field theory, usually representing an anomaly cancellation condition. Before attempting to understand \((A.65)\) from the point of view of anomalies, we would first like to recall some properties of anomalies for $SU(N)$ gauge theories in four dimensions. Consider left-handed Weyl fermions in a representation $R$ of $SU(N)$ and let

$$\chi(R) = \#R - \#\bar{R}$$ \((A.66)\)
be its chiral index. Then the \( SU(N)^3 \) anomaly cancellation takes the form
\[
\sum_{R} A_R \chi(R) = 0, \tag{A.67}
\]
where \( A_R \) is the anomaly coefficient associated to the representation \( R \). For the fundamental, \( \fund \), and \( \antifund \) representations it is 1, \( (N + 4) \), and \( (N - 4) \), respectively.

However, for this constraint to be non-trivial, complex conjugate representations \( \bar{R} \) must exist, so that \( \chi(\bar{R}) \) can be non-zero. For \( N > 2 \), such representations exist and the constraint is non-trivial. \( SU(2) \), on the other hand, only has self-conjugate representations, and accordingly there is no \( SU(2)^3 \) anomaly or associated cancellation condition. Furthermore, \( SU(2) \) gauge theories may have so-called Witten anomalies [376], which arises from the fact that
\[
\pi_4(SU(2)) = \mathbb{Z}_2, \tag{A.68}
\]
and in particular that it is non-zero. For example, if there are an odd number of Weyl fermion doublets of \( SU(2) \) the large gauge transformation associated with \( \pi_4(SU(2)) \neq 0 \) introduces a sign into the path integral, forcing it to be zero. \( SU(N > 2) \) gauge theories in four dimensions cannot have such anomalies, since
\[
\pi_4(SU(N > 2)) = 0. \tag{A.69}
\]

In summary, cubic non-abelian anomalies of type \( SU(N)^3 \) may only exist for \( N > 2 \), whereas Witten anomalies may only exist for \( N = 2 \). Anomaly cancellation conditions in 4d \( SU(N) \) gauge theories are fundamentally different according to whether \( N = 2 \) or \( N > 2 \).

Since \( U(N_a) \) gauge theories arising from \( N_a \) D6-branes in weakly coupled Type IIA compactifications may only have (chiral) fundamental representations (that arise as a bifundamental with another group) or two index symmetric or antisymmetric tensor representations, only those representations should appear in an anomaly cancellation condition for the \( SU(N_a) \) subgroup of \( U(N_a) \). This is precisely (A.65) in the case \( N_a > 2 \). The constraint \( T_b = 0 \), which is necessary for D6-brane tadpole cancellation, ensures the absence of \( SU(N_b)^3 \) anomalies in the case \( N_b > 2 \).

On the other hand, the constraints for \( U(2) \) and \( U(1)_a \), which we call \( T_2 = 0 \) and \( T_1 = 0 \), respectively, are less obvious from a field theoretic point of view. As is often said, tadpole cancellation is stronger than anomaly cancellation [377, 378], and these constraints \( T_2 = T_1 = 0 \) are often referred to as “stringy” constraints. These stringy constraints have been studied in a number of works, e.g., [34, 44, 222, 377, 378]. A summary of the current understanding of the stringy constraints is that they are:

- Necessary but not sufficient for tadpole cancellation.
- Necessary but not sufficient for \( U(1)_a \subset U(N_a) \) anomaly cancellation. This constraint on the chiral spectrum cancels the part of the anomaly not canceled by Chern-Simons terms.
- Necessary and sufficient for \( SU(M + N_a)^3 \) anomaly cancellation if the system were to nucleate \( M \) brane anti-brane pairs, embedding \( U(N_a) \) into \( U(M + N_a) \).

Each of these pathologies would render the theory inconsistent if the conditions were violated.
Before proceeding to a discussion of phenomenological implications of these constraints, we would like to make two additional comments.

First, similar constraints also arise in other corners of the landscape. For example, in Type IIB orientifold compactifications with magnetized D7-branes and D3-branes, the D7- and D5-tadpole cancellation conditions together play the same role that the D6-brane tadpole cancellation condition plays in Type IIA compactifications. From those conditions, \((A.65)\) can be derived in a different duality frame. In heterotic compactifications similar constraints are also known to exist [379]. From the Type IIA and Type IIB constraints it is also likely that they can be lifted to strong coupling, that is to M-theory and F-theory, though the form of \((A.65)\) may be more general due to the possible addition of more representations to the spectrum at strong coupling.

The constraints can be conceptually understood using one of the deep insights from the second string revolution: gauge theories that arise in string compactifications are often carried by charged objects. If these charged objects pair produce and change the structure of the gauge theory, there may be additional constraints necessary for consistency of the gauge theory after pair production. For example, if an \(SU(2)\) gauge theory is the stable end point of a process by which an \(SU(N)\) theory on \(N+2\) D-branes collides with an \(SU(2)\) theory on 2 anti-branes, the \(SU(N)\) theory prior to the collision must also be anomaly free. This can impose constraints on the \(SU(2)\) theory that would not be present from a purely bottom up perspective. See [222, 377, 380] for additional calculations and discussions that utilize this idea.

Finally, string constraints together with supersymmetry conditions can lead to interesting finiteness results. For example, in type IIA compactifications on a toroidal orbifold the D6-brane tadpole cancellation conditions together with the supersymmetry conditions yield a finite number of possible gauge sectors [381]. Similar results also apply in type IIB compactifications with magnetized D9-branes on a number of smooth Calabi-Yau threefold elliptic fibrations [382]. It is likely that such results hold quite generally in the landscape as a result of consistency conditions and supersymmetry conditions, which is an interesting direction for future work.

