Two hundred heterotic standard models on smooth Calabi-Yau threefolds
Lara B. Anderson, James Gray, Andre Lukas, and Eran Palti
Phys. Rev. D 84, 106005 — Published 7 November 2011
DOI: 10.1103/PhysRevD.84.106005
Two Hundred Heterotic Standard Models
on Smooth Calabi-Yau Threefolds

Lara B. Anderson\textsuperscript{1}, James Gray\textsuperscript{2}, Andre Lukas\textsuperscript{3}, Eran Palti\textsuperscript{4}

\textsuperscript{1}Department of Physics, University of Pennsylvania, Philadelphia, PA 19104-6395, U.S.A.
\textsuperscript{2}Arnold-Sommerfeld-Center for Theoretical Physics, Department für Physik, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80533 München, Germany
\textsuperscript{3}Rudolf Peierls Centre for Theoretical Physics, Oxford University, 1 Keble Road, Oxford, OX1 3NP, U.K.
\textsuperscript{4}Centre de Physique Theorique, Ecole Polytechnique, CNRS, 91128 Palaiseau, France,

Abstract

We construct heterotic standard models by compactifying on smooth Calabi-Yau three-folds in the presence of purely Abelian internal gauge fields. A systematic search over complete intersection Calabi-Yau manifolds with less than six Kähler parameters leads to over 200 such models which we present. Each of these models has precisely the matter spectrum of the MSSM, at least one pair of Higgs doublets, the standard model gauge group and no exotics. For about 100 of these models there are four additional U(1) symmetries which are Green-Schwarz anomalous and, hence, massive. In the remaining cases, three U(1) symmetries are anomalous while the fourth, massless one can be spontaneously broken by singlet vacuum expectation values. The presence of additional global U(1) symmetries, together with the possibility of switching on singlet vacuum expectation values, leads to a rich phenomenology which is illustrated for a particular example. Our database of standard models, which can be further enlarged by simply extending the computer-based search, allows for a detailed and systematic phenomenological analysis of string standard models, covering issues such as the structure of Yukawa couplings, R-parity violation, proton stability and neutrino masses.

andlara@physics.upenn.edu, james.gray@uni-muenchen.de, lukas@physics.ox.ac.uk, eran.palti@cpht.polytechnique.fr
1 Introduction

There is a long history of attempting to construct four dimensional theories, from smooth compactifications of the heterotic string, with a matter sector which precisely matches that of the Minimal Supersymmetric Standard Model (MSSM). Indeed, the subject of string phenomenology started in this way in the 1980s when various attempts were made to build models based upon the “standard embedding”. In that approach, the gauge bundle was taken to be the holomorphic tangent bundle, with $SU(3)$ structure group, or deformations of the tangent bundle [1]. In recent years, more general gauge configurations have been used in an attempt to achieve phenomenologically viable physics. Slope-stable bundles with $SU(n)$ structure groups (for $n = 3, 4, 5$), unrelated to the tangent bundle, have been used in the attempt to build stringy standard models [2–12]. These constructions were based upon the use of non-Abelian gauge field configurations on smooth Calabi-Yau three-folds.

In this paper we adopt a different approach to constructing standard models in smooth Calabi-Yau three-fold compactifications of heterotic string and M-theory. Instead of using the non-Abelian constructions mentioned in the preceding paragraph, we shall construct models where the gauge field configuration in the internal dimensions is simply a sum of line bundles - that is a set of $U(1)$ fluxes. This is the extremal form of the so-called “split” or reducible bundles first studied in Refs. [26, 27].

There are two key aspects to this approach that differentiate it from the traditional non-Abelian one. The first is a practical one: it is much simpler to construct, and calculate the resulting spectrum of, Abelian bundles than non-Abelian ones. As a result, an algorithmic and systematic approach to such (heterotic) string model building is relatively straightforward and can be used to analyse vast numbers of line bundle sums over Calabi-Yau manifolds. Rather than attempting to fine-tune the construction of a single example, this large data set can be scanned for realistic models, using methods of computational algebraic geometry [30–32]. This paper presents our first results from an investigation along these lines. We have systematically scanned line bundle sums on Calabi-Yau three-folds (defined as complete intersections in products of projective spaces) with Hodge number $h^{1,1}(X) \leq 5$ and have found 208 heterotic standard models. It is important to note that these models are all “global” in that they correspond to explicit Calabi-Yau threefolds and holomorphic vector bundles leading to fully consistent heterotic theories. All 208 models have the precise matter spectrum of the MSSM, at least one pair of Higgs doublets, the standard model gauge group and no exotics charged under the standard model group of any kind. The number of models constructed should be considered with the knowledge that to date, only 3 other smooth heterotic standard models have been produced in the literature [2, 5, 12].

---

1 Slope-stable bundles satisfy the Hermitian-Yang-Mills equations required for $N = 1$ supersymmetry in 4-dimensions [1].

2 Another class of models are based on non-smooth CY orbifolds, these have been shown to also allow for an appropriate massless spectrum as well as other phenomenological features [13–22]. There are also constructions based on non-geometric settings such as the free-fermionic models as studied in [23–25].

3 Similar scans for non-Abelian constructions have been started in Refs. [10–12, 28] and further results will be presented in Ref. [29].
The second key aspect of heterotic line bundle model building is related to additional $U(1)$ symmetries. We will consider line bundle sums with structure group $S(U(1)^5)$ whose commutant within $E_8$ is $SU(5) \times S(U(1)^5) \cong SU(5) \times U(1)^4$. Hence, before Wilson line breaking, our models are based on $SU(5)$ GUT theories with four additional $U(1)$ symmetries. Phenomenologically, the vector bosons associated with those $U(1)$ symmetries should of course be massive. Fortunately, there are two mechanisms to generate such masses, both within our control. The first is the Green-Schwarz mechanism: the $U(1)$ vector bosons can acquire a large mass, close to the compactification scale, due to a gauging of axion shift symmetries. For 105 of our 208 models this happens for all four $U(1)$ symmetries, so that the low-energy gauge group is precisely that of the standard model. The remaining models have three anomalous and, hence, massive $U(1)$ symmetries while the fourth Abelian gauge factor remains massless, as long as the internal bundle is a sum of line bundles. In this case, we can invoke the second mechanism, namely moving away from the split locus in bundle moduli space such that the bundle structure group becomes non-Abelian, thus removing the extra $U(1)$ from the low energy gauge group. In the effective field-theory this amounts to giving supersymmetric vacuum expectation values (VEVs) to bundle moduli fields. We have explicit control over the spectrum of such bundle moduli and can, therefore, analyse this effect in detail.

Another important physical implication, which is tied to the above discussion, is that the Green-Schwarz anomalous $U(1)$ symmetries give rise to residual $U(1)$ global symmetries in the effective theory. These global symmetries impose constraints on the possible operators present in the theory and may forbid problematic operators such as those that lead to proton decay or R-parity violation. They may also serve as Froggatt-Nielsen type symmetries to explain the patterns of observed quark and lepton masses. This interplay between $U(1)$ symmetries, their spontaneous breaking through bundle moduli VEVs, and the resulting operators in the low energy theory, leads to a rich arena for phenomenology [33, 34].

In this paper, we present the physical ideas behind our work, the database of 208 standard models, and an exploration some of the phenomenological issues by focusing on a particular example. A more comprehensive study will be presented in a forthcoming paper [35].

The plan of this paper is as follows. In the next section we briefly explain the basic model-building set-up. Section 3 reviews the Green-Schwarz mechanism and its particular implications for our models. In section 4 we describe our scanning procedure and its main results. As an illustrative example, one of our standard models is presented in Section 5. Section 6 discusses the phenomenological implications of the anomalous $U(1)$ symmetries and bundle moduli VEVs in more detail, focusing on the particular example introduced earlier. We present a brief summary and an outlook in Section 7. The data for all 208 standard models is listed in the Appendix.