**Phenomenological Implications**

The conditions \((A.65)\) constraint the chiral spectrum of \(U(N_a)\) gauge theories realized in the landscape, and as we have discussed the constraints \(T_2 = 0\) and \(T_1 = 0\) on \(U(2)\) and \(U(1)\) gauge theories are “stringy,” i.e., they do not correspond to any obvious or typical constraint in quantum field theory.

In this section we study the implications of the stringy constraints for possible embeddings of the MSSM into string theory. We will aim for brevity and clarity, since the details of this section have been discussed elsewhere [44].

Consider a bottom-up embedding of an MSSM-like model, where stacks of D-branes are engineered such that the visible sector gauge group is

\[
G = U(3)_a \times U(2)_b \times \prod_{i=1}^{N} U(1)_i,
\]

where \(SU(3)_{QCD} \subset U(3)_a\) and \(SU(2)_L \subset U(2)_b\). These are \((N+2)\)-stack models. The weak hy-
percharge arises as a linear combination
\[ Y = y_a U(1)_a + y_b U(1)_b + \sum_{i=1}^{N} y_i U(1)_i, \]  
and for models with a low number of stacks the possible hypercharge linear combinations, which are referred to as hypercharge embeddings, have been classified [34]. There are additional conditions that must be satisfied [378] for this linear combination to remain massless.

Given a fixed \( N \) and hypercharge embedding, together with the fact that chiral superfields may arise as bifundamental or two-index symmetric or antisymmetric tensor representations, possible realizations of each MSSM chiral superfield may be classified. From that classification, the possible embeddings of the MSSM into the string construction may be classified, and by direct calculation one finds that nearly all violate the \( T^2 = 0 \) or \( T^1 = 0 \) conditions.

In a number of works, including works by the authors, only models where the MSSM superfields satisfy the \( T^2 = 0 \) and \( T^1 = 0 \) conditions were studied. However, models that violate these stringy constraints are not ruled out and the associated MSSM spectra may arise as spectra in consistent string compactifications. The caveat is that they must be accompanied by other fields charged under \( G \) that also contribute to the \( T^2 \) or \( T^1 \) condition, so that
\[ T^2 = T^2_{MSSM} + T^2_{exotics} = 0 \quad T^1 = T^1_{MSSM} + T^1_{exotics} = 0, \]
even if \( T^2_{MSSM} \neq 0 \) or \( T^1_{MSSM} \neq 0 \). In these models, the string constraints are requiring the existence of \( G \)-charged exotics, which usually are also charged under \( G_{SM} = SU(3) \times SU(2) \times U(1) \).

In [44], a systematic study was done of all consistent quivers of this type that contain a three family MSSM spectrum with a single pair of Higgs doublets and up to five \( G \)-charged chiral superfields beyond the MSSM. There are 89964 such models, and nearly all have exotics beyond the MSSM that are required so that \( T^2_{MSSM} \neq 0 \) or \( T^1_{MSSM} \neq 0 \) is canceled; i.e., these exotics are added for string consistency.

The results are presented in Table 4. Relevant comments on the results include:

- String constraints have clear preferences for some exotic multiplets over others.
- Chiral fourth families and hyper-charge shifted fourth families occur relatively infrequently.
- The most common exotic is a standard model singlet, and it occurs far more than any other exotic. Though they are a singlet with respect to the standard model, they are always charged under anomalous \( U(1) \)'s, which dictate their couplings to other MSSM fields.
- Next most frequent is a colorless \( SU(2)_L \) triplet with \( Y = 0 \).
- From there, we a see a variety of exotic quarks and leptons, and from their multiplicities we infer that they are coming in vector-like pairs with respect to the standard model. However, by the assumptions of the analysis they are always chiral with respect to some anomalous \( U(1) \).

In summary, we have seen that string consistency conditions can place constraints on chiral spectra that force the existence of additional exotics beyond the MSSM particles. They are often quasi-chiral, which are of phenomenological interest, as discussed in the main text.
Table 4: Displayed are the standard model representation of matter additions obtained in [44], together with their total multiplicity across all three-node and four-node quivers. The third column excludes quivers involving states that would lead to fractionally-charged color singlets, which are strongly constrained by cosmology. The fourth further excludes those where the matter additions correspond to a chiral fourth generation, while the last also excludes a hypercharge-shifted fourth generation. The additions that remain after imposing these cuts make up most of the set, and are comprised of MSSM singlets, $SU(2)$ triplets with $Y = 0$, and quasichiral pairs.

References


C. P. Burgess, J. P. Conlon, L.-Y. Hung, C. H. Kom, A. Maharana, and F. Quevedo,
“Continuous Global Symmetries and Hyperweak Interactions in String Compactifications,”


[hep-th].


and stabilization of the proton,” Phys. Rev. D71 (2005) 026009,


Densities of Dark Glueballs,” Phys. Rev. D95 (2017) no.~1, 015032,


Strongly-Coupled Dark Sectors,” Phys. Rev. D95 (2017) no.~4, 043526,


[213] A. Soni, H. Xiao, and Y. Zhang, “Cosmic selection rule for the glueball dark matter relic


arxiv:0801.4015 [hep-ph].


T. Han, Z. Si, K. M. Zurek, and M. J. Strassler, “Phenomenology of hidden valleys at


[220] V. Barger, P. Langacker, M. McCaskey, M. Ramsey-Musolf, and G. Shaughnessy,


77