## 2 Model building set-up

We consider compactifications of the $E_8 \times E_8$ heterotic string on a smooth Calabi-Yau three-fold, $X$, with a freely acting discrete symmetry, $\Gamma$. In practice, we will use complete intersection Calabi-Yau manifolds (CICYs) which are defined as the common zero locus of homogeneous polynomials in an ambient product.
of projective spaces $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$. These manifolds have been classified \cite{36,37} and their freely-acting symmetries are known \cite{38}. In the present paper, we will explore all CICYs with freely acting symmetries and Hodge number satisfying $h^{1,1}(X) \leq 5$. It turns out, all these manifolds are “favourable” in the sense that $h^{1,1}(X) = m$, so that their whole second cohomology is spanned by the restrictions of the Kähler forms, $J_i$, of the ambient projective spaces. Line bundles, $L$, on $X$, the main building blocks of our bundle construction, can hence be denoted as $L = \mathcal{O}_X(k)$, where $k$ is an $m$-dimensional integer vector such that $c_1(\mathcal{O}_X(k)) = k^i J_i$.

As mentioned earlier, on $X$ we consider vector bundles $V$ with structure group $S(U(1)^5)$, that is, sums of line bundles $V = \bigoplus_{a=1}^5 L_a$ where $L_a = \mathcal{O}_X(k_a)$, (2.1) satisfying

$$c_1(V) = \sum_{a=1}^5 c_1^a(L_a)J_i = 0 .$$

Hence, for a given three-fold, $X$, and a given symmetry, $\Gamma$, a model is specified by the $5 h^{1,1}(X)$ integers $k^a_i$. In our model scan, we will restrict ourselves to bundles, $V$, for which

$$c_2(TX) - c_2(V) = [C] , \quad [C] \text{ an effective class in } H_2(X,\mathbb{Z}) \quad (2.3)$$

which allows for an anomaly-free supersymmetric completion by addition of an appropriate number of five-branes wrapping $C$. Supersymmetry conditions on the bundle $V$ itself will be discussed in the next section.

The structure group is embedded into $E_8$ via the sub-group chain $S(U(1)^5) \subset SU(5) \subset E_8$, so that the four-dimensional gauge group, before Wilson line breaking, is the GUT group $SU(5) \times S(U(1)^5)$. In general, the low-energy theory contains the standard $SU(5)$ multiplets $10$, $\bar{5}$ (and their conjugates) and bundle moduli singlets $1$. In addition, the above multiplets are labeled by $S(U(1)^5)$ charges, which can be represented as integers vectors $q = (q_1, \ldots, q_5)$. Due to the unit determinant condition in $S(U(1)^5)$, two such charge vectors $q$ and $\tilde{q}$ have to be identified if $q - \tilde{q} \in \mathbb{Z}(1,1,1,1,1)$ and, as a result, each charge vector with five same entries corresponds to the trivial representation. This fact will be of importance later on when we discuss the $S(U(1)^5)$ invariant operators in the four-dimensional effective theory. With this notation, the matter multiplet content of the GUT group is

$$10_{e_a}, \bar{10}_{-e_a}, 5_{e_a+e_b}, 5_{-e_a-e_b}, 1_{e_a-e_b}, 1_{-e_a+e_b} , \quad (2.4)$$

where $a < b$. Here, the subscripts are $S(U(1)^5)$ charges with $e_a$ the $a^{th}$ standard unit vector in five dimensions. These multiplets are associated to particular line bundle cohomology groups, as summarised in Table 1, and their numbers can be determined by computing the dimensions of these cohomology groups. For CICYs, line bundle cohomology can be explicitly computed by applying the methods described in Refs \cite{11,12,39}. Compared to a standard $SU(5)$ GUT theory, the multiplet content of our models is split into sub-sectors, labeled by different $S(U(1)^5)$ charges. Invariance under $S(U(1)^5)$ restricts the allowed
operators in the low-energy theory and this will be of importance for the phenomenological discussion later on. In particular, we note that the bundle moduli singlets carry non-trivial $S(U(1)^5)$ charges, so operators involving these singlets are constrained as well. This leads to an interesting interplay between $S(U(1)^5)$ invariance and switching on singlet VEVs. In the language of vector bundles, non-zero singlet VEVs corresponds to moving away from the Abelian locus in bundle moduli space to a bundle with non-Abelian structure group.

The further breaking of the GUT theory to the standard model proceeds in the standard way via Wilson lines. For the bundle $V$ to descend to the quotient Calabi-Yau manifold, $X/\Gamma$, it has to be equivariant under the symmetry $\Gamma$ [40], a property which can be explicitly checked for line bundles using the methods described in Ref. [12]. Note that for an equivariant line bundle, $L$, the cohomology groups $H^i(X, L)$ form representations under the group $\Gamma$. A Wilson line on the quotient, pointing into the standard hypercharge direction then breaks the GUT group into the standard model group times the massive $S(U(1)^5)$ symmetry. Let us consider a standard model multiplet with Wilson line representation $R_W$ which originates from a GUT multiplet with associated line bundle, $L$. The number of these multiplets can be computed from the $\Gamma$ invariant part of $H^1(X, L) \otimes R_W$. In essence, once the GUT multiplet content is known, computing the particle content after Wilson line breaking is a matter of applying representation theory of the finite group $\Gamma$.

### Table 1: Multiplet content, charges and associated line bundles of the $SU(5) \times S(U(1)^5)$ GUT theory. The indices $a, b, \ldots$ are in the range $1, \ldots, 5$ and $e_a$ denotes the standard five-dimensional unit vector in the $a$th direction. The number of each type of multiplet is obtained from the first cohomology, $H^1(X, L)$, of the associated line bundle $L$.

<table>
<thead>
<tr>
<th>multiplet $10_{e_a}$</th>
<th>$S(U(1)^5)$ charge $e_a$</th>
<th>associated line bundle $L_a$</th>
<th>contained in $V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10_{-e_a}$</td>
<td>$-e_a$</td>
<td>$L_a^*$</td>
<td>$V^*$</td>
</tr>
<tr>
<td>$5_{e_a+e_b}$</td>
<td>$e_a + e_b$</td>
<td>$L_a \otimes L_b$</td>
<td>$\wedge^2 V$</td>
</tr>
<tr>
<td>$5_{-e_a-e_b}$</td>
<td>$-e_a - e_b$</td>
<td>$L_a^* \otimes L_b^*$</td>
<td>$\wedge^2 V^*$</td>
</tr>
<tr>
<td>$1_{e_a-e_b}$</td>
<td>$e_a - e_b$</td>
<td>$L_a \otimes L_b^*$</td>
<td>$V \otimes V^*$</td>
</tr>
<tr>
<td>$1_{-e_a+e_b}$</td>
<td>$-e_a + e_b$</td>
<td>$L_a^* \otimes L_b$</td>
<td></td>
</tr>
</tbody>
</table>

3 Additional $U(1)$ symmetries and Green-Schwarz mechanism

We turn now to the fate of the four additional $U(1)$ symmetries in $S(U(1)^5) \cong U(1)^4$ which arise in our models. The Green-Schwarz mechanism in heterotic theories has been understood for many years (see [41] and [27, 42–44] for some recent papers on the subject). It is known that Abelian factors in the bundle
structure group give rise to a gauging of certain axion shift symmetries in the four dimensional effective theory. In our context, for each line bundle, $L^a$, in $V$, the Kähler axions, $\chi^i$, the supersymmetric partners of the Kähler moduli, $t^i$, acquire the following transformation

$$\delta \chi^i = -c^i_1(L_a)\eta_a ,$$  \hspace{1cm} (3.5)

with transformation parameter $\eta_a$. Note that, from Eq. (2.2), only four of these transformation, corresponding to the four $U(1)$ symmetries, are independent. Each such transformation leads to a D-term which schematically reads

$$D_a = \frac{\mu(L_a)}{\kappa} - \sum_I Q_{aI}|C_I|^2 .$$  \hspace{1cm} (3.6)

Here, $\kappa = d_{ijk}t^it^jt^k$ is the Kähler moduli space pre-potential with the triple intersection numbers $d_{ijk}$ of $X$ and $C_I$ are matter fields and bundle moduli with charges $Q_{aI}$ under $S(U(1)^5)$. The slope, $\mu(L_a)$, of the line bundle $L_a$ is defined as

$$\mu(L_a) = c^i_1(L_a)\kappa_i \text{ with } \kappa_i = d_{ijk}t^jt^k .$$  \hspace{1cm} (3.7)

We can now discuss the conditions on the line bundle sum $V$ arising from $N = 1$ supersymmetry. From a four-dimensional point of view, for a supersymmetric vacuum, all D-terms (3.6) must vanish. The locus in bundle moduli space where $V$ is split into a sum of line bundles corresponds to setting all VEVs of the fields $C_I$ to zero. Hence, all slopes, $\mu(L_a)$, must vanish simultaneously, somewhere in Kähler moduli space. This is, of course, the well-known condition for line bundle sums to preserve supersymmetry. For the equations $k^i_a\kappa_i = 0$ to have a non-trivial solution it must be the case that the

$$\text{(number of lin. independent } k_a) < h^{1,1}(X) .$$  \hspace{1cm} (3.8)

This implies strong model building constraints for Calabi-Yau manifolds with a small Hodge number $h^{1,1}(X)$ and explains why we were not able to find standard models on CICYs with $h^{1,1}(X) = 2, 3$.

At the split locus in bundle moduli space, the mass matrix for the $S(U(1)^5)$ vector bosons is given by

$$M_{ab} = G_{ij}c^i_1(L_a)c^j_1(L_b) ,$$  \hspace{1cm} (3.9)

where $G_{ij} = -\partial_i\partial_j \ln \kappa$ is the Kähler moduli space metric of $X$. Since $G_{ij}$ is positive definite, the number of massless $U(1)$ vector fields must equal $4 - \text{rank}(k_a^i)$ and can, hence, be easily determined from the integers $k_a^i$ which specify our models. Combining this statement with the inequality (3.8) we learn that

$$\text{(number of massless } U(1) \text{ vector fields}) > 4 - h^{1,1}(X) .$$  \hspace{1cm} (3.10)

Hence, for Calabi-Yau manifolds with $h^{1,1}(X) = 4$ at least one massless $U(1)$ vector field remains, while $h^{1,1}(X) = 5$ is the smallest Hodge number for which all $U(1)$ vector fields can receive masses from the Green-Schwarz mechanism.

---

4The equations below receive a one loop correction due to a non-trivial shift of the dilatonic and $M5$-brane axions. This has been explicitly studied in Ref. [27, 43] but will be neglected in the present context as it does not affect our discussion.
4 Searching for line bundle standard models

Our scanning procedure involves the following basic steps. For a given Calabi-Yau manifold \( X \) with freely-acting Abelian symmetry, \( \Gamma \), we generate a large number of line bundle sums, \( V = \bigoplus_{a=1}^{5} L_a \), satisfying \( c_1(V) = 0 \), each specified by an integer matrix \( k_a^i = c_1^i(L_a) \). In practice, we restrict the entries \( k_a^i \) to run in a certain finite range. In a first filtering step, we extract all line bundle sums which are supersymmetric (that is, all slope conditions \( \mu(L_a) = 0 \) can be satisfied for some Kähler parameters of \( X \)) and which satisfy (2.3). This ensures that all remaining models give rise to consistent heterotic vacua on \( X \). Subsequently, we extract all line bundle sums which are equivariant under \( \Gamma \), so that the model can be quotiented by \( \Gamma \).

The second step involves imposing physical constraints on the spectrum of the \( SU(5) \times S(U(1)^5) \) GUT theory. These conditions can be easily inferred from Table 1. First we impose that \( h^1(X, V) = 3|\Gamma| \) and \( h^1(X, V^*) = 0 \), where \( |\Gamma| \) is the order of the discrete symmetry group \( \Gamma \). This is to ensure that downstairs we have precisely three \( SU(5) \) families of \( 10 \) multiplets and no \( 10 \) anti-families. As can easily be proved, it then follows that \( h^1(X, \Lambda^2 V) - h^1(X, \Lambda^2 V^*) = 3|\Gamma| \) so that there is a downstairs chiral asymmetry of three \( 5 \) families. Secondly, we need at least one vector-like \( \tilde{5} - \bar{5} \) pair in order to retain a pair of Higgs doublets so we also require that \( h^1(X, \Lambda^2 V^*) > 0 \).

With these conditions imposed we have a model with the standard model gauge group (times four \( U(1) \) symmetries, some or all massive), three families of quarks and leptons and whatever remains from the \( \tilde{5} - \bar{5} \) pair. To increase the chance that the Higgs triplets can be removed we demand that \( h^1(L_a \otimes L_b^*) < |\Gamma| \) for all \( a < b \), so that the number of such pairs is smaller than the group order in each sector. In this case, it can be shown that for appropriate choices of equivariant structure and Wilson line, for all 208 models, the Higgs triplets can be projected out and at least one pair of Higgs doublets can be kept [35].

As a first step, the above procedure has been carried out for all CICYs with symmetries and \( h^{1,1}(X) \leq 5 \) in the standard list [36]. We recall that \( h^{1,1}(X) = 5 \) is the smallest value for which all four additional \( U(1) \) symmetries can become massive due to the Green-Schwarz mechanism, so it is sensible to scan up to this Hodge number at least. For the 6 CICYs with \( h^{1,1}(X) = 2 \) this has been done for line bundle entries in the range \(-10 \leq k_a^i \leq 10 \) and for the 12 CICYs with \( h^{1,1}(X) = 3 \) the range \(-3 \leq k_a^i \leq 3 \) has been covered. No model passing all the above tests has been found. As indicated earlier, this can be traced back to the stability constraint (3.8) which is particularly strong for low \( h^{1,1}(X) \).

The 19 CICYs with symmetries at \( h^{1,1}(X) = 4 \) have been scanned in the range \(-3 \leq k_a^i \leq 3 \) and 28 models passing all tests have been found. The scan over the 23 CICYs with \( h^{1,1}(X) = 5 \) in the range \(-2 \leq k_a^i \leq 2 \) resulted in 180 models. Altogether, we have found 208 heterotic line bundle standard models which are explicitly listed in the Appendix. For 105 of these models, all for \( h^{1,1}(X) = 5 \), all additional \( U(1) \) symmetries are Green-Schwarz anomalous and super-heavy. For the remaining models we have three anomalous, massive \( U(1) \) symmetries and one massless one. As indicated earlier, this remaining \( U(1) \) can be easily broken spontaneously by switching on singlet VEVs and, for this reason, these models have been included.

These results have been obtained from a scan over roughly \( 10^{12} \) integer matrices \( k_a^i \) generated initially. Hence, a “one in a billion” rule of thumb [45] is not too far from the truth in this part of the heterotic
vacuum space. It should be mentioned that this task has not required high performance computing but was completed (within several weeks) on a standard desktop machine. Extending to larger ranges for the \( k^i_a \) and to CICYs with larger \( h^{1,1}(X) \) is merely a question of computing power.

5 A standard model example

In order to illustrate our result and to set up a more explicit context for the subsequent phenomenological discussion, we will now present one of our 208 standard models in more detail. This will be sufficient for the main purpose of this paper which is to merely indicate the rich structure of model building possibilities. A detailed analysis for all standard models in our database will be carried out in a forthcoming paper [35].

Our example lives on the \( h^{1,1}(X) = 5 \) CICY with configuration matrix

\[
X = \begin{pmatrix} \mathbb{P}^1 & 1 & 1 & 0 & 0 \\ \mathbb{P}^1 & 0 & 0 & 0 & 2 \\ \mathbb{P}^1 & 0 & 0 & 2 & 0 \\ \mathbb{P}^1 & 2 & 0 & 0 & 0 \\ \mathbb{P}^3 & 1 & 1 & 1 & 1 \end{pmatrix}^{5,37}_{-64}
\]

defined in the ambient space \((\mathbb{P}^1)^4 \times \mathbb{P}^3\), as indicated in the first column of the configuration matrix. We denote the homogeneous coordinates of the four \( \mathbb{P}^1 \) by \( x_{ia} \), where \( i = 1, 2, 3, 4 \) and \( \alpha = 0, 1 \) and the \( \mathbb{P}^3 \) coordinates by \( y_\alpha \), where \( \alpha = 0, \ldots, 3 \). The remaining columns of the above matrix specify the multi-degrees of four homogeneous polynomials on the ambient space whose common zero locus defines the CICY, \( X \). The subscript is the Euler number and the superscripts provide the Hodge numbers \( h^{1,1}(X) \) and \( h^{2,1}(X) \), which count the number of Kähler and complex structure moduli, respectively. The second cohomology of \( X \) is spanned by the five ambient space Kähler forms \( J_i \) and the cone of allowed Kähler forms \( J = t^i J_i \) is specified by \( t^i > 0 \) for all \( i \). The triple intersection numbers of \( X \) have the following non-zero components (as well as those related by symmetry of the indices)

\[
d_{123} = d_{124} = d_{134} = d_{234} = d_{235} = 2 \\
d_{125} = d_{135} = d_{145} = d_{245} = d_{255} = d_{345} = d_{355} = 4 \\
d_{155} = d_{455} = d_{555} = 8 .
\]

The second Chern class of the tangent bundle is \( c_2(TX) = (24, 24, 24, 24, 56) \), relative to a basis of four-forms dual to the ambient space Kähler forms \( J_i \). The manifold is simply connected but can be divided by a freely acting \( \Gamma = \mathbb{Z}_2 \) symmetry which transforms the ambient space coordinates as

\[
(x_{i0}, x_{i1}) \to (-x_{i0}, x_{i1}) , \quad (y_0, y_1, y_2, y_3) \to (-y_0, -y_1, y_2, y_3)
\]
Our model is specified by the sum of line bundles

\[
V = \bigoplus_{a=1}^{5} L_a = \mathcal{O}_X(1, 0, 0, -1, 0) \oplus \mathcal{O}_X(1, -1, -2, 0, 1) \oplus \mathcal{O}_X(0, 1, 1, 1, -1) \oplus \mathcal{O}_X(0, -1, 1, 0, 0) \oplus \mathcal{O}_X(-2, 1, 0, 0, 0) .
\] (5.14)

This bundle satisfies \(c_1(V) = 0\) and \((2.3)\). In addition, using the above intersection numbers, it can be verified that the slope conditions \(\mu(L_a) = 0\) can be simultaneously satisfied at a locus in the Kähler cone of \(X\). It can also be verified that \(V\) is \(Z_2\) equivariant and, hence, descends to a bundle on the “downstairs” quotient space \(X/Z_2\). The bundle \((5.14)\) has four linearly independent Chern classes \(c_1(L_a) = k_a\). From our earlier discussion this means that all four additional \(U(1)\) symmetries are Green-Schwarz anomalous and, hence, massive. Consequently, the downstairs gauge group is precisely the standard model gauge group.

The non-vanishing cohomology groups of the constituent line bundles \(L_a\) are given by

\[
h^1(X, L_2) = 4 , \quad h^1(X, L_5) = 2 .
\] (5.15)

We recall from Table 1 that the cohomology groups \(H^1(X, L_a)\) count the number of GUT multiplets \(10_{e_a}\). Hence, after dividing by the symmetry order, \(|\Gamma| = 2\), this leads to three multiplets, \(10_{e_2}, 10_{e_2}, 10_{e_5}\), in the downstairs spectrum.

The non-vanishing first cohomology groups of tensor products \(L_a \otimes L_b\) and \(L_a^* \otimes L_b^*\) are

\[
h^1(X, L_2 \otimes L_4) = 4 , \quad h^1(X, L_4 \otimes L_5) = 2 , \quad h^1(X, L_2 \otimes L_5) = 1 , \quad h^1(X, L_2^* \otimes L_5^*) = 1 .
\] (5.16)

From Table 1, the cohomology groups \(H^1(X, L_a \otimes L_b)\) and \(H^1(X, L_a^* \otimes L_b^*)\) count the number of \(\bar{5}_{e_a+e_b}\) and \(5_{-e_a-e_b}\) GUT multiplets, respectively. This means downstairs we have three multiplets, \(\bar{5}_{e_2+e_4}, 5_{e_2+e_4}, \bar{5}_{e_4+e_5}\) plus whatever remains from the vector-like pair of \(\bar{5}_{e_2+e_4}\) and \(5_{-e_2-e_4}\) multiplets after Wilson line breaking. It turns out, in line with general arguments above, that both Higgs triplets can be projected out while the pair of Higgs doublets can be kept. As a result, the complete spectrum of multiplets charged under the standard model group is precisely that of the MSSM, as summarised in Table 2 below.

<table>
<thead>
<tr>
<th>name ((S(1)^5)) charge</th>
<th>(10_1)</th>
<th>(10_2)</th>
<th>(10_3)</th>
<th>(\bar{5}_1)</th>
<th>(\bar{5}_2)</th>
<th>(\bar{5}_3)</th>
<th>(H_u)</th>
<th>(H_d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e_2)</td>
<td>(e_2)</td>
<td>(e_5)</td>
<td>(e_2+e_4)</td>
<td>(e_2+e_4)</td>
<td>(e_4+e_5)</td>
<td>(-e_2-e_5)</td>
<td>(e_2+e_5)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Charges of the standard model multiplets in our example model. Each multiplet arises with multiplicity one. For simplicity, families are denoted by \(SU(5)\) representations but should be thought of as broken up into standard model multiplets, keeping the \((S(1)^5)\) charge unchanged.
first cohomology groups in this sector are
\[ h^1(X, L_2 \otimes L_1^*) = 4, \quad h^1(X, L_5 \otimes L_1^*) = 8, \quad h^1(X, L_2 \otimes L_3^*) = 4, \quad h^1(X, L_2 \otimes L_4^*) = 12 \]
\[ h^1(X, L_5 \otimes L_2^*) = 11, \quad h^1(X, L_5 \otimes L_4^*) = 3, \quad h^1(X, L_4 \otimes L_5^*) = 6. \]

After Wilson line breaking, this gives rise to seven types of singlets, denoted by \( C_1, \ldots, C_7 \), whose charges and multiplicities are listed in Table 3.

<table>
<thead>
<tr>
<th>name ( S(U(1)^5) ) charge</th>
<th>( C_1 )</th>
<th>( C_2 )</th>
<th>( C_3 )</th>
<th>( C_4 )</th>
<th>( C_5 )</th>
<th>( C_6 )</th>
<th>( C_7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>multiplicity</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3: Charges and multiplicities for the seven types of bundle moduli singlets in our example model.

To summarise, our example model has the exact spectrum and gauge group of the MSSM, plus seven types of bundle moduli fields which are singlets under the standard model group. All those fields carry charges under the remnant global \( S(U(1)^5) \) symmetry which constrains the four-dimensional effective theory. The phenomenology resulting from the interplay between this global symmetry and switching on VEVs for the singlet fields will be discussed in the next section.

6 Residual symmetries and singlet VEVs

In the previous section we presented an example from our standard model database which has exactly the matter spectrum of the MSSM along with some gauge singlet fields. In this model, all four additional \( U(1) \) symmetries are Green-Schwarz anomalous, so that their associated gauge bosons are super-heavy and, hence, absent from the low-energy theory. However, they leave behind global \( U(1) \) symmetries (see [33,34] for recent explorations of such symmetries in heterotic theories) which allow us to constrain the operator spectrum of the theory [34] and push the phenomenological study beyond the mere computation of the spectrum. Similar considerations apply to the other standard models in our database. In this section, we would like to discuss some of these phenomenological issues in general and illustrate our points within the context of the example model. A systematic study for all models will be presented in a forthcoming paper [35]. We also note that the themes presented in this section are recurrent within the F-theory GUT literature, see for example [46–50].

Before proceeding it is important to note that although we refer to the \( U(1) \) symmetries as global, since they are fundamentally gauge symmetries, there is no pseudo-Goldstone boson associated to their spontaneous breaking by the GUT singlet vevs. Rather the would be flat mode is a combination of the GUT singlets and the closed-string axion which gets eaten by the \( U(1) \). The remaining gauge-neutral combination can possibly remain a flat direction perturbatively but will be lifted by non-perturbative
operators which realise the gauge symmetry non-linearly. Whether all flat directions can be lifted in this way is a question of moduli stabilisation which we do not address here.

The study of allowed operators in the theory involves finding $S(U(1)^5)$ invariant field combinations. We recall that $S(U(1)^5)$ charges are labeled by integer vectors $\mathbf{q} = (q_1, \ldots, q_5)$ and, as a result of the determinant one condition in $S(U(1)^5)$, two such integer vectors $\mathbf{q}$ and $\tilde{\mathbf{q}}$ have to be identified if $\mathbf{q} - \tilde{\mathbf{q}} \in \mathbb{Z}(1,1,1,1,1)$. A particular operator is therefore allowed if its charge vector is entirely zero or if it is non-zero but with all entries equal. For our example, the explicit charge vectors of the MSSM fields and the seven singlet fields $C_I$ are given in Tables 2 and 3. We note that these charges are not flavour-universal, a feature which is generic for heterotic line bundle models. In our analysis, we also allow the singlets $C_I$ to develop a VEV $\epsilon_I$ which we denote by

$$\epsilon_I = \langle C_I \rangle .$$

As a result, the allowed terms involve higher dimension operators with singlet insertions - much like in the Froggatt-Nielsen setup [54]. As mentioned earlier, $S(U(1)^5)$ gauge bosons which did not receive a mass from the Green-Schwarz mechanism can become massive due to the spontaneous breaking induced by these VEVs. This is the reason why we have included such models in our list of 208 standard models given in the Appendix. In the following, we will frequently write down operators in terms of $SU(5)$ GUT multiplets, for simplicity. This is appropriate because every standard model field within a given $SU(5)$ multiplet carries the same $S(U(1)^5)$ charge. However, we should keep in mind that, even though we use the language of $SU(5)$ GUTs, the subsequent discussion applies to heterotic standard models.

It is important to note that an operator allowed by the $S(U(1)^5)$ symmetries is not necessarily present in the theory - this would require further calculations to determine [51]. In particular, the theory might have further discrete symmetries which forbid some operators allowed by $S(U(1)^5)$. However, an $S(U(1)^5)$ non-invariant operator is definitely forbidden at the perturbative level. It can still be generated by non-perturbative effects but one would expect such a contribution to be suppressed.\(^7\)

### 6.1 Proton decay

One of the strongest constraints on supersymmetric theories comes from dimension four proton decay, induced by superpotential operators of the form $\bar{5}\bar{5}10$ with matter multiplets $\bar{5}$ and $10$. In our context, these operators can be written as $\bar{5}e_a + e_b\bar{5}e_c + e_d 10e_f$ and, hence, have a total $S(U(1)^5)$ charge $e_a + e_b + e_c + e_d + e_f$. Such an operator is allowed precisely if all five charge vectors involved are different in which case the total charge is $(1,1,1,1,1)$. Whether this happens depends on the precise charges of the matter fields and has to be analysed in detail for each of our standard models. For our example, the matter field

\(^5\)Note that this shows that the approach adopted in [47] within the F-theory framework of allowing different families to come from different matter curves is in fact rather generic.

\(^6\)So long as this VEV remains small compared to the compactification scale, we can define a valid perturbative theory near the Abelian locus in moduli space. For more details on the mass scales associated to these VEVs, see [44].

\(^7\)Note that the $U(1)$ symmetries have discrete subgroups that are preserved even non-perturbatively such that the group $\mathbb{Z}_k$ is determined by the charge of the axion participating in the Green-Schwarz mass for the $U(1)$ [55, 56].
charges in Table 2 show that all such operators are forbidden and, hence, this particular model is safe from dimension four proton decay at the Abelian split locus. What happens if we move away from this locus by switching on singlet VEVs $\epsilon_I$? In this case, we have to worry about re-creating such operators by singlet VEV insertions. Again this is a matter of detailed analysis for each given model, but for our example model the singlet charges in Table 3 show that they are never re-created for any number of singlet insertions. Our example model is therefore safe from dimension four proton decay in at least a neighbourhood of the Abelian locus in bundle moduli space.

A less-constrained but nevertheless important effect is dimension five proton decay, induced by operator of the form $\bar{5}_e + e_b 10_e + 10_{e_d} e_f$ with total charge $e_a + e_b + e_c + e_d + e_f$. For our example, such operators are forbidden, as the $S(U(1)^5)$ charges in Table 2 show and, from the singlet charges in Table 3, they are not re-created by singlet insertions.

The above results regarding proton decay are promising. However within our models, forbidding proton decay using the $S(U(1)^5)$ symmetry comes at a price. From the neutrality of the Yukawa couplings in the MSSM, it is easy to show that the only $U(1)$ symmetry that can forbid dimension five proton decay is one that is not vector-like for the up- and down-Higgs. Such a symmetry is often referred to as a Pecci-Quinn symmetry, $U(1)_{PQ}$. In our example, the Higgs pair is indeed vector-like and so there is no $U(1)_{PQ}$. The reason for the absence of dimension five proton decay in this model is that, as discussed below, the down-type Yukawa couplings are forbidden by $S(U(1)^5)$ and, hence, the standard MSSM reasoning based on the presence of these couplings does not apply. Of course, this may not be a real problem as the down-type Yukawa couplings may be generated by non-perturbative effects. Such non-perturbative effects may or may not re-introduce proton decay. Whether or not this occurs can be decided at the present level of sophistication, relying on the information provided by the $S(U(1)^5)$ symmetry, by writing down the relevant gauge invariant non-perturbative contributions to the theory given the axion transformations (3.5). As with the perturbative terms being discussed in this section, whether or not such terms actually appear in the theory, as opposed to simply being allowed by gauge invariance, requires more detailed calculation to determine.

In fact, we find that, under fairly general assumptions, the issue discussed in the proceeding paragraph is generic in heterotic line bundle standard models. Assuming that the low-energy spectrum does not contain exotic states, such as Higgs triplets, Higgs pairs are always vector-like under $S(U(1)^5)$ and, hence, there is no $U(1)_{PQ}$ symmetry. The underlying model-building reasons for this will be discussed in Ref. [35]. Here, we present a more intuitive argument which follows from anomaly cancellation. The key observation is that, since the Green-Schwarz couplings only depend on the gauge field-strength, the GUT-breaking Wilson-line cannot affect Green-Schwarz anomaly cancellation. Considering the mixed anomalies of two standard model gauge factors with one of the additional $U(1)$ symmetries, together with the MSSM matter spectrum, these can only match the GUT anomalies if the Higgs fields are vector like under the
U(1) symmetry. Consequently, there is either no \(U(1)_{PQ}\) symmetry or the theory contains exotic matter fields \(^8\).

### 6.2 R-parity violation

There is a set of superpotential operators which violate the MSSM R-parity and which lead to too large neutrino masses, namely operators of the form \(5^H u_\alpha - e_b \bar{5}_c + e_d\) with \(S(U(1)^5)\) charge \(-e_a - e_b + e_c + e_d\).

For our example, an inspection of the charges in Table 2 shows that these operators are forbidden. This is consistent with our cohomology calculation which shows that, at the Abelian split locus, the three \(\bar{5}\) matter multiplets and the up-Higgs are indeed massless. However, the dimension four operator \(C_3 \bar{5}_3 H_u\) is allowed so it is possible to induce some of these R-parity violating terms by switching on a VEV for \(C_3\). To be safe we have to demand that \(\epsilon_3 = \langle C_3 \rangle = 0\) and this is sufficient to remove all similar operators with any number of singlet insertions.

### 6.3 \(\mu\)-term

A related discussion applies to the \(\mu\)-term, \(\mu H_u H_d\). As we have argued above, for our models Higgs doublets come in vector-like pairs under the \(S(U(1)^5)\) symmetry. Consequently, the \(\mu\)-term is allowed by \(S(U(1)^5)\). However, as the cohomology calculation shows, all our 208 standard models have at least one massless Higgs pair at the Abelian locus in bundle moduli space. Hence, for all these models, the \(\mu\)-term is absent from the superpotential for reasons unrelated to the \(S(U(1)^5)\) symmetry. What happens when we move away from the Abelian locus by switching on singlet VEVs? A quick glance at Table 3 shows that our example model has no singlets which are completely uncharged under \(S(U(1)^5)\), so dimension four terms of the form \(C_I H_u H_d\) are forbidden. In fact, this is generic for all our models. Bundle moduli with charge \(e_a - e_b\) are counted by the first cohomology of \(L_a \otimes L_b^*\). Singlets under \(S(U(1)^5)\) can only arise for \(a = b\) but in this case \(H^1(X, L_a \otimes L_a^*) = H^1(X, O_X) = 0\).

As a result, the lowest dimension at which a \(\mu\)-term can be generated is five. The relevant operators are of the form \(C_I C_J H_u H_d\) where \(C_I\) and \(C_J\) need to have opposite \(S(U(1)^5)\) charge. For sufficiently small VEVs, \(\epsilon_I, \epsilon_J\), this can provide a string-theoretical realisation of the solution to the \(\mu\)-problem proposed in Ref. [53]. In our example model such a dimension five operator, \(C_5 C_6 H_u H_d\), is allowed and, if indeed present, could give rise to a \(\mu\)-term of an acceptable size provided the product \(\epsilon_5 \epsilon_6\) is sufficiently small. A small value for this product is independently suggested by the pattern of up-type Yukawa couplings discussed below.

---

\(^8\)There is a very similar story in F-theory, for which we note our Wilson-line argument above also applies, in the case of hypercharge flux doublet-triplet splitting [48–50, 52]. Also note that this argument applies to an unbroken \(U(1)_{PQ}\) and can be evaded by having an approximate symmetry, i.e. breaking it well below the cutoff scale.
6.4 Yukawa couplings

Three possible types of contributions to the (superpotential) Yukawa coupling arise in our models. First we have regular dimension four Yukawa couplings. In the up sector they are of the form

\[ 5^H u^a e_e + 10 e_e^c e_e^d \]

and allowed provided \( e_a + e_b = e_c + e_d \). The down sector Yukawa couplings, \( \bar{5}^H d^a e_e + \bar{5} e_e^c + \bar{5} e_e^d e_f \), are allowed if \( e_a + e_b + e_c + e_d + e_f = (1, 1, 1, 1, 1) \). As we have mentioned earlier, the \( SU(1)^5 \) symmetry is not flavour-universal, so this generates a pattern of order one entries in the Yukawa matrices. Further contributions, proportional to the VEVs \( \epsilon_I \) or products thereof, can be generated by vacuum insertions once singlet VEVs are switched on. This amounts to a string-theoretical realisation of a Froggatt-Nielsen [54] type model for fermion masses. Finally, we may have non-perturbative contributions. Here, we will only consider the first two types of effects explicitly and we stress that they can be straightforwardly analysed for all our standard models.

However, when discussing the results, we should keep in mind that non-perturbative corrections to Yukawa couplings are rather common in string theory and provide a possible mechanism to generate small fermion masses. It is, therefore, not absolutely necessary to explain the full structure of Yukawa couplings from a Froggatt-Nielsen approach based on the \( SU(1)^5 \) symmetry. However, we should certainly require that the top Yukawa coupling is generated perturbatively at order one.

For our example model, the charges in Table (2) show that, in the absence of singlet VEVs, the up-type Yukawa matrix has rank two while the down-type Yukawa matrix vanishes identically. Switching on VEVs \( \epsilon_5 = \langle C_5 \rangle \) and \( \epsilon_6 = \langle C_6 \rangle \) the Yukawa matrices take the form

\[
Y^U = \begin{pmatrix}
\epsilon_5 & 1 & 1 \\
1 & \epsilon_6 & \epsilon_6 \\
1 & \epsilon_6 & \epsilon_6 \\
\end{pmatrix}, \quad Y^D = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}.
\]

(6.19)

Note that order one coefficients have been omitted so that \( Y^U \) generically has rank three. The eigenvalues of \( Y^U \) are of order 1, 1 and \( \epsilon_6 \), giving two heavy generations and one potentially lighter one, depending on the position in moduli space. The down-type Yukawa couplings are vanishing and so require non-perturbative effects in order to be generated.

Generally, when giving VEVs to singlets we must ensure that supersymmetry is preserved, that is, the D-terms (3.6) and F-terms must remain zero. The D-terms form a very mild restriction as the Kähler moduli can adjust themselves to minimise the D-term potential for many choices of singlet VEVs. For our example, it is even simpler to prove the existence of VEVs compatible with supersymmetry. Since \( C_5 \) and \( C_6 \) are vector-like we can set \( \epsilon_5 = \epsilon_6 \) and keep the Kähler moduli fixed, so that the FI and matter

\[ \text{We note that, as show in [47], the group theory of } E_8 \text{ allows for an accurate recreation of the observed masses and mixing of the quarks and leptons. Note though that a more detailed treatment would also require a mechanism for the observed mass splitting of the } SU(5) \text{ multiplets for the lighter generations.} \]
field contributions to the D-term vanish independently. In [35] we show that also more general vevs are supersymmetric with appropriate adjustment of the Kahler moduli.\(^{10}\)

The F-terms correspond to superpotential operators and are more difficult to study explicitly since we do not know the co-efficients of these operators. Including a possible vev for \(C_7\) the most general superpotential compatible with gauge invariance up to quartic terms is

\[
W \supset \lambda_{ij} C_6^i C_6^j + \gamma_{ij} C_6 C_4^i C_7^j .
\]

The indices on the singlets count generations. There are no quadratic terms since we know that all the fields are massless. The cubic and quartic terms are allowed by gauge invariance and may or may not be present. An explicit analysis of F-term stability is difficult due to the unknown matrices \(\lambda_{ij}\) and \(\gamma_{ij}\). However a sufficient condition for F-term stability is that the matrices do not have maximal rank since then we can set the vevs such that

\[
\lambda_{ij} \epsilon_5^j = \gamma_{ij} \epsilon_7^j = 0 .
\]

Note that the second condition in (6.21) is not strictly necessary as we will require \(\epsilon_4 = 0\) anyway, but we include the solution with it vanishing since it has relevance for the following section in which \(C_4\) are considered as potential right-handed neutrino candidates. Apart from these solutions we can also take the simpler setup with \(\epsilon_5 = 0\), which would only affect the \(\mu\)-term discussion above and would require Kahler moduli to adjust in order to solve the D-term (we have explicitly checked that this is possible within the Kahler cone up to the constraint \(|\epsilon_6| > |\epsilon_7|\)).

It is worth noting a practical advantage originating from the \(S(U(1)^5)\) symmetry, in relation to the physical Yukawa couplings in heterotic compactifications. It is generally very difficult to calculate the structure of the kinetic terms of the matter fields and so deducing the physical Yukawa couplings from the holomorphic ones is non-trivial. The additional \(U(1)\) symmetries can be of help in this regard because they can restrict the matter field kinetic terms severely.

### 6.5 Neutrino physics

The bundle moduli serve as good candidates for right-handed neutrinos [34]. For our example model, we can consider the fields \(C_4\) as forming the right handed neutrinos. In this case we have the superpotential operators, in GUT field notation,

\[
W \supset 5_{H_u} \bar{5}_3 C_4 + \epsilon_6 5_{H_u} \bar{5}_2 C_4 + \epsilon_6 5_{H_u} \bar{5}_1 C_4 + (\epsilon_6 \epsilon_7)^2 C_4 C_4 .
\]

\(^{10}\)In order to recreate the hierarchy between the top and up quark masses, and solve the D-terms, we should take \(\epsilon_5 = \epsilon_6 \sim 10^{-6}\), which interestingly implies the \(\mu\)-term operator discussed in section 6.3 is naturally at the TeV scale. Note though that this assumes that whatever mechanism induces the Charm-Top mass splitting leaves the Up coupling unaffected.
The first three terms provide Dirac neutrinos masses while the last gives a Majorana mass to $C_4$ thereby realising the see-saw mechanism. However note that there is also a possible linear term $\epsilon_6\epsilon_7C_4$ which must vanish in some way (in the MSSM this is done using matter-parity).\textsuperscript{11}

7 Conclusions and outlook

In this paper, we have presented a database of 208 heterotic standard models based on smooth Calabi-Yau manifolds and Abelian bundles over them. All of these models have the precise matter spectrum of the MSSM, one or more pairs of Higgs doublets, the standard model gauge group with possibly one additional $U(1)$ symmetry and no exotic matter charged under the standard model of any kind. For 105 of these models, there is no additional $U(1)$ symmetry so that the gauge group is exactly the standard model group. For the remaining models this $U(1)$ can be spontaneously broken by switching on singlet VEVs. We have presented an example model from our database with the exact gauge group and spectrum of the MSSM in more detail.

An interesting additional feature of our heterotic line bundle models is the presence of a global, flavour non-universal $S(U(1)^5) \cong U(1)^4$ symmetry which restricts the structure of the four-dimensional effective theory. Standard model fields as well as bundle moduli singlets are charged under $S(U(1)^5)$. The interplay between this symmetry and switching on singlet VEVs, thereby moving away from a purely Abelian bundle, provides a rich phenomenological setting for issues such as proton stability, R-parity violation, the $\mu$-problem and fermion masses. We have discussed some of these issues and have illustrated them with our example. It turns out, in this model, that the $S(U(1)^5)$ symmetry stabilises the proton, allows for an order one top Yukawa coupling, facilitates a possible solution to the $\mu$-problem and may provide a realisation of the see-saw mechanism for neutrino masses. However the down-type Yukawa couplings vanished perturbatively for this example and, as a further goal, it is important to search for models where such $U(1)$ flavour non-universal symmetries can accomodate both proton stability and a realistic Yukawa sector for both up and down.

We believe that our results raise the phenomenology of heterotic Calabi-Yau compactifications to a new level. Phenomenological problems beyond the calculation of the spectrum can now be addressed within a sizable class of quasi-realistic explicit models, rather than for a small number of individual models which are likely to fail more sophisticated phenomenological requirements. Such a systematic phenomenological analysis, for the standard models presented here, will be carried out in a forthcoming paper [35].

Our work can be extended in a variety of ways. Scans over CICYs with Hodge numbers $h^{1,1} > 5$ and larger ranges of bundles are underway and are likely to lead to more standard models. It would be interesting to perform a similar scan for heterotic line bundle models on Calabi-Yau hypersurfaces in toric varieties, as classified in Ref. [57,58], although this requires developing a number of technical tools [59,60].

\textsuperscript{11}Note that the numerical values required for the vevs $\epsilon_6$ and $\epsilon_7$ for the quark masses and neutrino majorana masses are in tension.
Acknowledgments

E. P. would like to thank Angel Uranga for correspondence and the University of Oxford for hospitality during the completion of the work. A. L. is supported by the EC 6th Framework Programme MRTN-CT-2004-503369 and would like to thank Ecole Polytechnique for hospitality while part of this work was being carried out. The work of EP was supported by the European ERC Advanced Grant 226371 MassTeV and the PITN contract PITN-GA-2009-237920. L. A. is supported in part by the DOE under Grant No. DE-AC02-76-ER-03071. J. G. would like to acknowledge support by the NSF-Microsoft grant NSF/CCF-1048082.

A Line bundle standard models on $h^{1,1}(X) = 4, 5$ CICYs

In this Appendix we provide tables with all 208 line bundle standard models which we have found on CICYs with $h^{1,1}(X) = 4$ and $h^{2,1}(X) = 5$. The scan has been performed over all line bundle sums $V = \bigoplus_{i=1}^{5} \mathcal{O}_X(k_i)$ with entries in the range $-3 \leq k_i \leq 3$ for $h^{1,1}(X) = 4$ and $-2 \leq k_i \leq 2$ for $h^{1,1}(X) = 5$. The methodology and the general results of this scan have already been described in Section 4.

The notation in the tables is as follows. The first row contains information about the CICY, namely the CICY identifier (that is, its position in the standard CICY list [36]), the standard configuration matrix with the Euler number as sub-script and $h^{1,1}, h^{2,1}(X)$ as super-scripts and the freely acting symmetry by which the model is divided. Each subsequent table entry specifies a line bundle sum by providing the five vectors $k_i$. As explained in Section 3, the number of massless $U(1)$ symmetries at the Abelian locus in bundle moduli space is given by $4$ minus the number of linearly independent vectors $k_i$ and can, hence, be directly read off from the data provided here.

| CICY 6784: | $\begin{pmatrix} p_1 & 1 & 0 & 0 & 0 & 0 \\ p_2 & 2 & 0 & 0 & 0 & 0 \\ p_3 & 1 & 1 & 2 & 0 & 0 \\ \end{pmatrix}$ |
| Z$_2 \times$ Z$_2$ |
| (3,2,-1)(1,-1,0,0)(-1,0,1,0)(-1,0,1,0)(-2,1,0,1) |
| (2,1,0,-1)(0,1,-3,0)(-1,2,1,1)(-1,0,1,0) |
| (1,0,-1)(1,0,1,0)(-2,1,1,-1)(3,1,0,0) |
| (1,1,0,-1)(1,0,1,0)(-1,2,1,1)(-2,1,1,1) |
| CICY 7435: | $\begin{pmatrix} p_1 & 1 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 1 & 1 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 1 & 1 & 0 \\ p_4 & 1 & 1 & 1 & 1 & 1 & 2 \\ \end{pmatrix}$ |
| Z$_2 \times$ Z$_2$ |
| (2,1,1,-1)(2,1,-3,0)(-1,0,1,0)(-1,0,1,0)(-2,1,0,1) |
| (2,1,1,-1)(2,1,-3,0)(-1,0,1,0)(-1,0,1,0)(-2,1,0,1) |
| (1,2,1,-1)(1,0,1,0)(0,2,1,0)(-3,2,1,0) |
| (1,2,1,-1)(1,2,-3,0)(0,1,1,0)(-3,2,0,0) |
| (1,2,1,-1)(1,2,-3,0)(0,0,0,0)(-2,0,0,0) |
### CICY 7862

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)

### CICY 5256

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)

### CICY 5256

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \)

### CICY 5256

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)

### CICY 5452

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)

### CICY 5452

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)

### CICY 5256

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \)

### CICY 5452

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)

### CICY 5452

<table>
<thead>
<tr>
<th>Vector</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0, 1, 2</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>0, 1, 2</td>
</tr>
</tbody>
</table>

\( \mathbb{Z}_2 \times \mathbb{Z}_2 \)
CICY 6770: 

$$\begin{pmatrix}
  p_1^1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  p_1^2 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
  p_1^3 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
  p_1^4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
  p_1^5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{pmatrix}$$ 

$\mathbb{Z}_2 \times \mathbb{Z}_2$

$$\begin{pmatrix}
  1 & 0 & 1 & 1 & -1 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
\end{pmatrix}$$ 

$$\begin{pmatrix}
  p_1^1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  p_1^2 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
  p_1^3 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
  p_1^4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
  p_1^5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{pmatrix}$$ 

$\mathbb{Z}_2 \times \mathbb{Z}_2$

$$\begin{pmatrix}
  1 & 0 & 1 & 1 & -1 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
\end{pmatrix}$$ 

$\mathbb{Z}_2 \times \mathbb{Z}_2$

$$\begin{pmatrix}
  1 & 0 & 1 & 1 & -1 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
  1 & 0 & 1 & 1 & 0 & 1 & -2 & 0 & 1 & 0 \\
\end{pmatrix}$$ 

$\mathbb{Z}_2 \times \mathbb{Z}_2$
| CICY 6777: | \[
\begin{pmatrix}
\begin{array}{cccc}
p^1 & 1 & 1 & 0 \\
p^1 & 1 & 0 & 0 \\
p^1 & 0 & 0 & 1 \\
p^3 & 2 & 0 & 0 \\
p^4 & 1 & 1 & 1
\end{array}
\end{pmatrix}
\] | $\mathbb{Z}_2$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1,1,0,-1)(1,0,-2,1,1)(1,-2,0,1)(-1,1,1,-1)(-2,0,0,1)</td>
<td>(1,1,0,1,-1)(1,0,1,0,0)(0,1,1,0)(0,0,1,0)(-2,1,0,1)</td>
<td>(1,1,0,1,-1)(1,0,1,0,0)(0,1,1,0)(0,0,1,0)(-2,1,0,1)</td>
</tr>
<tr>
<td>(1,0,1,1)(1,0,0,0,0)(0,1,0,0,0)(-1,2,0,1)(-2,1,0,0)</td>
<td>(1,0,1,1)(1,0,1,0,0)(0,1,1,0)(-2,1,0,1)(-2,1,0,0)</td>
<td>(1,0,1,1)(1,0,1,0,0)(0,1,1,0)(-2,1,0,1)(-2,1,0,0)</td>
</tr>
<tr>
<td>(1,0,0,1,-1)(0,1,1,1,-1)(0,1,-1,0)(-2,0,1,0)</td>
<td>(1,1,1,0,-1)(0,1,1,0,0)(0,1,1,0)(-2,1,0,1)(-1,0,1,0)</td>
<td>(1,1,1,0,-1)(0,1,1,0,0)(0,1,1,0)(-2,1,0,1)(-1,0,1,0)</td>
</tr>
<tr>
<td>(1,1,1,0,-1)(0,1,1,0,0)(0,-1,2,1,1)(-1,0,1,0)</td>
<td>(1,1,1,0,-1)(0,1,1,0,0)(0,-1,2,1,1)(-1,0,1,0)</td>
<td>(1,1,1,0,-1)(0,1,1,0,0)(0,-1,2,1,1)(-1,0,1,0)</td>
</tr>
<tr>
<td>(1,-1,1,0)(0,1,1,1,-1)(0,-1,2,1,0)(-1,0,0,1)</td>
<td>(1,-1,1,0)(0,1,1,1,-1)(0,-1,2,1,0)(-1,0,0,1)</td>
<td>(1,-1,1,0)(0,1,1,1,-1)(0,-1,2,1,0)(-1,0,0,1)</td>
</tr>
</tbody>
</table>

| CICY 6890: | \[
\begin{pmatrix}
\begin{array}{cccc}
p^1 & 1 & 1 & 0 \\
p^1 & 0 & 0 & 1 \\
p^1 & 0 & 0 & 2 \\
p^4 & 1 & 1 & 1
\end{array}
\end{pmatrix}
\] | $\mathbb{Z}_2$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1,1,0,-1)(1,0,1,0,0)(-1,1,1,-1)(-2,0,1,0)</td>
<td>(1,1,0,1,-1)(1,0,1,0,0)(-1,0,1,0)(-2,0,1,0)</td>
<td>(1,1,0,1,-1)(1,0,1,0,0)(-1,0,1,0)(-2,0,1,0)</td>
</tr>
<tr>
<td>(1,1,1,0,-1)(1,0,1,0,0)(-1,1,1,-1)(-2,1,0,1)</td>
<td>(1,1,0,1,-1)(1,0,1,0,0)(-1,0,1,0)(-2,0,1,1)</td>
<td>(1,1,0,1,-1)(1,0,1,0,0)(-1,0,1,0)(-2,0,1,1)</td>
</tr>
<tr>
<td>(1,1,1,1,0)(1,0,1,0,0)(0,1,0,0,0)(-2,1,0,1)</td>
<td>(1,1,1,1,0)(1,0,1,0,0)(0,1,0,0,0)(-2,1,0,1)</td>
<td>(1,1,1,1,0)(1,0,1,0,0)(0,1,0,0,0)(-2,1,0,1)</td>
</tr>
<tr>
<td>(1,0,1,1,1)(1,0,1,0,0)(0,1,0,0,0)(-1,0,1,0)</td>
<td>(1,0,1,1,1)(1,0,1,0,0)(0,1,0,0,0)(-1,0,1,0)</td>
<td>(1,0,1,1,1)(1,0,1,0,0)(0,1,0,0,0)(-1,0,1,0)</td>
</tr>
<tr>
<td>(1,0,1,1,1)(1,0,1,1,0)(0,1,0,0,0)(-1,0,0,0)</td>
<td>(1,0,1,1,1)(1,0,1,1,0)(0,1,0,0,0)(-1,0,0,0)</td>
<td>(1,0,1,1,1)(1,0,1,1,0)(0,1,0,0,0)(-1,0,0,0)</td>
</tr>
</tbody>
</table>

| CICY 7474: | \[
\begin{pmatrix}
\begin{array}{cc}
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
one & 0
\end{array}
\end{pmatrix}
\] | $\mathbb{Z}_2 \times \mathbb{Z}_2$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,1,0,-2,1)(0,-1,2,1,0)(-1,0,1,0,0)(-1,0,0,1)</td>
<td>(1,-2,0,0,1)(0,1,-2,1,0)(0,0,1,1,0)(0,0,1,1,0)</td>
<td>(1,-2,0,0,1)(0,1,-2,1,0)(0,0,1,1,0)(0,0,1,1,0)</td>
</tr>
<tr>
<td>(1,-2,0,0,1)(0,1,-2,1,0)(0,0,1,1,0)(0,0,1,1,0)</td>
<td>(1,-2,0,0,1)(0,1,-2,1,0)(0,0,1,1,0)(0,0,1,1,0)</td>
<td>(1,-2,0,0,1)(0,1,-2,1,0)(0,0,1,1,0)(0,0,1,1,0)</td>
</tr>
</tbody>
</table>

| CICY 7487: | \[
\begin{pmatrix}
\begin{array}{cc}
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
p^1 & 1 \\
one & 0
\end{array}
\end{pmatrix}
\] | $\mathbb{Z}_2 \times \mathbb{Z}_2$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
</tr>
<tr>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
</tr>
<tr>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
</tr>
<tr>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
<td>(1,-2,0,1,0)(0,1,2,0,1)(0,0,1,1,0)(-1,0,0,1)</td>
</tr>
</tbody>
</table>

| CICY 6828: | \[
\begin{pmatrix}
\begin{array}{cccc}
p^1 & 0 & 0 & 2 \\
p^1 & 1 & 1 & 0 \\
p^3 & 1 & 1 & 2
\end{array}
\end{pmatrix}
\] | $\mathbb{Z}_2 \times \mathbb{Z}_2$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,0,2,1)(1,-3,0,0)(0,1,1,0)(-1,0,1,0)(-2,1,0,1)</td>
<td>(1,2,0,1)(1,-3,2,0)(0,1,1,0)(-1,0,1,0)(-2,1,0,1)</td>
<td>(1,2,0,1)(1,-3,2,0)(0,1,1,0)(-1,0,1,0)(-2,1,0,1)</td>
</tr>
<tr>
<td>(2,-2,3,-1)(0,1,1,0)(0,1,1,0)(-1,0,2,1)</td>
<td>(2,1,2,-1)(0,1,3,0)(-2,-1,1)(-1,0,1,0)(-1,0,1,0)</td>
<td>(2,1,2,-1)(0,1,3,0)(-2,-1,1)(-1,0,1,0)(-1,0,1,0)</td>
</tr>
</tbody>
</table>
References


